Semi-definite extended formulations and sums of squares

Lecture notes for the Summer School on Combinatorial Optimization, Bonn, 2015

(Preliminary draft)

James R. Lee
University of Washington

1 Extended formulations and lifts of polytopes

1.1 Polytopes and inequalities

A *d-dimensional convex polytope* $P \subseteq \mathbb{R}^d$ is the convex hull of a finite set of points in $\mathbb{R}^d$. Equivalently, it is a compact set defined by a family of linear inequalities

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

for some matrix $A \in \mathbb{R}^{m \times d}$.

Let us define a measure of complexity for P: Define $\gamma(P)$ to be the smallest number m such that for some $C \in \mathbb{R}^{s \times d}$, $y \in \mathbb{R}^s$, $A \in \mathbb{R}^{m \times d}$, $b \in \mathbb{R}^m$, we have

$$P = \{x \in \mathbb{R}^d : Cx = y \text{ and } Ax \leqslant b\}.$$

In other words, this is the minimum number of *inequalities* needed to describe P. If P is full-dimensional, then this is precisely the number of *facets* of P (a facet is a maximal proper face of P).

Thinking of $\gamma(P)$ as a measure of complexity makes sense from the point of view of optimization: Interior point methods can efficiently optimize linear functions over P (to arbitrary accuracy) in time that is polynomial in $\gamma(P)$.

1.2 Lifts of polytopes

Many simple polytopes require a large number of inequalities to describe. For instance, the *cross-polytope*

$$C_d = \{x \in \mathbb{R}^d : \|x\|_1 \leqslant 1\} = \{x \in \mathbb{R}^d : \pm x_1 \pm x_2 \cdots \pm x_d \leqslant 1\}$$

has $\gamma(C_d) = 2^d$. On the other hand, C_d is the *projection* of the polytope

$$Q_d = \left\{(x, y) \in \mathbb{R}^{2d} : \sum_{i=1}^{n} y_i = 1,\ -y_i \leqslant x_i \leqslant y_i\ \forall i\right\}$$

onto the x coordinates, and manifestly, $\gamma(Q_d) \leqslant 2d$. Thus C_d is the (linear) shadow of a much simpler polytope in a higher dimension.

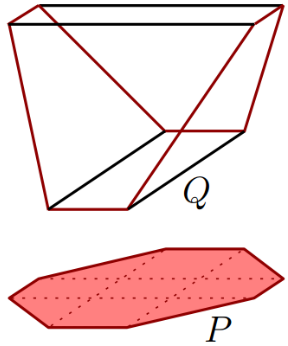

Figure 1: A lift Q of a polytope P. [Source: Fiorini, Rothvoss, and Tiwary]

A polytope Q is called a *lift* of the polytope P if P is the image of Q under a linear projection. Again, from an optimization stand point, lifts are important: If we can optimize linear functionals over Q, then we can optimize linear functionals over P. Define now $\bar{\gamma}(P)$ to be the minimal value of $\gamma(Q)$ over all lifts Q of P. (The value $\bar{\gamma}(P)$ is sometimes called the *(linear) extension complexity* of P.)

1.2.1 The permutahedron

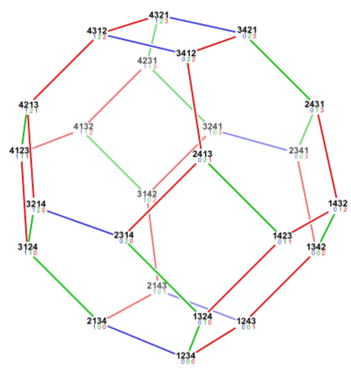

Figure 2: The permutahedron of order 4. [Source: Wikipedia]

Here is a somewhat more interesting family of examples where lifts reduce complexity. The *permutahedron* $\Pi_n \subseteq \mathbb{R}^n$ is the convex hull of the vectors $(i_1, i_2, \ldots, i_n)$ where $\{i_1, \ldots, i_n\} = \{1, \ldots, n\}$. It is known that $\gamma(\Pi_n) = 2^n - 2$.

Let $B_n \subseteq \mathbb{R}^{n^2}$ denote the convex hull of the $n \times n$ permutation matrices. The Birkhoff-von Neumann theorem tells us that B_n is precisely the set of doubly stochastic matrices, thus $\gamma(B_n) \leqslant n^2$ (corresponding to the non-negativity constraints on each entry).

Observe that Π_n is the linear image of B_n under the map $A \mapsto A(1, 2, \ldots, n)^T$, i.e. we multiply a matrix $A \in B_n$ on the right by the column vector $(1, 2, \ldots, n)$. Thus B_n is a lift of Π_n, and we conclude that $\bar{\gamma}(\Pi_n) \leqslant n^2 \ll \gamma(\Pi_n)$.

1.2.2 The cut polytope

If $P \neq NP$, there are certain combinatorial polytopes we should not be able to optimize over efficiently. A central example is the *cut polytope:* $\mathrm{CUT}_n \subseteq \mathbb{R}^{n^2}$ is the convex hull of all matrices of the form $(A_S)_{ij} = |\mathbf{1}_S(i) - \mathbf{1}_S(j)|$ for some subset $S \subseteq \{1, \dots, n\}$. Here, $\mathbf{1}_S$ denotes the characteristic function of S.

Note that the MAX-CUT problem on a graph $G = (V, E)$ can be encoded in the following way: Let $W_{ij} = 1$ if $\{i, j\} \in E$ and $W_{ij} = 0$ otherwise. Then the value of the maximum cut in G is precisely the maximum of $\langle W, A \rangle$ for $A \in \mathrm{CUT}_n$. Accordingly, we should expect that $\bar{\gamma}(\mathrm{CUT}_n)$ cannot be bounded by any polynomial in n (lest we violate a basic tenet of complexity theory).

1.2.3 Exercises

Exercise 1.1. Define the *bipartite matching polytope* $\mathrm{BM}_n \subseteq \mathbb{R}^{n^2}$ as the convex hull of all the indicator vectors of perfect matchings in the complete bipartite graph $K_{n,n}$. Show that $\bar{\gamma}(\mathrm{BM}_n) \leqslant O(n^2)$ using the standard connection between perfect matchings and s-t flows.

Exercise 1.2 (Goemans)**.** Show that for any polytope P,

$$\# \text{ faces of } P \leqslant 2^{\# \text{ facets of } P}\,.$$

Recall that a facet of P is a face of largest dimension. (Thus if $P \subseteq \mathbb{R}^n$ is full-dimensional, then a facet of P is an $(n-1)$-dimensional face.) Use this to conclude that $\bar{\gamma}(\Pi_n) \geqslant \log(n!) \geqslant \Omega(n \log n)$.

Exercise* 1.3 (Martin, 1991)**.** Define the *spanning tree polytope* $\mathrm{ST}_n \subseteq \mathbb{R}^{\binom{n}{2}}$ as the convex hull of all the indicator vectors of spanning trees in the complete graph K_n. Show that $\bar{\gamma}(\mathrm{ST}_n) \leqslant O(n^3)$ by introducing new variables $\{z_{uv,w} : u, v, w \in \{1, 2, \dots, n\}\}$ meant to represent whether the edge $\{u, v\}$ is in the spanning tree T and w is in the component of v when the edge $\{u, v\}$ is removed from T.

1.3 Non-negative matrix factorization

The key to understanding $\bar{\gamma}(\mathrm{CUT}_n)$ comes from Yannakakis' factorization theorem.

Consider a polytope $P \subseteq \mathbb{R}^d$ and let us write in two ways: As a convex hull of vertices

$$P = \mathrm{conv}\left(x_1, x_2, \dots, x_n\right)\,,$$

and as an intersection of half-spaces: For some $A \in \mathbb{R}^{m \times d}$,

$$P = \left\{x \in \mathbb{R}^d : Ax \leqslant b\right\}\,.$$

Given this pair of representations, we can define the corresponding *slack matrix* of P by

$$S_{ij} = b_i - \langle A_i, x_j \rangle \qquad i \in \{1, 2, \dots, m\}, j \in \{1, 2, \dots, n\}\,.$$

Here, $A_1, \dots, A_m$ denote the rows of A.

We need one more definition. In what follows, we will use $\mathbb{R}_+ = [0, \infty)$. If we have a non-negative matrix $M \in \mathbb{R}_+^{m \times n}$, then a *rank-$r$ non-negative factorization of M* is a factorization $M = AB$ where $A \in \mathbb{R}_+^{m \times r}$ and $B \in \mathbb{R}_+^{r \times n}$. We then define the *non-negative rank of M*, written $\mathrm{rank}_+(M)$, to be the smallest r such that M admits a rank-r non-negative factorization.

Exercise° 1.4. Show that $\text{rank}_+(M)$ is the smallest r such that $M = M_1 + \cdots + M_r$ where each M_i is a non-negative matrix satisfying $\text{rank}_+(M_i) = 1$.

The next result gives a precise connection between non-negative rank and extension complexity.

Theorem 1.5 (Yannakakis Factorization Theorem)**.** *For every polytope P, it holds that $\bar{\gamma}(P) = \text{rank}_+(S)$ for any slack matrix S of P.*

The key fact underlying this theorem is Farkas' Lemma. It asserts that if $P = \{x \in \mathbb{R}^d : Ax \leqslant b\}$, then every valid linear inequality over P can be written as a non-negative combination of the defining inequalities $\langle A_i, x\rangle \leqslant b_i$.

Exercise° 1.6. Use Farkas' Lemma to prove that if S and S' are two different slack matrices for the same polytope P, then $\text{rank}_+(S) = \text{rank}_+(S')$.

There is an interesting connection here to proof systems. The theorem says that we can interpret $\bar{\gamma}(P)$ as the minimum number of axioms so that every valid linear inequality for P can be proved using a conic (i.e., non-negative) combination of the axioms.

1.4 Slack matrices and the correlation polytope

Thus to prove a lower bound on $\bar{\gamma}(\text{CUT}_n)$, it suffices to find a valid set of linear inequalities for CUT_n and prove a lower bound on the non-negative rank of the corresponding slack matrix.

Toward this end, consider the correlation polytope $\text{CORR}_n \subseteq \mathbb{R}^{n^2}$ given by

$$\text{CORR}_n = \text{conv}\left(\{xx^T : x \in \{0,1\}^n\}\right) .$$

Exercise° 1.7. Show that for every $n \geqslant 1$, CUT_{n+1} and CORR_n are linearly isomorphic.

Now we identify a particularly interesting family of valid linear inequalities for CORR_n. (In fact, it turns out that this will also be an exhaustive list.) A *quadratic multi-linear function* on $\mathbb{R}^n$ is a function $f : \mathbb{R}^n \to \mathbb{R}$ of the form

$$f(x) = a_0 + \sum_i a_{ii}x_i + \sum_{i<j} a_{ij}x_i x_j ,$$

for some real numbers a_0 and $\{a_{ij}\}$.

Suppose f is a quadratic multi-linear function that is also non-negative on $\{0,1\}^n$. Then "$f(x) \geqslant 0 \ \forall x \in \{0,1\}^n$" can be encoded as a valid linear inequality on CORR_n. To see this, write $f(x) = \langle A, xx^T\rangle + a_0$ where $A = (a_{ij})$. (Note that $\langle\cdot,\cdot\rangle$ is intended to be the standard inner product on $\mathbb{R}^{n^2}$.)

Now let QML_n^+ be the set of all quadratic multi-linear functions that are non-negative on $\{0,1\}^n$, and consider the matrix (represented here as a function) $\mathcal{M}_n : \text{QML}_n^+ \times \{0,1\}^n \to \mathbb{R}_+$ given by

$$\mathcal{M}_n(f, x) = f(x) .$$

Then from the above discussion, $\mathcal{M}_n$ is a valid sub-matrix of some slack matrix of CORR_n. To summarize, we have the following theorem.

Theorem 1.8. *For all $n \geqslant 1$, it holds that $\bar{\gamma}(\text{CUT}_{n+1}) \geqslant \text{rank}_+(\mathcal{M}_n)$.*

Exercise° 1.9. Show that we have equality in the above theorem: For every $n \geqslant 1$,

$$\bar{\gamma}(\mathrm{CUT}_{n+1}) = \bar{\gamma}(\mathrm{CORR}_n) = \mathrm{rank}_+(\mathcal{M}_n)\,.$$

The following result represents a breakthrough in our understanding of extension complexity.

Theorem 1.10 (Fiorini, Massar, Pokutta, Tiwari, de Wolf 2012)**.** *There is a constant $c > 0$ such that for every $n \geqslant 1$, $\bar{\gamma}(\mathrm{CUT}_n) \geqslant e^{cn}$.*

We will examine a somewhat weaker lower bound following Chan-Lee-Raghavendra-Steurer (2013) and Lee-Raghavendra-Steurer (2015). This method is only currently capable of proving that $\bar{\gamma}(\mathrm{CUT}_n) \geqslant e^{cn^{1/3}}$, but it has the advantage of being somewhat more general—it extends well to the setting of approximate lifts and spectrahedral lifts (those coming from semi-definite programs).

1.5 NMF and positivity certificates

If $r = \mathrm{rank}_+(\mathcal{M}_n)$, it means we can write

$$f(x) = \mathcal{M}_n(f, x) = \sum_{i=1}^{r} A_i(f) B_i(x) \tag{1.1}$$

for some functions $A_i : \mathrm{QML}_n^+ \to \mathbb{R}_+$ and $B_i : \{0,1\}^n \to \mathbb{R}_+$. (Here we are using a factorization $\mathcal{M}_n = AB$ where $A_{f,i} = A_i(f)$ and $B_{x,i} = B_i(x)$.)

Thus the low-rank factorization gives us a "proof system" for QML_n^+. Every $f \in \mathrm{QML}_n^+$ can be written as a conic combination of the functions $B_1, B_2, \ldots, B_r$, thereby certifying its positivity (since the B_i's are positive functions).

If we think about natural families $\mathcal{B} = \{B_i\}$ of "axioms," then since QML_n^+ is invariant under the natural action of S_n (the symmetric group on $\{1, 2, \ldots, n\}$), we might expect that our family $\mathcal{B}$ should share this invariance. Once we entertain this expectation, there are natural small families of axioms to consider: The space of non-negative k-juntas for some $k \ll n$.

A *k-junta* $b : \{0,1\}^n \to \mathbb{R}$ is a function whose value only depends on k of its input coordinates. For a subset $S \subseteq \{1, \ldots, n\}$ with $|S| = k$ and an element $z \in \{0,1\}^k$, let $q_{S,z} : \{0,1\}^n \to \{0,1\}$ denote the function given by $q_{S,z}(x) = 1$ if and only if $x|_S = z$.

We let $\mathcal{J}_k = \{q_{S,z} : |S| \leqslant k, z \in \{0,1\}^{|S|}\}$. Observe that $|\mathcal{J}_k| \leqslant O(n^k)$. Let us also define $\mathrm{cone}(\mathcal{J}_k)$ as the set of all conic combinations of functions in $\mathcal{J}_k$.

Exercise° 1.11. Show that $\mathrm{cone}(\mathcal{J}_k)$ is precisely the set of all non-negative combinations of non-negative k-juntas.

If it were true that $\mathrm{QML}_n^+ \subseteq \mathcal{J}_k$ for some k, we could immediately conclude that $\mathrm{rank}_+(\mathcal{M}_n) \leqslant |\mathcal{J}_k| \leqslant O(n^k)$ by writing $\mathcal{M}_n$ in the form (1.1) where now $\{B_i\}$ ranges over the elements of $\mathcal{J}_k$ and $\{A_i(f)\}$ gives the corresponding non-negative coefficients that follow from $f \in \mathcal{J}_k$.

1.6 Junta degree and the dual cone

Clearly $\mathrm{QML}_n^+ \subseteq \mathrm{cone}(\mathcal{J}_n)$. We will now see that juntas cannot a smaller set of axioms for QML_n^+.

Theorem 1.12. *Consider the function* $f : \{0,1\}^n \to \mathbb{R}_+$ *given by* $f(x) = (x_1 + x_2 + \cdots + x_n - 1)^2$. *Then* $f \notin \mathrm{cone}(\mathcal{J}_{n-1})$.

Proof. Suppose we write $f = \sum_{i=1}^N q_i$ where each q_i is non-negative. Clearly if $\sum_{i=1}^n x_i = 1$, then $f(x_1, \ldots, x_n) = 0$, hence $q_i(x_1, \ldots, x_n) = 0$ for every i. But if $q_i \in \mathcal{J}_{n-1}$, then there is some coordinate on which it does not depend. Without loss, suppose it is the first coordinate. Then $0 = q_i(1, 0, \ldots, 0) = q_i(0, 0, \ldots, 0)$. But $f(0, 0, \ldots, 0) = 1$. We conclude that $f \notin \mathcal{J}_{n-1}$. □

Let us now prove this in a more roundabout way by introducing a few definitions. First, for $f : \{0,1\}^n \to \mathbb{R}_+$, define the *junta degree of* f to be

$$\deg_J(f) = \min\{k : f \in \mathrm{cone}(\mathcal{J}_k)\} .$$

Since every f is an n-junta, we have $\deg_J(f) \leqslant n$.

Now because $\{f : \deg_J(f) \leqslant k\}$ is a cone (spanned by $\mathcal{J}_k$), there is a universal way of proving that $\deg_J(f) > k$. Say that a functional $\varphi : \{0,1\}^n \to \mathbb{R}$ is *k-locally positive* if for all $|S| \leqslant k$ and $z \in \{0,1\}^{|S|}$, we have

$$\sum_{x \in \{0,1\}^n} \varphi(x) q_{S,z}(x) \geqslant 0 .$$

These are precisely the linear functionals separating a function f from $\mathrm{cone}(\mathcal{J}_k)$: We have $\deg_J(f) > k$ if and only if there is a k-locally positive functional φ such that $\sum_{x \in \{0,1\}^n} \varphi(x) f(x) < 0$. Now we are ready to prove Theorem 1.12 in a different way.

Second proof of Theorem 1.12. We will use an appropriate k-locally positive functional. Define

$$\varphi(x) = \begin{cases} -1 & |x| = 0 \\ 1 & |x| = 1 \\ 0 & |x| > 1 , \end{cases}$$

where $|x|$ denotes the hamming weight of $x \in \{0,1\}^n$.

Recall the the function f from the statement of the theorem and observe that by opening up the square, we have

$$\begin{aligned} \sum_{x \in \{0,1\}^n} \varphi(x) f(x) &= \sum_{x \in \{0,1\}^n} \varphi(x) \left(1 - 2 \sum_i x_i + \sum_i x_i^2 + 2 \sum_{i \neq j} x_i x_j \right) \\ &= \sum_{x \in \{0,1\}^n} \varphi(x) \left(1 - \sum_i x_i \right) = -1 . \end{aligned} \tag{1.2}$$

Now consider some $S \subseteq \{1, \ldots, n\}$ with $|S| = k \leqslant n - 1$ and $z \in \{0,1\}^k$. If $z = \mathbf{0}$, then

$$\sum_{x \in \{0,1\}^n} \varphi(x) q_{S,z}(x) = -1 + 1 \cdot (n - k) \geqslant 0 .$$

If $|z| > 1$, then the sum is 0. If $|z| = 1$, then the sum is non-negative because in that case $q_{S,z}$ is only supported on non-negative values of φ. We conclude that φ is k-locally positive for $k \leqslant n - 1$. Combined with (1.2), this yields the statement of the theorem. □

Exercise 1.13. Consider the *knapsack polynomial:* For $n \geqslant 1$ odd,

$$f(x) = \left(x_1 + x_2 + \cdots + x_n - \frac{n}{2}\right)^2 - \frac{1}{4}.$$

It is straightforward to check that $f(x) \geqslant 0$ for all $x \in \{0,1\}^n$. Define an appropriate locally positive functional to show that $\deg_J(f) \geqslant \lfloor \frac{n}{2} \rfloor$.

1.7 From juntas to general factorizations

So far we have seen that we cannot achieve a low non-negative rank factorization of $\mathcal{M}_n$ using k-juntas for $k \leqslant n-1$.

Remark 1.14. If one translates this into the setting of lift-and-project systems, it says that the k-round Sherali-Adams lift of the polytope

$$P = \left\{x \in [0,1]^{n^2} : x_{ij} = x_{ji},\, x_{ij} \leqslant x_{jk} + x_{ki} \quad \forall i,j,k \in \{1,\ldots,n\}\right\}$$

does not capture CUT_n for $k \leqslant n-1$.

In the next lecture, we will show that a non-negative factorization of $\mathcal{M}_n$ would lead to a k-junta factorization with k small (which we just saw is impossible). This will yield a lower bound on $\bar{\gamma}(\mathrm{CUT}_n)$.

For now, let us state the theorem we want to prove. We first define a submatrix of $\mathcal{M}_n$. Fix some integer $m \geqslant 1$ and a function $g : \{0,1\}^m \to \mathbb{R}_+$. Now define the matrix $M_n^g : \binom{[n]}{m} \times \{0,1\}^n \to \mathbb{R}_+$ given by

$$M_n^g(S,x) = g(x|_S).$$

The matrix is indexed by subsets $S \subseteq [n]$ with $|S| = m$ and elements $x \in \{0,1\}^n$. Here, $x|_S$ represents the (ordered) restriction of x to the coordinates in S.

Theorem 1.15 (Chan-Lee-Raghavendra-Steurer 2013)**.** *For every $m \geqslant 1$ and $g : \{0,1\}^m \to \mathbb{R}_+$, there is a constant $C = C(g)$ such that for all $n \geqslant 2m$,*

$$\mathrm{rank}_+(M_n^g) \geqslant C\left(\frac{n}{\log n}\right)^{\deg_J(g)-1}.$$

Note that if $g \in \mathrm{QML}_m^+$ then M_n^g is a submatrix of $\mathcal{M}_n$. Since Theorem 1.12 furnishes a sequence of quadratic multi-linear functions $\{g_j\}$ with $\deg_J(g_j) \to \infty$, the preceding theorem tells us that $\mathrm{rank}_+(\mathcal{M}_n)$ cannot be bounded by any polynomial in n. A more technical version of the theorem is able to achieve a lower bound of $e^{cn^{1/3}}$, but it is unclear if this method can match the lower bound of Theorem 1.10. Its main advantage is that it extends to both the setting of approximation algorithms, and also to semi-definite extended formulations.

2 Entropy maximization and approximation by juntas

Our goal is now to prove Theorem 1.15.

2.1 Analytic non-negative rank

Before getting to the proof, let us discuss the situation in somewhat more generality. Consider finite sets X and Y and a matrix $M : X \times Y \to \mathbb{R}_+$. Our goal is to show that $\text{rank}_+(M)$ is large. In other words, that we cannot write

$$M(x, y) = \sum_{i=1}^{r} A_i(x) B_i(y)$$

for some small value $r \geqslant 0$ and $A_1, \ldots, A_r : X \to \mathbb{R}_+, B_1, \ldots, B_r : Y \to \mathbb{R}_+$.

It would be nice if we could argue that M cannot be too correlated with any map $(x, y) \mapsto A_i(x)B_i(y)$ and therefore r must be large. This would avoid having to argue about a subtle relationship between $\{A_i\}$ and $\{B_i\}$ for different values of i. For instance, we could try to find a functional $F : X \times Y \to \mathbb{R}$ such that $\sum_{x,y} F(x, y)M(x, y) < 0$ while $\sum_{x,y} F(x, y)A_i(x)B_i(y) \geqslant 0$ for all $i = 1, \ldots, r$.

In other words, we would like to define a convex set of "low non-negative rank" matrices and show that M is not in this set (by convex duality, this separation would always be accomplished with such a linear functional F). Note that matrices of the form $(x, y) \mapsto A_i(x)B_i(y)$ are exactly those of non-negative rank 1. But the convex hull of $\{N \in \mathbb{R}_+^{X\times Y} : \text{rank}_+(N) = 1\}$ is precisely the set of all non-negative matrices (which certainly contains M!).

Instead, let us proceed analytically. For simplicity, let us equip both X and Y with the uniform measure. Let $\mathcal{Q} = \{b : Y \to \mathbb{R}_+ \mid \|b\|_1 = 1\}$ denote the set of probability densities on Y, where we define $\|b\|_1 = \frac{1}{|Y|} \sum_{y\in Y} |b(y)|$.

Now define

$$\alpha_+(N) = \min \left\{ \alpha : \exists A \in \mathbb{R}_+^{X\times k},\ B \in \mathbb{R}_+^{k\times Y} \text{ with } N = AB, \{B_1, \ldots, B_k\} \subseteq \mathcal{Q}, \text{ and } \max_{i\in[k]} \|B_i\|_\infty \leqslant \alpha \text{ and } \sum_{i=1}^{k} \|A^{(i)}\|_\infty \leqslant \alpha \right\}.$$

Here $\{A^{(i)}\}$ are the columns of A and $\{B_i\}$ are the rows of B. Note that now k is unconstrained.

Observe that for any $c > 0$, the set $\{N : \alpha_+(N) \leqslant c\}$ is convex. To see this, given a pair $N = AB$ and $N' = A'B'$, write

$$\frac{N + N'}{2} = \begin{pmatrix} \frac{1}{2}A & \frac{1}{2}A' \end{pmatrix} \begin{pmatrix} B \\ B' \end{pmatrix},$$

witnessing the fact that $\alpha_+(\frac{1}{2}(N + N')) \leqslant \max\{\alpha_+(N), \alpha_+(N')\}$.

2.2 A truncation argument: Relating α_+ and rank_+

We will see now that low non-negative rank matrices are close to matrices with α_+ small. In standard communication complexity/discrepancy arguments, this corresponds to discarding "small rectangles."

Lemma 2.1. *For every non-negative $M \in \mathbb{R}_+^{X\times Y}$ with $\text{rank}_+(M) \leqslant r$ and every $\delta \in (0, 1)$, there is a matrix $\tilde{M} \in \mathbb{R}_+^{X\times Y}$ such that*

$$\|M - \tilde{M}\|_1 \leqslant \delta$$

and

$$\alpha_+(\tilde{M}) \leqslant \frac{r}{\delta} \|M\|_\infty .$$

Proof. Suppose that $M = AB$ with $A \in \mathbb{R}_+^{X \times r}, B \in \mathbb{R}_+^{r \times Y}$, and let us interpret this factorization in the form

$$M(x, y) = \sum_{i=1}^{r} A_i(x) B_i(y) \tag{2.1}$$

(where $\{A_i\}$ are the columns of A and $\{B_i\}$ are the rows of B). By rescaling the columns of A and the rows of B, respectively, we may assume that $\mathbb{E}[B_i] = 1$ for every $i \in [r]$.

Let $\Lambda = \{i : \|B_i\|_\infty > \tau\}$ denote the "bad set" of indices (we will choose τ momentarily). Observe that if $i \in \Lambda$, then

$$\|A_i\|_\infty \leqslant \frac{\|M\|_\infty}{\tau},$$

from the representation (2.1) and the fact that all summands are positive.

Define the matrix $\tilde{M}(x, y) = \sum_{i \notin \Lambda} A_i(x) B_i(y)$. It follows that

$$\|M - \tilde{M}\|_1 = \mathop{\mathbb{E}}_{x,y}\left[|M(x, y) - \tilde{M}(x, y)|\right] = \sum_{i \in \Lambda} \mathop{\mathbb{E}}_{x,y}[A_i(x) B_i(y)].$$

Each of the latter terms is at most $\|A_i\|_\infty \|B_i\|_1 \leqslant \frac{\|M\|_\infty}{\tau}$ and $|\Lambda| \leqslant r$, thus

$$\|M - \tilde{M}\|_1 \leqslant r \frac{\|M\|_\infty}{\tau}.$$

Next, observe that

$$\mathop{\mathbb{E}}_{y}[M(x, y)] = \sum_{i=1}^{r} A_i(x) \|B_i\|_1 = \sum_{i=1}^{r} A_i(x),$$

implying that $\|A_i\|_\infty \leqslant \|M\|_\infty$ and thus $\sum_{i=1}^{r} \|A_i\|_\infty \leqslant r\|M\|_\infty$.

Setting $\tau = r\|M\|_\infty/\delta$ yields the statement of the lemma. □

Generally, the ratio $\frac{\|M\|_\infty}{\|M\|_1}$ will be small compared to r (e.g., polynomial in n vs. super-polynomial in n). Thus from now on, we will actually prove a lower bound on $\alpha_+(M)$. One has to verify that the proof is robust enough to allow for the level of error inherent in Lemma 2.1.

2.3 Simplifying the B-side of the factorization

Returning to Theorem 1.15 and the matrix M_n^g, we will assume, for the sake of an eventual contradiction, that $\alpha_+(M_n^g) \leqslant \alpha$. Thus we can write

$$M_n^g(S, x) = \sum_{i=1}^{k} A_i(S) B_i(x), \tag{2.2}$$

where $A_i, B_i \geqslant 0$ and we have $\|B_i\|_1 = 1$ and $\|B_i\|_\infty \leqslant \alpha$ for all $i \in [k]$, and finally $\sum_{i=1}^{k} \|A_i\|_\infty \leqslant \alpha$.

We will consider the uniform measures on $\binom{[n]}{m}$ and $\{0,1\}^n$. We use $\mathbb{E}_S$ and $\mathbb{E}_x$ to denote averaging with respect to these measures.

Let $d = \deg_J(g) - 1$. From Section 1.6, we know there exists a d-locally positive functional $\varphi : \{0,1\}^n \to \mathbb{R}$ such that $\beta := \mathbb{E}_x\, \varphi(x) g(x) < 0$, and $\mathbb{E}_x\, \varphi(x) q(x) \geqslant 0$ for every d-junta q. For $S \subseteq [n]$ with $|S| = m$, let us denote $\varphi_S(x) = \varphi(x|_S)$.

Following our observations in Section 1.6, we can see if each B_i in (2.2) were a d-junta, we would have a contradiction: For some constant $\beta < 0$ (depending only on g):

$$\mathop{\mathbb{E}}_{S,x}\left[\varphi_S(x) M_n^g(S,x)\right] = \mathop{\mathbb{E}}_{y\in\{0,1\}^m} \varphi(y) g(y) = \beta < 0\,, \tag{2.3}$$

and yet

$$\mathop{\mathbb{E}}_{S,x}\left[\varphi_S(x) \sum_{i=1}^{k} A_i(S) B_i(x)\right] = \sum_{i=1}^{k} \mathop{\mathbb{E}}_{S} A_i(S) \mathop{\mathbb{E}}_{x}\left[\varphi_S(x) B_i(x)\right] \geqslant 0 \tag{2.4}$$

because $\mathbb{E}_x\left[\varphi_S(x)B_i(x)\right] = \mathbb{E}_{y\in\{0,1\}^S} \varphi(y)\, \mathbb{E}_x\left[B_i(x) \,\middle|\, x|_S = y\right] \geqslant 0$ since φ is d-locally positive and the function $y \mapsto \mathbb{E}_x\left[B_i(x) \,\middle|\, x|_S = y\right]$ is a d-junta.

Thus our goal now will be to *approximate* each B_i by a junta such that the equality in (2.2) approximately holds in a suitable sense. In fact, the argument in (2.4) gives us a hint as to what kind of approximation we want: If $\tilde{B}_i$ is our approximator for B_i, we would like that $\mathbb{E}_x[\varphi_S(x)\tilde{B}_i(x)] \geqslant \mathbb{E}_x[\varphi_S(x)B_i(x)] - \varepsilon$ for some small $\varepsilon > 0$.

2.4 High entropy functions are close to juntas

Why should we expect that $B_i : \{0,1\}^n \to \mathbb{R}_+$ can be approximated by a "simple" function? Recall that $\mathbb{E}_x B_i(x) = 1$ and $\|B_i\|_\infty \leqslant \alpha$. In particular, this implies that if we think of B_i as a probability distribution on $\{0,1\}^n$, then its Shannon entropy is very high: At least $n - \log\alpha$ bits.

Instead of using Shannon entropy, let us define the *relative entropy* with respect to the uniform measure by: Given a density $b : \{0,1\}^n \to \mathbb{R}_+$ with $\mathbb{E}_x b(x) = 1$, define

$$\mathrm{Ent}(b) = \mathop{\mathbb{E}}_{x}[b(x) \log b(x)]\,.$$

It is an exercise to check that $0 \leqslant \mathrm{Ent}(b) \leqslant n$ holds for every such b. Moreover, $\mathrm{Ent}(b) = 0$ if and only if $b = \mathbf{1}$ is the constant function equal to 1 everywhere. And $\mathrm{Ent}(b) = n$ if and only if b is supported on a single point.

Most importantly, the function $b \mapsto \mathrm{Ent}(b)$ is *convex* on the space of probability densities.

Fix $i \in [k]$ and consider the following convex optimization problem. The variables are the values $\{\tilde{B}_i(x) : x \in \{0,1\}^n\}$. The value ε is not a variable, but a parameter we will choose later.

$$\begin{aligned}
\text{minimize} \quad & \mathrm{Ent}(\tilde{B}_i) = \mathbb{E}[\tilde{B}_i \log \tilde{B}_i] \\
\text{subject to} \quad & \mathbb{E}[\tilde{B}_i] = 1 && (2.5)\\
& \tilde{B}_i(x) \geqslant 0 \quad \forall x \in \{0,1\}^n && (2.6)\\
& \mathop{\mathbb{E}}_{x}\left[\varphi_S(x)\tilde{B}_i(x)\right] \leqslant \mathop{\mathbb{E}}_{x}\left[\varphi_S(x) B_i(x)\right] + \varepsilon \quad \forall |S| = m\,. && (2.7)
\end{aligned}$$

The idea is that we attempt to find a density $\tilde{B}_i$ that satisfies the constraints we identified earlier (2.7) up to accuracy ε. To encourage "simplicity" of the approximator, we minimize its relative entropy subject to these constraints.

It turns out that if we attempt to solve this program using the correct a particularly simple form sub-gradient descent, then the solution we obtain (while not necessarily optimal!) will satisfy constraints (2.5)–(2.7), and have a special form.

Claim 2.2. There exists a function $\tilde{B}_i : \{0,1\}^n \to \mathbb{R}^n_+$ satisfying all the preceding constraints and of the form

$$\tilde{B}_i = \frac{\exp\left(\sum_{j=1}^{M} \lambda_j \varphi_{S_j}\right)}{\mathbb{E}\exp\left(\sum_{i=1}^{M} \lambda_j \varphi_{S_j}\right)}$$

such that

$$M \leqslant C(g)\frac{\log \alpha}{\varepsilon^2},$$

where $C(g)$ is some constant depending only on g.

Now each φ_{S_j} only depends on m variables (those in S_j and $|S_j| = m$), meaning that our approximator $\tilde{B}_i$ is an h-junta for

$$h \leqslant m \cdot C(g)\frac{\log \alpha}{\varepsilon^2}. \tag{2.8}$$

This doesn't seem very good! The calculation in (2.4) needs that $\tilde{B}_i$ is a d-junta, and certainly $d < m$ (since g is a function on $\{0,1\}^m$).

So far we have only used the fact that each B_i has small relative entropy. We did not use the other fact we know from our assumption $\alpha_+(M_n^g) \leqslant \alpha$, namely that $\sum_{i=1}^k \|A_i\|_\infty \leqslant \alpha$. Now we employ this assumption to drastically reduce the junta size in (2.8).

2.5 Random restriction: Using our bounds on the A-side of the factorization

Let's try to apply the logic of (2.4) to the $\tilde{B}_i$ approximators anyway. Fix some $i \in [k]$ and let J_i be the set of coordinates on which $\tilde{B}_i$ depends. Then:

$$\mathop{\mathbb{E}}_{S,x}\left[\varphi_S(x)A_i(S)\tilde{B}_i(x)\right] = \mathop{\mathbb{E}}_{S}\mathop{\mathbb{E}}_{y\in\{0,1\}^S}\varphi(y)\mathop{\mathbb{E}}_{x\in\{0,1\}^n}\left[B_i(x)\,\middle|\,x|_S = y\right]$$

Note that the map $y \mapsto \mathbb{E}_{x\in\{0,1\}^n}\left[B_i(x)\,\middle|\,x|_S = y\right]$ is a junta on $J_i \cap S$. Thus if $|J_i \cap S| \leqslant d$, then the contribution from this term is non-negative since φ is d-locally positive. But $|S| = m$ is fixed and n is growing, thus $|J_i \cap S| > d$ is quite rare! Formally,

$$\mathop{\mathbb{E}}_{S,x}\left[\varphi_S(x)A_i(S)\tilde{B}_i(x)\right] \geqslant -\|A_i\|_\infty \mathbb{P}_S[|J_i \cap S| > d] \geqslant -\|A_i\|_\infty \frac{h^d(2m)^d}{n^d}.$$

In the last estimate, we have used a simple union bound and $n \geqslant 2m$.

Putting everything together and summing over $i \in [k]$, we conclude that

$$\sum_{i=1}^{k}\mathop{\mathbb{E}}_{S,x}\left[\varphi_S(x)A_i(S)\tilde{B}_i(x)\right] \geqslant -\alpha\frac{h^d(2m)^d}{n^d}.$$

Note that by choosing n only moderately large, we will make this error term very small.

Moreover, since each $\tilde{B}_i$ satisfies (2.7), we have

$$\sum_{i=1}^{k}\mathop{\mathbb{E}}_{S,x}\left[\varphi_S(x)A_i(S)B_i(x)\right] = \sum_{i=1}^{k}\mathop{\mathbb{E}}_{S}A_i(S)\mathop{\mathbb{E}}_{x}\left[\varphi_S(x)B_i(x)\right]$$

$$\geqslant \sum_{i=1}^{k} \mathop{\mathbb{E}}_{S} A_i(S) \left(-\varepsilon + \mathop{\mathbb{E}}_{x} \left[\varphi_S(x)\tilde{B}_i(x)\right]\right)$$

$$\geqslant -\varepsilon \sum_{i=1}^{k} \|A_i\|_1 - \alpha \frac{h^d (2m)^d}{n^d} .$$

Almost there: Now observe that

$$\|g\|_1 = \mathop{\mathbb{E}}_{S,x}[M_n^g(S,x)] = \sum_{i=1}^{k} \|A_i\|_1 \|B_i\|_1 = \sum_{i=1}^{k} \|A_i\|_1 .$$

Plugging this into the preceding line yields

$$\sum_{i=1}^{k} \mathop{\mathbb{E}}_{S,x} \left[\varphi_S(x) A_i(S) B_i(x)\right] \geqslant -\varepsilon \|g\|_1 - \alpha \frac{h^d (2m)^d}{n^d} .$$

Recalling now (2.3), we have derived a contradiction to $\alpha_+(M) \leqslant \alpha$ if we can choose the right-hand side to be bigger than β (which is a constant depending only on g). Setting $\varepsilon = -\beta/(2\|g\|_1)$, we consult (2.8) to see that

$$h \leqslant C'(g) m \log \alpha \tag{2.9}$$

for some other constant $C'(g)$ depending only on g.

We thus arrive at a contradiction as long as $\alpha = o((n/\log n)^d)$, recalling that m, d depend only on g. This completes the argument.

2.6 Exercises

2.6.1 Stronger lower bounds

So far, we only saw how to prove a lower bound of $\bar{\gamma}(\mathrm{CUT}_n) \geqslant n^{\omega(1)}$. To obtain stronger quantitative lower bounds, one has to analyze carefully the parts of the argument that read "some constant depending only on g." To do this properly, it turns out that one needs an appropriate definition of "approximate" junta degree. Basically, the "proof" that a function has large junta degree (the locally positive functional) has to be robust.

For a function $f : \{0,1\}^n \to \mathbb{R}_+$ and $\varepsilon > 0$, define

$$\deg_{\mathrm{J}}^{\varepsilon}(f) = 1 + \max\left\{d : \exists \text{ a } d\text{-locally positive functional } \varphi : \{0,1\}^n \to \mathbb{R} \text{ s.t. } \langle \varphi, f\rangle < -\varepsilon \|\varphi\|_\infty \|f\|_1\right\} .$$

We take the maximum to be -1 if no such functional exists. One can prove the following (see Lee-Raghavendra-Steurer 2014).

Exercise 2.3. Give an equivalent characterization of $\deg_{\mathrm{J}}^{\varepsilon}$ in terms of approximating f by a non-negative sum of juntas, where the approximation is in the L^1-norm.

Theorem 2.4. *For any $g : \{0,1\}^m \to \mathbb{R}_+$ and $\varepsilon > 0$, the following holds. For all $n \geqslant 2m$,*

$$\mathrm{rank}_+(M_n^g) \geqslant \left(\frac{c\varepsilon^2 n}{m^2(d \log n + \log(\|g\|_\infty/\|g\|_1))}\right)^{\deg_{\mathrm{J}}^{\varepsilon}(g)-1} ,$$

where $c > 0$ is a universal constant.

The point of this result is that there is no hidden constant. This allows one to prove much stronger bounds.

Exercise 2.5. Prove that there is a constant $\varepsilon > 0$ such that for all $n \geqslant 3$, we have

$$\deg_J^{\varepsilon}(f) \geqslant \frac{n}{2} + 1\,,$$

where f is the function defined in Theorem 1.12. Combined with Theorem 2.4, what lower bound does this yield for $\bar{\gamma}(\mathrm{CUT}_n)$?

2.6.2 Approximation

We have been concerned so far with *exact* characterization of polytopes (and, mainly, CUT_n). But in this model, one can also talk about approximate lifts. For instance, consider the MAX-CUT problem: Given a non-negative weight $w : \binom{[n]}{2} \to \mathbb{R}_+$ on the edges of the (undirected) complete graph, the goal is to compute the *maximum-cut value*

$$\mathrm{opt}(w) \stackrel{\mathrm{def}}{=} \max_{z \in \mathrm{CUT}_n} \frac{\langle w, z\rangle}{\|w\|_1}\,.$$

The objective is the (normalized) weight of edges cut. (Strictly speaking, MAX-CUT involves finding the optimizer, not just its value.)

Fix the number of vertices n. For some constants $1 \geqslant c > s \geqslant 0$, let us consider the matrix

$$M^{c,s}(w, z) = c - \frac{\langle z, w\rangle}{\|w\|_1}\,,$$

where w ranges over all weighted graphs w with $\mathrm{opt}(w) \leqslant s$ and z ranges over all cuts (the extreme rays of CUT_n).

Exercise 2.6. Argue that if $\mathrm{rank}_+(M^{c,s}) \geqslant r$, then the following holds. For any polytope P defined by at most r inequalities, if P linearly projects to a polytope $\hat{P} \subseteq \mathbb{R}^{\binom{n}{2}}$ such that $\hat{P} \supseteq \mathrm{CUT}_n$, then there exists a weighted graph w such that $\mathrm{opt}(w) \leqslant s$, but

$$\max_{z \in \hat{P}} \frac{\langle z, w\rangle}{\|w\|_1} \geqslant c\,.$$

In other words, $\hat{P}$ does a poor job of capturing the MAX-CUT optimization over CUT_n.

3 PSD rank and sums-of-squares degree

We have previously explored whether the cut polytope can be expressed as the linear projection of a polytope with a small number of facets (i.e., whether it has a small linear programming extended formulation).

For many cut problems, semi-definite programs (SDPs) are able to achieve better approximation ratios than LPs. The most famous example is the Goemans-Williamson 0.878-approximation for MAX-CUT. The techniques we have seen so far (see also Section 2.6) are able to show that no polynomial-size LP can achieve better than factor 1/2.

3.1 Spectrahedral lifts

The feasible regions of LPs are polyhedra. Up to linear isomorphism, every polyhedron P can be represented as $P = \mathbb{R}_+^n \cap V$ where $\mathbb{R}_+^n$ is the positive orthant and $V \subseteq \mathbb{R}^n$ is an affine subspace.

In this context, it makes sense to study any cones that can be optimized over efficiently. A prominent example is the positive semi-definite cone. Let us define $\mathcal{S}_+^n \subseteq \mathbb{R}^{n^2}$ as the set of $n \times n$ real, symmetric matrices with non-negative eigenvalues. A *spectrahedron* is the intersection $\mathcal{S}_+^n \cap V$ with an affine subspace V. The value n is referred to as the *dimension* of the spectrahedron.

In analogy with the γ parameter we defined for polyhedral lifts, let us define $\bar{\gamma}_{\mathsf{sdp}}(P)$ for a polytope P to be the minimal dimension of a spectrahedron that linearly projects to P.

Exercise° 3.1. Show that $\bar{\gamma}_{\mathsf{sdp}}(P) \leqslant \bar{\gamma}(P)$ for every polytope P. In other words, spectahedral lifts are at least as powerful as polyhedral lifts in this model.

In fact, spectrahedral lifts can be strictly more powerful. Certainly there are many examples of this in the setting of approximation (like the Goemans-Williamson SDP mentioned earlier), but there are also recent gaps between $\bar{\gamma}$ and $\bar{\gamma}_{\text{sdp}}$ for exact characterizations of polytopes; see the work of Fawzi, Saunderson, and Parrilo (2015).

Nevertheless, we are recently capable of proving strong lower bounds on the dimension of such lifts. Let us consider the cut polytope CUT_n as in previous posts.

Theorem 3.2 (Lee-Raghavendra-Steurer 2015)**.** *There is a constant $c > 0$ such that for every $n \geqslant 1$, $\bar{\gamma}_{\mathsf{sdp}}(\mathrm{CUT}_n) \geqslant e^{cn^{2/11}}$.*

Our goal now is to understand how the general framework we have seen for LP lower bounds extends to the SDP setting.

3.2 PSD rank and factorizations

Just as in the setting of polyhedra, there is a notion of "factorization through a cone" that characterizes the parameter $\bar{\gamma}_{\mathsf{sdp}}(P)$. Let $M \in \mathbb{R}_+^{m\times n}$ be a non-negative matrix. One defines the *psd rank* of M as the quantity

$$\mathrm{rank}_{\mathsf{psd}}(M) = \min\left\{r : M_{ij} = \mathrm{Tr}(A_i B_j) \text{ for some } A_1, \dots, A_m, B_1, \dots, B_n \in \mathcal{S}_+^r\right\} .$$

The following theorem was independently proved by Fiorini-Massar-Pokutta-Tiwari-de Wolf and Gouveia-Parrilo-Thomas. The proof is a direct analog of Yannakakis' proof for non-negative rank.

Theorem 3.3. *For every polytope P, it holds that $\bar{\gamma}_{\mathsf{sdp}}(P) = \mathrm{rank}_{\mathsf{psd}}(M)$ for any slack matrix M of P.*

Recall the class QML_n^+ of non-negative quadratic multi-linear functions that are positive on $\{0,1\}^n$ and the matrix $\mathcal{M}_n : \mathrm{QML}_n^+ \times \{0,1\}^n \to \mathbb{R}_+$ given by

$$\mathcal{M}_n(f, x) = f(x) .$$

We saw previously that $\mathcal{M}_n$ is a submatrix of some slack matrix of CUT_n. Thus our goal is to prove a lower bound on $\mathrm{rank}_{\mathsf{psd}}(\mathcal{M}_n)$.

3.3 Sum-of-squares certificates

Just as in the setting of non-negative matrix factorization, we can think of a low psd rank factorization of $\mathcal{M}_n$ as a small set of "axioms" that can prove the non-negativity of every function in QML_n^+. But now our proof system is considerably more powerful.

For a subspace of functions $\mathcal{U} \subseteq L^2(\{0,1\}^n)$, let us define the cone

$$\mathsf{sos}(\mathcal{U}) = \mathrm{cone}\left(q^2 : q \in \mathcal{U}\right) .$$

This is the cone of squares of functions in $\mathcal{U}$. We will think of $\mathcal{U}$ as a set of axioms of size $\dim(\mathcal{U})$ that is able to assert non-negativity of every $f \in \mathrm{sos}(\mathcal{U})$ by writing

$$f = \sum_{i=1}^{k} q_i^2$$

for some $q_1, \ldots, q_k \in \mathsf{sos}(\mathcal{U})$.

Fix a subspace $\mathcal{U}$ and let $r = \dim(\mathcal{U})$. Fix also a basis $q_1, \ldots, q_r : \{0,1\}^n \to \mathbb{R}$ for $\mathcal{U}$.

Define $B : \{0,1\}^n \to \mathcal{S}_+^r$ by setting $B(x)_{ij} = q_i(x)q_j(x)$. Note that $B(x)$ is PSD for every x because $B(x) = \vec{q}(x)\vec{q}(x)^T$ where $\vec{q}(x) = (q_1(x), \ldots, q_r(x))$.

We can write every $p \in \mathcal{U}$ as $p = \sum_{i=1}^r \lambda_i q_i$. Defining $\Lambda(p^2) \in \mathcal{S}_+^r$ by $\Lambda(p^2)_{ij} = \lambda_i \lambda_j$, we see that

$$\mathrm{Tr}(\Lambda(p^2)Q(x)) = \sum_{i,j} \lambda_i \lambda_j q_i(x) q_j(x) = p(x)^2 .$$

Now every $f \in \mathsf{sos}(\mathcal{U})$ can be written as $\sum_{i=1}^k c_i p_i^2$ for some $k \geqslant 0$ and $\{c_i \geqslant 0\}$. Therefore if we define $\Lambda(f) = \sum_{i=1}^k c_i \Lambda(p_i^2)$, we arrive at the representation

$$f(x) = \mathrm{Tr}(\Lambda(f)Q(x)) .$$

In conclusion, if $\mathrm{QML}_+^n \subseteq \mathsf{sos}(\mathcal{U})$, then $\mathrm{rank}_{\mathsf{psd}}(\mathcal{M}_n) \leqslant \dim(\mathsf{sos}(\mathcal{U}))$.

Exercise 3.4. Show that an approximate converse holds: $\dim(\mathsf{sos}(\mathcal{U})) \leqslant \mathrm{rank}_{\mathsf{psd}}(\mathcal{M}_n)^2$.

3.4 The canonical axioms

And just as d-juntas were the canonical axioms for our NMF proof system, there is a similar canonical family in the SDP setting: Let Q_d be the subspace of all degree-d multi-linear polynomials on $\mathbb{R}^n$. We have

$$\dim(Q_d) \leqslant \sum_{k=0}^{d} \binom{n}{k} \leqslant 1 + n^d . \tag{3.1}$$

For a function $f : \{0,1\}^n \to \mathbb{R}_+$, one defines

$$\deg_{\mathsf{sos}}(f) = \min\{d : f \in \mathsf{sos}(Q_d)\} .$$

(One could debate whether the definition of sum-of-squares degree should have $d/2$ or d. The most convincing arguments suggest that we should use membership in the cone of squares over $Q_{\lfloor d/2 \rfloor}$ so that the sos-degree will be at least the real-degree of the function.)

On the other hand, our choice has the following nice property.

Lemma 3.5. *For every* $f : \{0,1\}^n \to \mathbb{R}$, *we have* $\deg_{\text{sos}}(f) \leqslant \deg_J(f)$.

Proof. If q is a non-negative d-junta, then $\sqrt{q}$ is also a non-negative d-junta. It is elementary to see that every d-junta on $\{0,1\}^n$ has a multi-linear polynomial representation of degree at most d, thus q is the square of a multi-linear polynomial of degree at most d. □

3.5 The dual cone

As with junta-degree, there is a simple characterization of sos-degree in terms of separating functionals. Say that a functional $\varphi : \{0,1\}^n \to \mathbb{R}$ is *degree-d pseudo-positive* if

$$\langle \varphi, q^2 \rangle = \mathop{\mathbb{E}}_{x \in \{0,1\}^n} \varphi(x) q(x)^2 \geqslant 0$$

whenever $q : \{0,1\}^n \to \mathbb{R}$ satisfies $\deg(q) \leqslant d$ (and by deg here, we mean degree as a multi-linear polynomial on $\{0,1\}^n$).

Again, since $\mathsf{sos}(Q_d)$ is a convex set, these separating functionals are the only way of exhibiting non-membership. The following characterization is elementary.

Lemma 3.6. *For every* $f : \{0,1\}^n \to \mathbb{R}_+$, *it holds that* $\deg_{\mathsf{sos}}(f) > d$ *if and only if there is a degree-d pseudo-positive functional* $\varphi : \{0,1\}^n \to \mathbb{R}$ *such that* $\langle \varphi, f \rangle < 0$.

3.6 The connection to psd rank

Following the analogy with non-negative rank, we have two objectives left: (1) to exhibit a function $f \in \mathrm{QML}_n^+$ with $\deg_{\mathsf{sos}}(f)$ large, and (ii) to give a connection between the sum-of-squares of f and the psd rank of an associated matrix.

Notice that the function $g(x) = (1 - \sum_{i=1}^m x_i)^2$ we used for junta-degree has $\deg_{\mathsf{sos}}(g) = 1$, making it a poor candidate. In fact, this implies that $\mathrm{rank}_{\mathsf{psd}}(M_n^g) \leqslant O(n)$, while we have seen that $\mathrm{rank}_+(M_n^g) \geqslant \Omega((n/\log n)^{m-1})$ as $n \to \infty$.

Fortunately, Grigoriev has shown that the *knapsack polynomial* (from 1.13) has large sos-degree.

Theorem 3.7. *For every odd* $m \geqslant 1$, *the function*

$$f(x) = \left(\frac{m}{2} - \sum_{i=1}^{m} x_i\right)^2 - \frac{1}{4}$$

has $\deg_{\mathsf{sos}}(f) \geqslant \lfloor m/2 \rfloor$.

Observe that this f is non-negative over $\{0,1\}^m$ (because m is odd), but it is manifestly *not* non-negative on $\mathbb{R}^m$.

Finally, we recall the submatrices of $\mathcal{M}_n$ defined as follows. Fix some integer $m \geqslant 1$ and a function $g : \{0,1\}^m \to \mathbb{R}_+$. Then $M_n^g : \binom{[n]}{m} \times \{0,1\}^n \to \mathbb{R}_+$ is given by

$$M_n^g(S, x) = g(x|_S)\,.$$

Our goal now is to sketch the following analog of Theorem 1.15.

Theorem 3.8 (Lee-Raghavendra-Steurer 2015)**.** *For every $m \geqslant 1$ and $g : \{0,1\}^m \to \mathbb{R}_+$, there exists a constant $C(g)$ such that the following holds. For every $n \geqslant 2m$,*

$$1 + n^{d/2} \geqslant \operatorname{rank}_{\mathsf{psd}}(M_n^g) \geqslant C\left(\frac{n}{\log n}\right)^{(d-1)/2},$$

where $d = \deg_{\mathsf{sos}}(g)$.

Note that the upper bound is from (3.1) and the non-trivial content is contained in the lower bound.

3.7 Exercise: Proving a lower bound on $\deg_{\mathsf{sos}}$

[This exercise follows an elegant argument of J. Kaniewski, T. Lee, and R. de Wolf (2014).]

You will prove a lower bound on the sum-of-squares degree of the function $f : \{0,1\}^n \to \mathbb{R}_+$ given by

$$f(x) = (|x| - 1)(|x| - 2), \tag{3.2}$$

where we use $|x| = \sum_{i=1}^n x_i$ for the hamming weight of $x \in \{0,1\}^n$.

Suppose that we can write

$$f(x) = \sum_{i=1}^N p_i(x)^2,$$

where $\deg(p_i) \leqslant d$ for every $i \in [N]$. Define the function $q_i : [n] \to \mathbb{R}$ by

$$q_i(k) = \mathop{\mathbb{E}}_{x \in \{0,1\}^n : |x| = k} \left[p_i(x)\right].$$

The first step can be accomplished using the Fourier representation of functions on $\{0,1\}^n$, or using an appropriate averaging procedure.

(a) Show that there is a function $\tilde{q}_i : \mathbb{R} \to \mathbb{R}$ that agrees with q_i on $[n]$ and such that $\deg(\tilde{q}_i) \leqslant d$.

Now let us define $Q(t) = \sum_{i=1}^N \tilde{q}_i(t)^2$, which is a polynomial of degree at most $2d$. We also have $Q(1) = Q(2) = 0$ since $f(x) = 0$ for $|x| \in \{1,2\}$. The zeroes of a non-negative real polynomial must have multiplicity at least 2, thus we can write

$$Q(t) = (t-1)^2(t-2)^2 q(t)$$

for some polynomial q with $\deg(q) \leqslant 2d - 4$.

(b) Your goal now is to prove a lower bound $\deg(q) \geqslant \Omega(\sqrt{n})$, implying that $\deg_{\mathsf{sos}}(f) \geqslant \Omega(\sqrt{n})$. [Note that plugging this into Theorem 3.8 is enough to show that $\bar{\gamma}_{\mathsf{sdp}}(\mathrm{CUT}_n)$ must grow faster than any polynomial.]

You should be able to do this using the following oft-employed lemma of A. A. Markov.

Lemma 3.9. *If $q : \mathbb{R} \to \mathbb{R}$ is a polynomial, then for every $T \geqslant 0$,*

$$\deg(q) \geqslant \sqrt{\frac{T}{2} \frac{\max_{x \in [0,T]} |q'(x)|}{\max_{x \in [0,T]} |q(x)|}}.$$

Remark 3.10. Note that the function f defined in (3.2) is symmetric with respect to permutations of the coordinates. Thus we can symmetrize (3.2) by averaging over all permutations. It would be great if we could have symmetrized *inside* the square, since then we would have an exact univariate representation of f.

A natural way to attempt this might be to write

$$p_i(x)^2 = \left(\mathop{\mathbb{E}}_{y \in \{0,1\}^n : |y| = |x|} p_i(y) \right)^2 + \text{ other stuff}\,,$$

where "other stuff" captures the non-symmetric part of p_i. This can be done using a little representation theory of the symmetric group, and was pursued by Blekherman, Gouveia, and Pfeiffer (2014). This method is capable of proving Theorem 3.7, for instance.

3.8 Finite-dimensional operator norms

Let H denote a finite-dimensional Euclidean space over $\mathbb{R}$ equipped with inner product $\langle \cdot, \cdot \rangle$ and norm $|\cdot|$. For a linear operator $A : H \to H$, we define the operator, trace, and Frobenius norms by

$$\|A\| = \max_{x \neq 0} \frac{|Ax|}{x}, \quad \|A\|_* = \mathrm{Tr}(\sqrt{A^T A}), \quad \|A\|_F = \sqrt{\mathrm{Tr}(A^T A)}\,.$$

Let $\mathcal{M}(H)$ denote the set of self-adjoint linear operators on H. Note that for $A \in \mathcal{M}(H)$, the preceding three norms are precisely the ℓ_∞, ℓ_1, and ℓ_2 norms of the eigenvalues of A. For $A, B \in \mathcal{M}(H)$, we use $A \succeq 0$ to denote that A is positive semi-definite and $A \succeq B$ for $A - B \succeq 0$. We use $\mathcal{D}(H) \subseteq \mathcal{M}(H)$ for the set of density operators: Those $A \in \mathcal{M}(H)$ with $A \succeq 0$ and $\mathrm{Tr}(A) = 1$.

One should recall that $\mathrm{Tr}(A^T B)$ is an inner product on the space of linear operators, and we have the operator analogs of the Hölder inequalities: $\mathrm{Tr}(A^T B) \leqslant \|A\| \cdot \|B\|_*$ and $\mathrm{Tr}(A^T B) \leqslant \|A\|_F \|B\|_F$.

3.9 John's theorem and factorization rescaling

As in the case of non-negative rank, consider finite sets X and Y and a matrix $M : X \times Y \to \mathbb{R}_+$. For the purposes of proving a lower bound on the psd rank of some matrix, we would like to have a nice analytic description.

To that end, suppose we have a rank-r psd factorization

$$M(x, y) = \mathrm{Tr}(A(x)B(y))$$

where $A : X \to \mathcal{S}_+^r$ and $B : Y \to \mathcal{S}_+^r$. The following result of Briët, Dadush and Pokutta (2013) gives us a way to "scale" the factorization so that it becomes nicer analytically. (The improved bound stated here is from an article of Fawzi, Gouveia, Parrilo, Robinson, and Thomas, and we follow their proof.)

Lemma 3.11. *Every M with $\mathrm{rank}_{\mathsf{psd}}(M) \leqslant r$ admits a factorization $M(x, y) = \mathrm{Tr}(P(x)Q(y))$ where $P : X \to \mathcal{S}_+^r$ and $Q : Y \to \mathcal{S}_+^r$ and, moreover,*

$$\max\{\|P(x)\| \cdot \|Q(y)\| : x \in X, y \in Y\} \leqslant r\|M\|_\infty\,,$$

where $\|M\|_\infty = \max_{x \in X, y \in Y} M(x, y)$.

Proof. Start with a rank-r psd factorization $M(x,y) = \mathrm{Tr}(A(x)B(y))$. Observe that there is a degree of freedom here, because for any invertible operator J, we get another psd factorization $M(x,y) = \mathrm{Tr}\left((JA(x)J^T)\cdot((J^{-1})^T B(y)J^{-1})\right)$.

Let $U = \{u \in \mathbb{R}^r : \exists x \in X\ A(x) \succeq uu^T\}$ and $V = \{v \in \mathbb{R}^r : \exists y \in X\ B(y) \succeq vv^T\}$. Set $\Delta = \|M\|_\infty$. We may assume that U and V both span $\mathbb{R}^r$ (else we can obtain a lower-rank psd factorization). Both sets are bounded by finiteness of X and Y.

Let $C = \mathrm{conv}(U)$ and note that C is centrally symmetric and contains the origin. Now John's theorem tells us there exists a linear operator $J : \mathbb{R}^r \to \mathbb{R}^r$ such that

$$B_{\ell_2} \subseteq JC \subseteq \sqrt{r} B_{\ell_2}\,, \tag{3.3}$$

where B_{ℓ_2} denotes the unit ball in the Euclidean norm. Let us now set $P(x) = JA(x)J^T$ and $Q(y) = (J^{-1})^T B(y) J^{-1}$.

Eigenvalues of $P(x)$:. Let w be an eigenvector of $P(x)$ normalized so the corresponding eigenvalue is $\|w\|_2^2$. Then $P(x) \succeq ww^T$, implying that $J^{-1}w \in U$ (here we use that $A \succeq 0 \implies SAS^T \succeq 0$ for any S). Since $w = J(J^{-1}w)$, (3.3) implies that $\|w\|_2 \leqslant \sqrt{r}$. We conclude that every eigenvalue of $P(x)$ is at most r.

Eigenvalues of $Q(y)$:. Let w be an eigenvector of $Q(y)$ normalized so that the corresponding eigenvalue is $\|w\|_2^2$. Then as before, we have $Q(y) \succeq ww^T$ and this implies $J^T w \in V$. Now, on the one hand we have

$$\max_{z \in JC} \langle z, w\rangle \geqslant \|w\|_2 \tag{3.4}$$

since $JC \supseteq B_{\ell_2}$.

On the other hand:

$$\max_{z \in JC} \langle z, w\rangle^2 = \max_{z \in C} \langle Jz, w\rangle^2 = \max_{z \in C} \langle z, J^T w\rangle^2\,. \tag{3.5}$$

Finally, observe that for any $u \in U$ and $v \in V$, we have

$$\langle u, v\rangle^2 = \langle uu^T, vv^T\rangle \leqslant \max_{x \in X, y \in Y} \langle A(x), B(y)\rangle \leqslant \Delta\,.$$

By convexity, this implies that $\max_{z \in C} \langle z, v\rangle^2 \leqslant \Delta$ for all $v \in V$, bounding the right-hand side of (3.5) by Δ. Combining this with (3.4) yields $\|w\|_2^2 \leqslant \Delta$. We conclude that all the eigenvalues of $Q(y)$ are at most Δ. □

3.10 Analytic psd rank

Again, in analogy with the non-negative rank setting, we can define an "analytic psd rank" parameter for matrices $N : X \times Y \to \mathbb{R}_+$:

$$\begin{aligned}
\alpha_{\mathsf{psd}}(N) = \min \Big\{ \alpha \mid \exists A : X \to \mathcal{S}_+^k, B : Y \to \mathcal{S}_+^k,\\
\mathop{\mathbb{E}}_{x \in X}[A(x)] = I,\\
\|B(y)\| \leqslant \frac{\alpha}{k} \mathop{\mathbb{E}}_{y \in Y}[\mathrm{Tr}(B(y))] \quad \forall y \in Y\\
\|A(x)\| \leqslant \alpha \quad \forall x \in X \Big\}\,.
\end{aligned}$$

Note that we have implicit equipped X and Y with the uniform measure. The main point here is that k can be arbitrary. One can verify that for every $c > 0$, the set $\{N : \alpha_{\mathsf{psd}(N)} \leqslant c\}$ is convex.

And there is a corresponding approximation lemma. We use $\|N\|_\infty = \max_{x,y} |N(x,y)|$ and $\|N\|_1 = \mathbb{E}_{x,y} |N(x,y)|$.

Lemma 3.12. *For every non-negative matrix $M : X \times Y \to \mathbb{R}_+$ and every $\eta \in (0,1]$, there is a matrix N such that $\|M - N\|_\infty \leqslant \eta \|M\|_\infty$ and*

$$\alpha_{\mathsf{psd}}(N) \leqslant O(\mathrm{rank}_{\mathsf{psd}}(M)) \frac{1}{\eta} \frac{\|M\|_\infty}{\|M\|_1}.$$

Using Lemma 3.11 in a straightforward way, it is not particularly difficult to construct the approximator N. The condition $\mathbb{E}_x[A(x)] = I$ poses a slight difficulty that requires adding a small multiple of the identity to the LHS of the factorization (to avoid a poor condition number), but this has a correspondingly small affect on the approximation quality.

3.11 Quantum entropy maximization

Given a density matrix $P \in \mathcal{D}(H)$, we define the *quantum relative entropy* by

$$\mathcal{S}(P) = \mathrm{Tr}(P \log P).$$

(Strictly speaking, this is the relative entropy between P and the maximally mixed state $\mathcal{U} = \frac{\mathrm{Id}}{\mathrm{Tr}(\mathrm{Id})}$.) Note that $\mathcal{S}(P) = \sum_i \lambda_i \log \lambda_i$ where $\{\lambda_i\}$ are the eigenvalues of P. In analogy with the classical case, $\mathcal{S}$ is a convex function on $\mathcal{D}(H)$.

Let us now suppose that $\alpha_{\mathsf{psd}}(M_n^g) \leqslant \alpha$, meaning that we can write

$$M_n^g(S,x) = \mathrm{Tr}(A(S)B(x)) \tag{3.6}$$

for some $A : \binom{[n]}{m} \to \mathcal{S}_+^k$ and $B : \{0,1\}^n \to \mathcal{S}_+^k$ with $\mathbb{E}_S A(S) = \mathrm{Id}$ and

$$\|B(y)\| \leqslant \frac{\alpha}{k} \mathop{\mathbb{E}}_x \mathrm{Tr}(B(x)) \quad \forall y \in \{0,1\}^n \tag{3.7}$$

$$\|A(S)\| \leqslant \alpha \quad \forall S \in \binom{[n]}{m}. \tag{3.8}$$

Note that since $\mathbb{E}_S A_S = \mathrm{Id}$, we have

$$\|g\|_1 = \mathop{\mathbb{E}}_{S,x} M_n^g(S,x) = \mathop{\mathbb{E}}_{S,x} \mathrm{Tr}(B(x)A(S)) = \mathop{\mathbb{E}}_x \mathrm{Tr}(B(x)).$$

Let us define $K = \mathbb{E}_x \mathrm{Tr}(B(x))$ and rescale so that

$$\mathop{\mathbb{E}}_x \mathrm{Tr}(B(x)) = 1, \tag{3.9}$$

$$\mathop{\mathbb{E}}_S A(S) = K \cdot \mathrm{Id}, \tag{3.10}$$

$$\|A(S)\| \leqslant K \cdot \alpha. \tag{3.11}$$

Recall for later that K is a constant depending only on g.

As in the setting of non-negative rank, we would like to approximate B by a simpler mapping with respect to a set of test functionals. Let $\varphi : \{0,1\}^m \to \mathbb{R}$ be a degree-d pseudo-positive functional such that $\langle \varphi, g \rangle < 0$. For $S \in \binom{[n]}{m}$, denote $\varphi_S(x) = \varphi(x|_S)$.

Suppose, for a moment, that

$$\deg(\sqrt{B}) \stackrel{\text{def}}{=} \max_{i,j} \deg\left(x \mapsto \sqrt{B(x)_{ij}}\right) \leqslant d\,.$$

In that case, we would have

$$\mathop{\mathbb{E}}_{S,x}[\varphi_S(x)\operatorname{Tr}(A(S)B(x))] = \mathop{\mathbb{E}}_{S}\mathop{\mathbb{E}}_{x}\left[\varphi_S(x)\sum_{i,j} A(S)_{ij}\left(\sqrt{B(x)_{ij}}\right)^2\right] \geqslant 0\,,$$

since φ_S is a degree-d pseudo-positive functional. On the other hand,

$$\mathop{\mathbb{E}}_{S,x}[\varphi_S(x)g(x|_S)] = \mathop{\mathbb{E}}_{x}[\varphi(x)g(x)] < 0\,.$$

We would thus arrive at a contradiction to our assumed factorization (3.6). So our goal will be to approximate B by a low-degree square.

Let us now imagine encoding a matrix-valued function $P : \{0,1\}^n \to \mathcal{S}_+^k$ as both a function and a $(k2^n) \times (k2^n)$ block-diagonal matrix: $P = 2^{-n}\sum_{x\in\{0,1\}^n} P(x) \otimes e_x e_x^T$ for some set of orthonormal vectors $\{e_x\}$ orthogonal to $\mathcal{S}_+^k$ (in some enlarged Hilbert space). Observe that when the function B is considered as a matrix in this way, we have $\operatorname{Tr}(B) = 1$.

Consider the convex optimization:

$$\begin{aligned}
&\text{minimize} && \mathcal{S}(P) && \\
&\text{subject to} && \operatorname{Tr}(P) = 1 && (3.12)\\
& && P \succeq 0 && (3.13)\\
& && \mathop{\mathbb{E}}_{S,x}[\varphi_S(x)\operatorname{Tr}(A(S)B(x))] \leqslant \mathop{\mathbb{E}}_{S,x}[\varphi_S(x)\operatorname{Tr}(A(S)P(x))] + \varepsilon\,. && (3.14)
\end{aligned}$$

Standard optimality conditions (the KKT conditions) tell us that the (unique) optimal solution P^* is of the form

$$P^*(x) = \frac{e^{-\lambda \mathbb{E}_S \varphi_S(x)A(S)}}{\operatorname{Tr}\left(e^{-\lambda \mathbb{E}_S \varphi_S(x)A(S)}\right)}\,, \tag{3.15}$$

and

$$\lambda \leqslant \frac{\mathcal{S}(B)}{\varepsilon^2} \leqslant \frac{\log \alpha}{\varepsilon^2}\,, \tag{3.16}$$

where the first inequality uses the fact that B is a feasible solution (thereby bounding the optimal value of the dual), and the second inequality uses (3.9).

Now we will try to approximate P^* by the square of a low-degree multi-linear polynomial in x. For a matrix valued function $P : \{0,1\}^n \to \mathbb{R}$, let us define $\deg(P)$ to be the maximum degree of the entries, i.e. the maximum degree of $x \mapsto P(x)_{ij}$ as i, j ranges over the dimensions of P.

We will use the fact that the Taylor expansion of $e^t = \sum_{k=0}^{\infty} \frac{t^k}{k!}$ converges fast. More specifically, e^t is well-approximated on the interval $[-\tau, \tau]$ by a polynomial of degree $O(r)$. This also implies that $e^{t/2}$ is well-approximated by the square of a low-degree polynomial. A similar fact is true for the matrix exponential e^Z when Z is a real-symmetric matrix, and we are concerned with matrices Z for which $\|Z\| \in [-\tau, \tau]$. Specifically, the following holds.

Lemma 3.13. *Let $\delta \in (0, \frac{1}{2}]$ and $\tau > 0$ be given. For every real, symmetric matrix $Z \in \mathcal{M}(H)$, there is a number $k \leqslant 3e\left(\|Z\| + \frac{\log(1/\varepsilon)}{\log\log(1/\varepsilon)}\right)$ and a univariate degree-k polynomial p with non-negative coefficients such that*

$$\left\| \frac{e^Z}{\mathrm{Tr}(e^Z)} - \frac{p(F)^2}{\mathrm{Tr}(p(F)^2)} \right\| \leqslant \delta\,.$$

Given the preceding lemma and the form of P^* from (3.15), we should now estimate the operator norm:

$$\sup_{x\in\{0,1\}^n} \left\| \mathop{\mathbb{E}}_S \varphi_S(x) A(S) \right\| \leqslant \|\varphi_S\|_\infty \left\| \mathop{\mathbb{E}}_S A(S) \right\| \overset{(3.10)}{=} K \cdot \|\varphi\|_\infty\,,$$

and recall that $\|\varphi\|_\infty$ and K depend only on g.

Putting all this together with (3.16), we see that P^* is well-approximated by a function of the form

$$P^*(x) \approx p\left(\mathop{\mathbb{E}}_S \varphi_S(x) A(S)\right)^2$$

where p is a polynomial of degree $O(\frac{\log \alpha}{\varepsilon^2})$. Moreover, $\deg(\varphi_S) = \deg(\varphi) \leqslant m$ (since φ is a function on m bits).

If we set $Q(x) = p\left(\mathbb{E}_S\, \varphi_S(x) A(S)\right)^2$, then

$$\deg(Q) \leqslant O\left(\frac{\log \alpha}{\varepsilon^2}\right) m\,,$$

where the implicit constant depends on the function g.

Just as with our bound (2.8) in the setting of non-negative factorizations, this bound is not nearly good enough since certainly we will have $d \leqslant m$ (since $g : \{0,1\}^m \to \mathbb{R}_+$).

But again random restriction is the key (and the way we employ the bound (3.11) on the A-side of the psd factorization).

Note that

$$\mathop{\mathbb{E}}_{S,x}\left[\varphi_S(x)\,\mathrm{Tr}(A(S)Q(x))\right] = \mathop{\mathbb{E}}_S \mathop{\mathbb{E}}_{y\in\{0,1\}^S}\left[\varphi(y)\,\mathrm{Tr}\left(A(S) \mathop{\mathbb{E}}_{x\in\{0,1\}^n : x|_S = y} Q(x)\right)\right].$$

Just as in Section 2.5, it turns out that with high probability over the choice of $|S| = m$, it holds that

$$y \mapsto \mathop{\mathbb{E}}_{x\in\{0,1\}^n : x|S = y} Q(x)$$

is well-approximated by the square of a much lower degree polynomial. The averaging over all bits outside of S again achieves a drastic reduction in the "simplicity" of the hypothesis. The quantitative calculation essentially parallels that in Section 2.5, with the added complication that we are averaging *outside* the square (recall that $Q = p(\cdot)^2$). This can be handled using a bit of Fourier analysis.