18.S096: Approximation Algorithms and Max-Cut

Topics in Mathematics of Data Science (Fall 2015)

Afonso S. Bandeira
bandeira@mit.edu
http://math.mit.edu/~bandeira

November 24, 2015

These are lecture notes not in final form and will be continuously edited and/or corrected (as I am sure it contains many typos). Please let me know if you find any typo/mistake. Also, I am posting the open problems on my Blog, see [Ban15].

8.1 The Max-Cut problem

Unless the widely believed $P \neq NP$ conjecture is false, there is no polynomial algorithm that can solve all instances of an NP-hard problem. Thus, when faced with an NP-hard problem (such as the `Max-Cut` problem discussed below) one has three options: to use an exponential type algorithm that solves exactly the problem in all instances, to design polynomial time algorithms that only work for some of the instances (hopefully the typical ones!), or to design polynomial algorithms that, in any instance, produce guaranteed approximate solutions. This section is about the third option. The second is discussed in later in the course, in the context of community detection.

The `Max-Cut` problem is the following: Given a graph $G = (V, E)$ with non-negative weights w_{ij} on the edges, find a set $S \subset V$ for which $\mathrm{cut}(S)$ is maximal. Goemans and Williamson [GW95] introduced an approximation algorithm that runs in polynomial time and has a randomized component to it, and is able to obtain a cut whose expected value is guaranteed to be no smaller than a particular constant α_{GW} times the optimum cut. The constant α_{GW} is referred to as the approximation ratio.

Let $V = \{1, \ldots, n\}$. One can restate `Max-Cut` as

$$\begin{array}{ll} \max & \frac{1}{2}\sum_{i<j} w_{ij}(1 - y_i y_j) \\ s.t. & |y_i| = 1 \end{array} \qquad (1)$$

The y_i's are binary variables that indicate set membership, i.e., $y_i = 1$ if $i \in S$ and $y_i = -1$ otherwise.

Consider the following relaxation of this problem:

$$\begin{array}{ll} \max & \frac{1}{2}\sum_{i<j} w_{ij}(1 - u_i^T u_j) \\ s.t. & u_i \in \mathbb{R}^n, \|u_i\| = 1. \end{array} \qquad (2)$$

This is in fact a relaxation because if we restrict u_i to be a multiple of e_1, the first element of the canonical basis, then (12) is equivalent to (1). For this to be a useful approach, the following two properties should hold:

(a) Problem (12) needs to be easy to solve.

(b) The solution of (12) needs to be, in some way, related to the solution of (1).

Definition 8.1 *Given a graph G, we define* $\mathrm{MaxCut}(G)$ *as the optimal value of* (1) *and* $\mathcal{R}\mathrm{MaxCut}(G)$ *as the optimal value of* (12).

We start with property (a). Set X to be the Gram matrix of $u_1, \ldots, u_n$, that is, $X = U^T U$ where the i'th column of U is u_i. We can rewrite the objective function of the relaxed problem as

$$\frac{1}{2}\sum_{i<j} w_{ij}(1 - X_{ij})$$

One can exploit the fact that X having a decomposition of the form $X = Y^T Y$ is equivalent to being positive semidefinite, denoted $X \succeq 0$. The set of PSD matrices is a convex set. Also, the constraint $\|u_i\| = 1$ can be expressed as $X_{ii} = 1$. This means that the relaxed problem is equivalent to the following semidefinite program (SDP):

$$\begin{array}{ll} \max & \frac{1}{2}\sum_{i<j} w_{ij}(1 - X_{ij}) \\ s.t. & X \succeq 0 \text{ and } X_{ii} = 1,\ i = 1, \ldots, n. \end{array} \tag{3}$$

This SDP can be solved (up to ϵ accuracy) in time polynomial on the input size and $\log(\epsilon^{-1})$[VB96].

There is an alternative way of viewing (3) as a relaxation of (1). By taking $X = yy^T$ one can formulate a problem equivalent to (1)

$$\begin{array}{ll} \max & \frac{1}{2}\sum_{i<j} w_{ij}(1 - X_{ij}) \\ s.t. & X \succeq 0\ ,\ X_{ii} = 1,\ i = 1, \ldots, n, \text{ and } \mathrm{Rank}(X) = 1. \end{array} \tag{4}$$

The SDP (3) can be regarded as a relaxation of (4) obtained by removing the non-convex rank constraint. In fact, this is how we will later formulate a similar relaxation for the minimum bisection problem.

We now turn to property (b), and consider the problem of forming a solution to (1) from a solution to (3). From the solution $\{u_i\}_{i=1,\ldots,n}$ of the relaxed problem (3), we produce a cut subset S' by randomly picking a vector $r \in \mathbb{R}^n$ from the uniform distribution on the unit sphere and setting

$$S' = \{i | r^T u_i \geq 0\}$$

In other words, we separate the vectors $u_1, \ldots, u_n$ by a random hyperplane (perpendicular to r). We will show that the cut given by the set S' is comparable to the optimal one.

Let W be the value of the cut produced by the procedure described above. Note that W is a random variable, whose expectation is easily seen (see Figure 1) to be given by

$$\begin{aligned} \mathbb{E}[W] &= \sum_{i<j} w_{ij} \Pr\left\{\mathrm{sign}(r^T u_i) \neq \mathrm{sign}(r^T u_j)\right\} \\ &= \sum_{i<j} w_{ij} \frac{1}{\pi} \arccos(u_i^T u_j). \end{aligned}$$

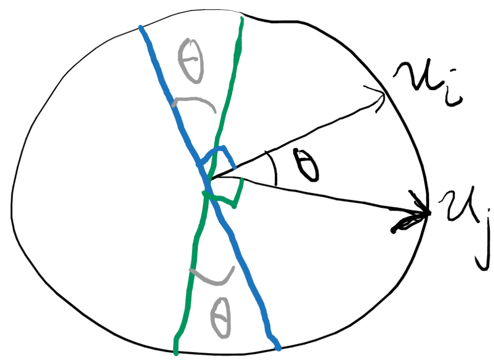

Figure 1: $\theta = \arccos(u_i^T u_j)$

If we define α_{GW} as

$$\alpha_{GW} = \min_{-1 \le x \le 1} \frac{\frac{1}{\pi}\arccos(x)}{\frac{1}{2}(1-x)},$$

it can be shown that $\alpha_{GW} > 0.87$.

It is then clear that

$$\mathbb{E}[W] = \sum_{i<j} w_{ij} \frac{1}{\pi} \arccos(u_i^T u_j) \ge \alpha_{GW} \frac{1}{2} \sum_{i<j} w_{ij}(1 - u_i^T u_j). \quad (5)$$

Let $\mathrm{MaxCut}(G)$ be the maximum cut of G, meaning the maximum of the original problem (1). Naturally, the optimal value of (12) is larger or equal than $\mathrm{MaxCut}(G)$. Hence, an algorithm that solves (12) and uses the random rounding procedure described above produces a cut W satisfying

$$\mathrm{MaxCut}(G) \ge \mathbb{E}[W] \ge \alpha_{GW} \frac{1}{2} \sum_{i<j} w_{ij}(1 - u_i^T u_j) \ge \alpha_{GW} \mathrm{MaxCut}(G), \quad (6)$$

thus having an approximation ratio (in expectation) of α_{GW}. Note that one can run the randomized rounding procedure several times and select the best cut.

Note that the above gives

$$\mathrm{MaxCut}(G) \ge \mathbb{E}[W] \ge \alpha_{GW} \mathcal{R}\mathrm{MaxCut}(G) \ge \alpha_{GW} \mathrm{MaxCut}(G)$$

8.2 Can α_{GW} be improved?

A natural question is to ask whether there exists a polynomial time algorithm that has an approximation ratio better than α_{GW}.

The unique games problem (as depicted in Figure 2) is the following: Given a graph and a set of k colors, and, for each edge, a matching between the colors, the goal in the unique games problem is to color the vertices as to agree with as high of a fraction of the edge matchings as possible. For example, in Figure 2 the coloring agrees with $\frac{3}{4}$ of the edge constraints, and it is easy to see that one cannot do better.

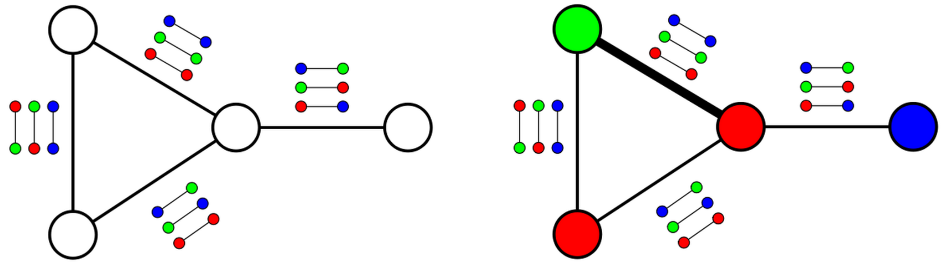

Figure 2: The Unique Games Problem

The Unique Games Conjecture of Khot [Kho02], has played a major role in hardness of approximation results.

Conjecture 8.2 *For any $\epsilon > 0$, the problem of distinguishing whether an instance of the Unique Games Problem is such that it is possible to agree with a $\geq 1 - \epsilon$ fraction of the constraints or it is not possible to even agree with a ϵ fraction of them, is NP-hard.*

There is a sub-exponential time algorithm capable of distinguishing such instances of the unique games problem [ABS10], however no polynomial time algorithm has been found so far. At the moment one of the strongest candidates to break the Unique Games Conjecture is a relaxation based on the Sum-of-squares hierarchy that we will discuss below.

Open Problem 8.1 *Is the Unique Games conjecture true? In particular, can it be refuted by a constant degree Sum-of-squares relaxation?*

Remarkably, approximating `Max-Cut` with an approximation ratio better than α_{GW} is has hard as refuting the Unique Games Conjecture (UG-hard) [KKMO05]. More generality, if the Unique Games Conjecture is true, the semidefinite programming approach described above produces optimal approximation ratios for a large class of problems [Rag08].

Not depending on the Unique Games Conjecture, there is a NP-hardness of approximation of $\frac{16}{17}$ for `Max-Cut` [Has02].

Remark 8.3 *Note that a simple greedy method that assigns membership to each vertex as to maximize the number of edges cut involving vertices already assigned achieves an approximation ratio of $\frac{1}{2}$ (even of $\frac{1}{2}$ of the total number of edges, not just of the optimal cut).*

8.3 A Sums-of-Squares interpretation

We now give a different interpretation to the approximation ratio obtained above. Let us first slightly reformulate the problem (recall that $w_{ii} = 0$).

$$
\begin{aligned}
\max_{y_i=\pm1} \frac{1}{2}\sum_{i<j} w_{ij}(1-y_iy_j) &= \max_{y_i=\pm1} \frac{1}{4}\sum_{i,j} w_{ij}(1-y_iy_j)\\
&= \max_{y_i=\pm1} \frac{1}{4}\sum_{i,j} w_{ij}\left(\frac{y_i^2+y_j^2}{2}-y_iy_j\right)\\
&= \max_{y_i=\pm1} \frac{1}{4}\left(-\sum_{i,j} w_{ij}y_iy_j + \frac{1}{2}\sum_i\left[\sum_j w_{ij}\right]y_i^2 + \frac{1}{2}\sum_j\left[\sum_i w_{ij}\right]y_j^2\right)\\
&= \max_{y_i=\pm1} \frac{1}{4}\left(-\sum_{i,j} w_{ij}y_iy_j + \frac{1}{2}\sum_i \deg(i)y_i^2 + \frac{1}{2}\sum_j \deg(j)y_j^2\right)\\
&= \max_{y_i=\pm1} \frac{1}{4}\left(-\sum_{i,j} w_{ij}y_iy_j + \sum_i \deg(i)y_i^2\right)\\
&= \max_{y_i=\pm1} \frac{1}{4}y^TL_Gy,
\end{aligned}
$$

where $L_G = D_G - W$ is the Laplacian matrix, D_G is a diagonal matrix with $(D_G)_{ii} = \deg(i) = \sum_j w_{ij}$ and $W_{ij} = w_{ij}$.

This means that we rewrite (1) as

$$
\begin{array}{ll}
\max & \frac{1}{4}y^TL_Gy\\
 & y_i = \pm1,\ i = 1,\ldots,n.
\end{array}
\tag{7}
$$

Similarly, (3) can be written (by taking $X = yy^T$) as

$$
\begin{array}{ll}
\max & \frac{1}{4}\operatorname{Tr}(L_GX)\\
s.t. & X \succeq 0\\
 & X_{ii} = 1,\ i = 1,\ldots,n.
\end{array}
\tag{8}
$$

Indeed, given

Next lecture we derive the formulation of the dual program to (8) in the context of recovery in the Stochastic Block Model. Here we will simply show weak duality. The dual is given by

$$
\begin{array}{ll}
\min & \operatorname{Tr}(D)\\
s.t. & D \text{ is a diagonal matrix}\\
 & D - \frac{1}{4}L_G \succeq 0.
\end{array}
\tag{9}
$$

Indeed, if X is a feasible solution to (8) and D a feasible solution to (9) then, since X and $D-\frac{1}{4}L_G$ are both positive semidefinite $\operatorname{Tr}\left[X\left(D-\frac{1}{4}L_G\right)\right] \geq 0$ which gives

$$
0 \leq \operatorname{Tr}\left[X\left(D-\frac{1}{4}L_G\right)\right] = \operatorname{Tr}(XD) - \frac{1}{4}\operatorname{Tr}(L_GX) = \operatorname{Tr}(D) - \frac{1}{4}\operatorname{Tr}(L_GX),
$$

since D is diagonal and $X_{ii} = 1$. This shows weak duality, the fact that the value of (9) is larger than the one of (8).

If certain conditions, the so called Slater conditions [VB04, VB96], are satisfied then the optimal values of both programs are known to coincide, this is known as strong duality. In this case, the Slater conditions ask whether there is a matrix strictly positive definite that is feasible for (8) and the identity is such a matrix. This means that there exists $D^\natural$ feasible for (9) such that

$$\mathrm{Tr}(D^\natural) = \mathcal{R}\mathrm{MaxCut}.$$

Hence, for any $y \in \mathbb{R}^n$ we have

$$\frac{1}{4} y^T L_G y = \mathcal{R}\mathrm{MaxCut} - y^T \left(D^\natural - \frac{1}{4} L_G \right)^T + \sum_{i=1}^{n} D_{ii} \left(y_i^2 - 1 \right). \tag{10}$$

Note that (10) certifies that no cut of G is larger than $\mathcal{R}$MaxCut. Indeed, if $y \in \{\pm 1\}^2$ then $y_i^2 = 1$ and so

$$\mathcal{R}\mathrm{MaxCut} - \frac{1}{4} y^T L_G y = y^T \left(D^\natural - \frac{1}{4} L_G \right)^T.$$

Since $D^\natural - \frac{1}{4} L_G \succeq 0$, there exists V such that $D^\natural - \frac{1}{4} L_G = V V^T$ with the columns of V denoted by $v_1, \ldots, v_n$. This means that meaning that $y^T \left(D^\natural - \frac{1}{4} L_G \right)^T = \left\| V^T y \right\|^2 = \sum_{k=1}^{n} (v_k^T y)^2$. This means that, for $y \in \{\pm 1\}^2$,

$$\mathcal{R}\mathrm{MaxCut} - \frac{1}{4} y^T L_G y = \sum_{k=1}^{n} (v_k^T y)^2.$$

In other words, $\mathcal{R}\mathrm{MaxCut} - \frac{1}{4} y^T L_G y$ is, in the hypercube ($y \in \{\pm 1\}^2$) a sum-of-squares of degree 2. This is known as a sum-of-squares certificate [BS14, Bar14, Par00, Las01, Sho87, Nes00]; indeed, if a polynomial is a sum-of-squares naturally it is non-negative.

Note that, by definition, $\mathrm{MaxCut} - \frac{1}{4} y^T L_G y$ is always non-negative on the hypercube. This does not mean, however, that it needs to be a sum-of-squares[1] of degree 2.

(A Disclaimer: the next couple of paragraphs are a bit hand-wavy, they contain some of intuition for the Sum-of-squares hierarchy but for details and actual formulations, please see the references.)

The remarkable fact is that, if one bounds the degree of the sum-of-squares certificate, it can be found using Semidefinite programming [Par00, Las01]. In fact, SDPs (9) and (9) are finding the smallest real number Λ such that $\Lambda - \frac{1}{4} y^T L_G y$ is a sum-of-squares of degree 2 over the hypercube, the dual SDP is finding a certificate as in (10) and the primal is constraining the moments of degree 2 of y of the form $X_{ij} = y_i y_j$ (see [Bar14] for some nice lecture notes on Sum-of-Squares, see also Remark 8.4). This raises a natural question of whether, by allowing a sum-of-squares certificate of degree 4 (which corresponds to another, larger, SDP that involves all monomials of degree ≤ 4 [Bar14]) one can improve the approximation of α_{GW} to Max-Cut. Remarkably this is open.

Open Problem 8.2 1. *What is the approximation ratio achieved by (or the integrality gap of) the Sum-of-squares degree 4 relaxation of the Max-Cut problem?*

[1]This is related with Hilbert's 17th problem [Sch12] and Stengle's Positivstellensatz [Ste74]

2. *The relaxation described above (of degree 2) (9) is also known to produce a cut of* $1 - \mathcal{O}(\sqrt{\epsilon})$ *when a cut of* $1-\epsilon$ *exists. Can the degree 4 relaxation improve over this?*

3. *What about other (constant) degree relaxations?*

Remark 8.4 (triangular inequalities and Sum of squares level 4) *A (simpler) natural question is wether the relaxation of degree 4 is actually strictly tighter than the one of degree 2 for Max-Cut (in the sense of forcing extra constraints). What follows is an interesting set of inequalities that degree 4 enforces and that degree 2 doesn't, known as triangular inequalities.*

Since $y_i \in \{\pm 1\}$ *we naturally have that, for all* i, j, k

$$y_iy_j + y_jy_k + y_ky_i \geq -1,$$

this would mean that, for $X_{ij} = y_iy_j$ *we would have,*

$$X_{ij} + X_{jk} + X_{ik} \geq -1,$$

however it is not difficult to see that the SDP (8) *of degree 2 is only able to constraint*

$$X_{ij} + X_{jk} + X_{ik} \geq -\frac{3}{2},$$

which is considerably weaker. There are a few different ways of thinking about this, one is that the three vector u_i, u_j, u_k *in the relaxation may be at an angle of 120 degrees with each other. Another way of thinking about this is that the inequality* $y_iy_j + y_jy_k + y_ky_i \geq -\frac{3}{2}$ *can be proven using sum-of-squares proof with degree 2:*

$$(y_i + y_j + y_k)^2 \geq 0 \quad \Rightarrow \quad y_iy_j + y_jy_k + y_ky_i \geq -\frac{3}{2}$$

However, the stronger constraint cannot.

On the other hand, if degree 4 monomials are involved, (let's say $X_S = \prod_{s\in S} y_s$*, note that* $X_\emptyset = 1$ *and* $X_{ij}X_{ik} = X_{jk}$*) then the constraint*

$$\begin{bmatrix} X_\emptyset \\ X_{ij} \\ X_{jk} \\ X_{ki} \end{bmatrix} \begin{bmatrix} X_\emptyset \\ X_{ij} \\ X_{jk} \\ X_{ki} \end{bmatrix}^T = \begin{bmatrix} 1 & X_{ij} & X_{jk} & X_{ki} \\ X_{ij} & 1 & X_{ik} & X_{jk} \\ X_{jk} & X_{ik} & 1 & X_{ij} \\ X_{ki} & X_{jk} & X_{ij} & 1 \end{bmatrix} \succeq 0$$

implies $X_{ij} + X_{jk} + X_{ik} \geq -1$ *just by taking*

$$\mathbf{1}^T \begin{bmatrix} 1 & X_{ij} & X_{jk} & X_{ki} \\ X_{ij} & 1 & X_{ik} & X_{jk} \\ X_{jk} & X_{ik} & 1 & X_{ij} \\ X_{ki} & X_{jk} & X_{ij} & 1 \end{bmatrix} \mathbf{1} \geq 0.$$

Also, note that the inequality $y_iy_j + y_jy_k + y_ky_i \geq -1$ *can indeed be proven using sum-of-squares proof with degree 4 (recall that* $y_i^2 = 1$*):*

$$(1 + y_iy_j + y_jy_k + y_ky_i)^2 \geq 0 \quad \Rightarrow \quad y_iy_j + y_jy_k + y_ky_i \geq -1.$$

Interestingly, it is known [KV13] that these extra inequalities alone will not increase the approximation power (in the worst case) of (3).

8.4 The Grothendieck Constant

There is a somewhat similar remarkable problem, known as the Grothendieck problem [AN04, AMMN05]. Given a matrix $A \in \mathbb{R}^{n\times m}$ the goal is to maximize

$$\begin{array}{ll} \max & x^T A y \\ s.t. & x_i = \pm, \forall_i \\ s.t. & y_j = \pm, \forall_j \end{array} \qquad (11)$$

Note that this is similar to problem (1). In fact, if $A \succeq 0$ it is not difficult to see that the optimal solution of (11) satisfies $y = x$ and so if $A = L_G$, since $L_G \succeq 0$, (11) reduces to (1). In fact, when $A \succeq 0$ this problem is known as the little Grothendieck problem [AN04, CW04, BKS13].

Even when A is not positive semidefinite, one can take $z^T = [x^T\, y^T]$ and the objective can be written as

$$z^T \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} z.$$

Similarly to the approximation ratio in Max-Cut, the Grothendieck constant [Pis11] K_G is the maximum ratio (over all matrices A) between the SDP relaxation

$$\begin{array}{ll} \max & \sum_{ij} A_{ij} u_i^T v_j \\ s.t. & u_i \in \mathbb{R}^{n+m}, \|u_i\| = 1, \\ & v_j \in \mathbb{R}^{n+m}, \|v_j\| = 1 \end{array} \qquad (12)$$

and 11, and its exact value is still unknown, the best known bounds are available here [] and are $1.676 < K_G < \frac{\pi}{2\log(1+\sqrt{2})}$. See also page 21 here [F+14]. There is also a complex valued analogue [Haa87].

Open Problem 8.3 *What is the real Grothendieck constant K_G?*

8.5 The Paley Graph

Let p be a prime such that $p \cong 1 \mod 4$. The Paley graph of order p is a graph on p nodes (each node associated with an element of $\mathbb{Z}_p$) where (i, j) is an edge if $i - j$ is a quadratic residue modulo p. In other words, (i, j) is an edge is there exists a such that $a^2 \cong i - j \mod p$. Let $\omega(p)$ denote the clique number of the Paley graph of order p, meaning the size of its largest clique. It is conjectured that $\omega(p) \lesssim \text{pollywog}(n)$ but the best known bound is $\omega(p) \leq \sqrt{p}$ (which can be easily obtained). The only improvement to date is that, infinitely often, $\omega(p) \leq \sqrt{p} - 1$, see [BRM13].

The theta function of a graph is a Semidefinite programming based relaxation of the independence number [Lov79] (which is the clique number of the complement graph). As such, it provides an upper bound on the clique number. In fact, this upper bound for Paley graph matches $\omega(p) \leq \sqrt{p}$.

Similarly to the situation above, one can define a degree 4 sum-of-squares analogue to $\theta(G)$ that, in principle, has the potential to giving better upper bounds. Indeed, numerical experiments in [GLV07] seem to suggest that this approach has the potential to improve on the upper bound $\omega(p) \leq \sqrt{p}$

Open Problem 8.4 *What are the asymptotics of the Paley Graph clique number $\omega(p)$? Can the the SOS degree 4 analogue of the theta number help upper bound it?* [2]

[2] The author thanks Dustin G. Mixon for suggesting this problem.

Interestingly, a polynomial improvement on Open Problem 6.4. is known to imply an improvement on this problem [BMM14].

8.6 An interesting conjecture regarding cuts and bisections

Given d and n let $G^{reg}(n,d)$ be a random d-regular graph on n nodes, drawn from the uniform distribution on all such graphs. An interesting question is to understand the typical value of the Max-Cut such a graph. The next open problem is going to involve a similar quantity, the Maximum Bisection. Let n be even, the Maximum Bisection of a graph G on n nodes is

$$\mathrm{MaxBis}(G) = \max_{S\colon |S|=\frac{n}{2}} \mathrm{cut}(S),$$

and the related Minimum Bisection (which will play an important role in next lectures), is given by

$$\mathrm{MinBis}(G) = \min_{S\colon |S|=\frac{n}{2}} \mathrm{cut}(S),$$

A typical bisection will cut half the edges, meaning $\frac{d}{4}n$. It is not surprising that, for large n, $\mathrm{MaxBis}(G)$ and $\mathrm{MinBis}(G)$ will both fluctuate around this value, the amazing conjecture [ZB09] is that their fluctuations are the same.

Conjecture 8.5 ([ZB09]) *Let $G \sim G^{reg}(n,d)$, then for all d, as n grows*

$$\frac{1}{n}\left(\mathrm{MaxBis}(G) + \mathrm{MinBis}(G)\right) = \frac{d}{2} + o(1),$$

where $o(1)$ is a term that goes to zero with n.

Open Problem 8.5 *Prove or disprove Conjecture 8.5.*

Recently, it was shown that the conjecture holds up to $o(\sqrt{d})$ terms [DMS15]. We also point the reader to this paper [Lyo14], that contains bounds that are meaningful already for $d = 3$.

References

[ABS10] S. Arora, B. Barak, and D. Steurer. Subexponential algorithms for unique games related problems. 2010.

[AMMN05] N. Alon, K. Makarychev, Y. Makarychev, and A. Naor. Quadratic forms on graphs. *Invent. Math*, 163:486–493, 2005.

[AN04] N. Alon and A. Naor. Approximating the cut-norm via Grothendieck's inequality. In *Proc. of the 36 th ACM STOC*, pages 72–80. ACM Press, 2004.

[Ban15] A. S. Bandeira. Relax and Conquer BLOG: Ten Lectures and Forty-two Open Problems in Mathematics of Data Science. 2015.

[Bar14] B. Barak. Sum of squares upper bounds, lower bounds, and open questions. *Available online at* `http://www.boazbarak.org/sos/files/all-notes.pdf`, 2014.

[BKS13] A. S. Bandeira, C. Kennedy, and A. Singer. Approximating the little grothendieck problem over the orthogonal group. *Available online at arXiv:1308.5207 [cs.DS]*, 2013.

[BMM14] A. S. Bandeira, D. G. Mixon, and J. Moreira. A conditional construction of restricted isometries. *Available online at arXiv:1410.6457 [math.FA]*, 2014.

[BRM13] C. Bachoc, I. Z. Ruzsa, and M. Matolcsi. Squares and difference sets in finite fields. *Available online at arXiv:1305.0577 [math.CO]*, 2013.

[BS14] B. Barak and D. Steurer. Sum-of-squares proofs and the quest toward optimal algorithms. *Survey, ICM 2014*, 2014.

[CW04] M. Charikar and A. Wirth. Maximizing quadratic programs: Extending grothendieck's inequality. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '04, pages 54–60, Washington, DC, USA, 2004. IEEE Computer Society.

[DMS15] A. Dembo, A. Montanari, and S. Sen. Extremal cuts of sparse random graphs. *Available online at arXiv:1503.03923 [math.PR]*, 2015.

[F^+14] Y. Filmus et al. Real analysis in computer science: A collection of open problems. *Available online at* `http://simons.berkeley.edu/sites/default/files/openprobsmerged.pdf`, 2014.

[GLV07] N. Gvozdenovic, M. Laurent, and F. Vallentin. Block-diagonal semidefinite programming hierarchies for 0/1 programming. *Available online at arXiv:0712.3079 [math.OC]*, 2007.

[GW95] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefine programming. *Journal of the Association for Computing Machinery*, 42:1115–1145, 1995.

[Haa87] U. Haagerup. A new upper bound for the complex Grothendieck constant. *Israel Journal of Mathematics*, 60(2):199–224, 1987.

[Has02] J. Hastad. Some optimal inapproximability results. 2002.

[Kho02] S. Khot. On the power of unique 2-prover 1-round games. *Thiry-fourth annual ACM symposium on Theory of computing*, 2002.

[KKMO05] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. Optimal inapproximability results for max-cut and other 2-variable csps? 2005.

[KV13] S. A. Khot and N. K. Vishnoi. The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into l1. *Available online at arXiv:1305.4581 [cs.CC]*, 2013.

[Las01] J. B. Lassere. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.

[Lov79] L. Lovasz. On the shannon capacity of a graph. *IEEE Trans. Inf. Theor.*, 25(1):1–7, 1979.

[Lyo14] R. Lyons. Factors of IID on trees. *Combin. Probab. Comput.*, 2014.

[Nes00] Y. Nesterov. Squared functional systems and optimization problems. *High performance optimization*, 13(405-440), 2000.

[Par00] P. A. Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. PhD thesis, 2000.

[Pis11] G. Pisier. Grothendieck's theorem, past and present. *Bull. Amer. Math. Soc.*, 49:237–323, 2011.

[Rag08] P. Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, STOC '08, pages 245–254. ACM, 2008.

[Sch12] K. Schmudgen. Around hilbert's 17th problem. *Documenta Mathematica - Extra Volume ISMP*, pages 433–438, 2012.

[Sho87] N. Shor. An approach to obtaining global extremums in polynomial mathematical programming problems. *Cybernetics and Systems Analysis*, 23(5):695–700, 1987.

[Ste74] G. Stengle. A nullstellensatz and a positivstellensatz in semialgebraic geometry. *Math. Ann. 207*, 207:87–97, 1974.

[VB96] L. Vanderberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38:49–95, 1996.

[VB04] L. Vanderberghe and S. Boyd. *Convex Optimization*. Cambridge University Press, 2004.

[ZB09] L. Zdeborova and S. Boettcher. Conjecture on the maximum cut and bisection width in random regular graphs. *Available online at arXiv:0912.4861 [cond-mat.dis-nn]*, 2009.