

### 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  $w_{ii} = 0$ ).

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

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 (66) as

$$\begin{aligned}
\max_{y_i = \pm 1, i = 1, \dots, n} \frac{1}{4} y^T L_G y
\end{aligned} \tag{72}$$

Similarly, (68) can be written (by taking  $X = y y^T$ ) as

$$\begin{aligned}
\max \quad & \frac{1}{4} \text{Tr}(L_G X) \\
s.t. \quad & X \succeq 0 \\
& X_{ii} = 1, i = 1, \dots, n.
\end{aligned} \tag{73}$$

Indeed, given

Next lecture we derive the formulation of the dual program to (73) in the context of recovery in the Stochastic Block Model. Here we will simply show weak duality. The dual is given by

$$\begin{aligned}
\min \quad & \text{Tr}(D) \\
s.t. \quad & D \text{ is a diagonal matrix} \\
& D - \frac{1}{4} L_G \succeq 0.
\end{aligned} \tag{74}$$

Indeed, if  $X$  is a feasible solution to (73) and  $D$  a feasible solution to (74) then, since  $X$  and  $D - \frac{1}{4} L_G$  are both positive semidefinite  $\text{Tr} \left[ X \left( D - \frac{1}{4} L_G \right) \right] \geq 0$  which gives

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

since  $D$  is diagonal and  $X_{ii} = 1$ . This shows weak duality, the fact that the value of (74) is larger than the one of (73).

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 (73) and the identity is such a matrix. This means that there exists  $D^\natural$  feasible for (74) such that

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

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

$$\frac{1}{4}y^T L_G y = \mathcal{R}\text{MaxCut} - y^T \left( D^\natural - \frac{1}{4}L_G \right)^T + \sum_{i=1}^n D_{ii} (y_i^2 - 1). \quad (75)$$

Note that (75) certifies that no cut of  $G$  is larger than  $\mathcal{R}\text{MaxCut}$ . Indeed, if  $y \in \{\pm 1\}^2$  then  $y_i^2 = 1$  and so

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

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

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

In other words,  $\mathcal{R}\text{MaxCut} - \frac{1}{4}y^T L_G y$  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,  $\text{MaxCut} - \frac{1}{4}y^T L_G y$  is always non-negative on the hypercube. This does not mean, however, that it needs to be a sum-of-squares<sup>33</sup> 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 (74) and (74) are finding the smallest real number  $\Lambda$  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 (75) 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  $\leq 4$  [Bar14]) one can improve the approximation of  $\alpha_{GW}$  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?*

*2. The relaxation described above (of degree 2) (74) 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?*

---

<sup>33</sup>This is related with Hilbert's 17th problem [Sch12] and Stengle's Positivstellensatz [Ste74]

## References

- [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.
- [BS14] B. Barak and D. Steurer. Sum-of-squares proofs and the quest toward optimal algorithms. *Survey, ICM 2014*, 2014.
- [Las01] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.
- [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.
- [Sho87] N. Shor. An approach to obtaining global extremums in polynomial mathematical programming problems. *Cybernetics and Systems Analysis*, 23(5):695–700, 1987.
- [VB04] L. Vanderberghe and S. Boyd. *Convex Optimization*. Cambridge University Press, 2004.
- [VB96] L. Vanderberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38:49–95, 1996.

MIT OpenCourseWare  
<http://ocw.mit.edu>

18.S096 Topics in Mathematics of Data Science  
Fall 2015

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.