A convex quadratic characterization of the Lovász theta number

Carlos J. Luz† and Alexander Schrijver‡

†Escola Superior de Tecnologia de Setúbal/Instituto Politécnico de Setúbal, Campus do IPS, Estefanilha, 2914-508 Setúbal, Portugal, e-mail: cluz@est.ips.pt ‡CWI, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands, email: lex@cwi.nl

Abstract

In previous works an upper bound on the stability number $\alpha(G)$ of a graph G based on convex quadratic programming was introduced and several of its properties were established. The aim of this investigation is to relate theoretically this bound (usually represented by v(G)) with the well known Lovász $\vartheta(G)$ number. To begin with, a new set of convex quadratic bounds on $\alpha(G)$ that generalize and improve the bound v(G) is proposed. Then it is proved that $\vartheta(G)$ is never worse than any bound belonging to this set of new bounds. The main result of this note states that one of these new bounds equals $\vartheta(G)$, a fact that leads to a new characterization of the Lovász theta number.

Keywords: Combinatorial Optimization, Graph Theory, Maximum Stable Set, Quadratic Programming.

1 Introduction

Let G = (V, E) be a simple undirected graph where $V = \{1, ..., n\}$ denotes the vertex set and E is the edge set. It will be supposed that G has at least one edge, i.e., E is not empty. We will write $ij \in E$ to denote the edge linking nodes i and j of V. The adjacency matrix $A_G = [a_{ij}]$ of G is defined by

$$a_{ij} = \left\{ \begin{array}{ll} 1 & \text{if } ij \in E \\ 0 & \text{if } ij \notin E \end{array} \right..$$

A stable set (independent set) of G is a subset of nodes of V whose elements are pairwise nonadjacent. The stability number (or independence number) of G is defined as the cardinality of a largest stable set and is usually denoted by $\alpha(G)$. A maximum stable set of G is a stable set with $\alpha(G)$ nodes. The problem of finding $\alpha(G)$ is NP-hard and thus it is suspected that it cannot be solved in polynomial time. Even, there exists $\epsilon > 0$ such that to approximate $\alpha(G)$ within a ratio of $n^{-\epsilon}$ is NP-hard (see [1]). However, several ways of approaching $\alpha(G)$ have been proposed in the literature (see for example [2, 6, 9, 16] and [3] for a survey).

For any graph G with at least one edge, it can easily be proved (see proposition 2.1 below) that $\alpha(G) \leq v(G)$, where v(G) is the optimal value of the following convex quadratic programming problem,

$$(P_G)$$
 $v(G) = \max\{2e^T x - x^T (H+I) x : x \ge 0\}.$

Here and hereinafter e is the $n \times 1$ all ones vector, T stands for the transposition operation, I is the identity matrix of order n and

$$H = \frac{1}{-\lambda_{\min}(A_G)} A_G,$$

where A_G is the adjacency matrix of G and $\lambda_{\min}(A_G)$ is its smallest eigenvalue. As the trace of A_G is zero and G has at least one edge, A_G is indefinite (see [5] for details). Thus $\lambda_{\min}(H) = -1$ and this guarantees the convexity of P_G because H + I is positive semidefinite.

Having in mind the nice properties of v(G) (see [14, 15]), the initial aim of this investigation was to relate theoretically v(G) with the well known Lovász $\vartheta(G)$ number introduced in [12] and discussed in many publications [8, 9, 10, 11, 13]. As a consequence of this effort, a new set of convex quadratic bounds on $\alpha(G)$ that generalize and improve the v(G) bound is now introduced. Also, it is shown that $\vartheta(G)$ is never worse than any bound belonging to this set of new bounds. The main result herein proved states that $\vartheta(G)$ is equal to the best bound in this set. In consequence, it leads to a characterization of $\vartheta(G)$ by convex quadratic programming.

This note is organized as follows. In section 2 the new family of upper bounds on $\alpha(G)$ is introduced and some of its properties are presented. In section 3, some different $\vartheta(G)$ formulations are recalled and the results relating the new introduced bounds with $\vartheta(G)$ are established in section 4.

2 Generalizing the v(G) bound

In order to improve the upper bound v(G) we define the following family of quadratic problems which are based on a perturbation in the Hessian of the convex quadratic programming problem P_G :

$$(P_G(C))$$
 $v(G,C) = \max\{2e^T x - x^T (H_C + I) x : x \ge 0\},$

where $C = [c_{ij}]$ is a non null real symmetric matrix such that $c_{ij} = 0$ if i = j or $ij \notin E$ and

$$H_C = \frac{C}{-\lambda_{\min}(C)},$$

denoting $\lambda_{\min}(C)$ the smallest eigenvalue of C. Any matrix satisfying the conditions imposed to matrix C will be called a weighted adjacency matrix of G. Note that as well as the adjacency matrix A_G , the matrix C is indefinite taking into account that its trace is null and not all entries c_{ij} are null. Consequently, since $\lambda_{\min}(H_C) = -1$, all problems $P_G(C)$ are convex. Note also that $v(G, A_G) = v(G)$ and thus P_G is included in the introduced family of quadratic problems.

Some basic facts about the $P_G(C)$ family of problems are given below.

Proposition 2.1 For any weighted adjacency matrix C of a graph G, the number v(G, C) is the optimal value of a convex quadratic problem and verifies $\alpha(G) \leq v(G, C)$, i.e., v(G, C) is an upper bound on $\alpha(G)$.

Proof. As $\lambda_{\min}(H_C) = -1$, the problem $P_G(C)$ is convex quadratic as stated. To see that v(G, C) is an upper bound on $\alpha(G)$ for all matrices C, let x be a characteristic vector of any maximum independent set S of G (defined by $x_i = 1$ if $i \in S$ and $x_i = 0$ otherwise). Since the vector x is a feasible solution of $P_G(C)$ and verifies $x^T H_C x = 0$ (note that $x_i x_j = 0$ if $ij \in E$), we have

$$v(G, C) \ge 2e^T x - x^T x - x^T H_C x = 2\alpha(G) - \alpha(G) = \alpha(G),$$

i.e., $\alpha(G) \leq v(G, C)$, for all weighted adjacency matrices C of G.

A clique of the graph G=(V,E) is any subset of V such that the induced subgraph is complete. A minimum clique cover of G is a set of cliques of G that cover V with the least cardinality. This minimum number of cliques can be denoted by $\bar{\chi}(G)$ and, like the stability number, it is NP-hard to compute $\bar{\chi}(G)$. The partial graph associated with a minimum clique cover of G is a graph with the same set of vertices as that of G, and whose edges are those of the complete subgraphs induced by the cliques forming the clique cover.

Proposition 2.2 Let G be a graph with at least one edge. If M is the adjacency matrix of the partial graph associated with a minimum clique cover of G, then $v(G, M) \leq \bar{\chi}(G)$.

Proof. Suppose that $\bar{\chi}(G) = k$ and denote by G_i , i = 1, ..., k, the complete subgraphs induced by the cliques forming a minimum clique cover of G. Let x be an optimal solution of $P_G(M)$, where M is the adjacency matrix of the partial graph associated with this minimum clique cover. Note that $\lambda_{\min}(M) = -1$ since M + I is formed by k all ones blocks on the diagonal (say $J_1, ..., J_k$), these blocks are positive semidefinite and any J_i -block of size at least two has a zero eigenvalue. Thus

$$v(G, M) = 2e^{T}x - x^{T}(M+I)x = \sum_{i=1}^{k} 2e_{i}^{T}x_{i} - x_{i}^{T}J_{i}x_{i},$$

where, for each i, e_i and x_i are respectively the subvectors of e and x whose components correspond to the vertices of G_i . As $J_i = e_i e_i^T$ and $\left(e_i^T x_i - 1\right)^2 \ge 0$, we have $2e_i^T x_i - x_i^T J_i x_i \le 1$ for all i, hence $v(G, M) \le k$, as required. \blacksquare

Note that for any graph G with at least one edge that satisfies $\alpha(G) = \bar{\chi}(G)$ (in particular for perfect graphs), the propositions 2.1 and 2.2 allow to define $\alpha(G)$ as follows:

$$\alpha(G) = \min_{C} v(G, C),$$

where C is a weighted adjacency matrix of G.

3 The Lovász $\vartheta(G)$ number

The Lovász $\vartheta(G)$ number was introduced in [12] and has been subsequently studied in several publications. It is generally considered the most famous upper bound on $\alpha(G)$, for which various different formulations were established in the literature (see [9, 11]). Some of these formulations are now recalled.

An orthonormal representation of a graph G = (V, E) with $V = \{1, 2, ..., n\}$ is a set of unit vectors $u_1, u_2, ..., u_n$ in a Euclidean space, which are orthogonal (i.e., $u_i^T u_j = 0$) whenever $ij \notin E$. Note that the vectors dimension is not fixed and that any graph has an orthonormal representation, considering for example a set of pairwise orthonormal vectors.

Lovász defined his theta number as follows:

$$\vartheta(G) = \min_{\substack{c, u_1, u_2, \dots, u_n \\ c \text{ unitary}}} \max_{i \in V} \frac{1}{(c^T u_i)^2},\tag{1}$$

where the minimum is taken over all vectors c with ||c|| = 1 and all orthonormal representations u_1, u_2, \ldots, u_n of G.

As mentioned before, the inequality $\alpha(G) \leq \bar{\chi}(G)$ holds true for any graph G. Both of these numbers are NP-hard to compute but they "sandwich" the number $\vartheta(G)$ which can be computed in polynomial time as proved by Grötschel, Lovász and Schrijver [7]. That is,

$$\alpha(G) \le \vartheta(G) \le \bar{\chi}(G),$$

a fact known as the Lovász's sandwich theorem (see [11]).

The paper [12] gives several characterizations of $\vartheta(G)$. One of them is the following:

$$\vartheta(G) = \min_{\Delta} \lambda_{\max}(A)$$

where $\lambda_{\max}(A)$ denotes the largest eigenvalue of A, and the minimum is taken over the set of all symmetric matrices $A = [a_{ij}] \in \mathbb{R}^{n \times n}$ such that $a_{ij} = 1$ if i = j or $ij \notin E$. Since we are assuming that G has at least one edge, we can eliminate the matrix ee^T from this set. In fact, if $\vartheta(G) = \lambda_{\max}(ee^T) = n$, then $\bar{\chi}(G) = n$ (recall the "sandwich" theorem) and thus G would have no edge.

Let A be one of the above symmetric matrices. As $A \neq ee^T$ we have that $Q = A - ee^T \neq 0$ is a weighted adjacency matrix of G. Consequently, setting $A = ee^T + Q$, $\vartheta(G)$ can be formulated as follows:

$$\vartheta(G) = \min_{Q} \lambda_{\max}(ee^{T} + Q), \tag{2}$$

where Q is a weighted adjacency matrix of G.

Another characterization of ϑ which is dual of (2) is the following (see [12]):

$$\vartheta(G) = \max_{\mathcal{B}} e^T B e,\tag{3}$$

where $B = [b_{ij}] \in \mathbb{R}^{n \times n}$ ranges over all positive semidefinite symmetric matrices such that $b_{ij} = 0$ for $ij \in E$ and Tr(B) = 1. (Tr(B) denotes the trace of B.)

4 Relating $\vartheta(G)$ and $\upsilon(G,C)$

In this section we relate $\vartheta(G)$ with the convex quadratic upper bounds v(G,C).

Theorem 4.1 Let G be a graph with at least one edge. Then for any weighted adjacency matrix C of graph G, we have $\vartheta(G) \leq \upsilon(G, C)$.

Proof. Let $C = [c_{ij}]$ be a weighted adjacency matrix of G = (V, E) and suppose that $P_G(C)$ is not unbounded for otherwise the theorem is true.

Let x be an optimal solution of $P_G(C)$. The Karush-Kuhn-Tucker conditions applied to this problem guarantee that the following conditions are true:

$$x \ge 0$$
, $(H_C + I) x \ge e$ and $x^T (H_C + I) x = e^T x = v(G, C)$. (4)

As $H_C + I$ is positive semidefinite we can write $H_C + I = U^T U$. Thus the columns of U can be thought of as an orthonormal representation of G.

Define $c = v^{-1/2}Ux$ where v abbreviates v(G, C). Then by (4), $c^Tc = v^{-1}x^T(H_C + I)x = 1$ and

$$U^T c = v^{-1/2} U^T U x > v^{-1/2} e.$$

This inequality implies

$$\frac{1}{(u_i^T c)^2} \le v$$
, for each i ,

where u_i denotes the column i of U. Recalling (1) we have $\vartheta(G) \leq \upsilon(G,C)$ as desired.

This theorem asserts that $\vartheta(G)$ is not worse that any v(G,C) bound. So, in particular, the inequality $\vartheta(G) \leq v(G)$ is always true. However there are many graphs for which the value of v(G) equals $\vartheta(G)$. In fact, it was proved in [4] that there is an infinite number of graphs that verify $\alpha(G) = v(G)$ and hence $\vartheta(G) = v(G)$. These graphs constitute the so called class of graphs with convex-QP stability number (one member of this class can be constructed by considering L(L(G)), where L(G) is the line graph of a connected graph G with an even number of edges).

We state now the main result of this note which gives the announced characterization of $\vartheta(G)$ by convex quadratic programming.

Theorem 4.2 Let G be a graph with at least one edge. If Q attains the optimum in (2) then $\vartheta(G) = \upsilon(G, C)$, where C = -Q.

Consequently, the following characterization of $\vartheta(G)$ is valid:

$$\vartheta(G) = \min_{C} \upsilon(G, C) = \min_{C} \max_{x>0} \left\{ 2e^{T}x - x^{T}(H_{C} + I)x \right\}, \tag{5}$$

where C is a weighted adjacency matrix of G.

Proof. Let Q be a weighted adjacency matrix of G attaining the optimum in (2) and let C = -Q. As $\vartheta(G) = \lambda_{\max}(ee^T + Q) \ge \lambda_{\max}(Q)$, we will divide the proof of the equality $\vartheta(G) = \upsilon(G, C)$ in two cases. (To simplify the notation we will sometimes use ϑ instead of $\vartheta(G)$.)

Case 1: $\vartheta(G) = \lambda_{\max}(Q)$.

Let x attain the optimum in $P_G(C)$. Then, using the positive semidefiniteness of $I-\vartheta^{-1}(ee^T+Q)$, we have

$$v(G,C) = 2e^{T}x - x^{T} (H_{C} + I) x = 2e^{T}x - x^{T} \left(\frac{-Q}{-\lambda_{\min}(-Q)} + I\right) x$$

$$= 2e^{T}x - x^{T} \left(\frac{-Q}{\lambda_{\max}(Q)} + I\right) x$$

$$= 2e^{T}x - x^{T} (I - \vartheta^{-1}Q + \vartheta^{-1}ee^{T}) x - \vartheta^{-1} (e^{T}x)^{2}$$

$$= 2e^{T}x - x^{T} [I - \vartheta^{-1} (ee^{T} + Q)] x - \vartheta^{-1} (e^{T}x)^{2}$$

$$\leq 2e^{T}x - \vartheta^{-1} (e^{T}x)^{2} \leq \vartheta,$$

since $(\vartheta^{1/2} - \vartheta^{-1/2}e^Tx)^2 \ge 0$. So by theorem 4.1, we have $\vartheta(G) = \upsilon(G, C)$ for this case.

Case 2: $\vartheta(G) > \lambda_{\max}(Q)$.

Let B attain the optimum in (3). Since $\vartheta I - ee^T - Q$ and B are positive semidefinite, we have

$$0 \le \operatorname{Tr} \left[B(\vartheta I - ee^T - Q) \right] = \vartheta \operatorname{Tr}(B) - \operatorname{Tr}(Bee^T) - \operatorname{Tr}(BQ) = \vartheta - \vartheta - 0 = 0.$$

So $\text{Tr}\left[B(\vartheta I-ee^T-Q)\right]=0$ and then $B(ee^T+Q-\vartheta I)=0$, i.e., the column space of B is orthogonal to the column space of $\vartheta I-ee^T-Q$. (In fact, if M and N are positive semidefinite matrices and Tr(MN)=0, then MN=0. To see this, let $M=U^TU$ and $N=W^TW$. Then $0=\text{Tr}(MN)=\text{Tr}(U^TUW^TW)=\text{Tr}(WU^TUW^T)$. Since WU^TUW^T is positive semidefinite, it implies that $UW^T=0$, hence MN=0.)

The inequality $\vartheta(G) > \lambda_{\max}(Q)$ implies that $\lambda_{\min}(\vartheta I - Q) > 0$ and hence $\operatorname{rank}(\vartheta I - Q) = n$. Then $\operatorname{rank}(\vartheta I - ee^T - Q) \ge n - 1$ and by the column spaces orthogonality, $\operatorname{rank}(B) \le 1$. As $\operatorname{Tr} B = 1$, $\operatorname{rank}(B) = 1$, and then $B = \vartheta^{-1}xx^T$ for some vector x whose support is a stable set S. Since $e^T B e = \vartheta$ and $\operatorname{Tr} B = 1$, we can choose $x \ge 0$ and thus we have $e^T x = x^T x = \vartheta$. Additionally, x is a characteristic vector of S. (To see this, let y be the characteristic vector of S. Then $y^T x = e^T x = \vartheta$ and, by the Cauchy-Schwarz inequality, $(y^T x)^2 \le (x^T x)(y^T y)$. So $\vartheta \le |S|$ and by the maximality of ϑ , we have $y^T y = \vartheta$. Hence, the Cauchy-Schwarz inequality is satisfied with equality and this implies x = y.)

Using once more the orthogonality of the column spaces of B and $\vartheta I - ee^T - Q$, we conclude that $\left(ee^T + Q\right)x = \vartheta x$, and hence $-Qx = \vartheta(e-x)$. Then x satisfies the Karush-Kuhn-Tucker conditions associated to $P_G(C)$ (recall (4)) as:

• $x \ge 0$;

- $(H_C + I)x = \left(\frac{-Q}{\lambda_{\max}(Q)} + I\right)x = \frac{-Qx}{\lambda_{\max}(Q)} + x = \frac{\vartheta}{\lambda_{\max}(Q)}(e x) + x \ge e$, since $\vartheta \ge \lambda_{\max}(Q)$; and
- $x^T(H_C+I)x = x^Tx = e^Tx = \vartheta$, since x is a characteristic vector of a stable set.

Consequently, by the positive semidefiniteness of $H_C + I$, $\vartheta(G) = \upsilon(G, C)$ is also true for case 2. Finally, the proved equality and the definition of Q imply the characterization (5).

Acknowledgements The first author thanks Domingos Cardoso and Carlos Sarrico for the stimulating discussions and valuable suggestions made to a previous version of this note. He is also grateful to the UI&D "Centro de Estudos de Optimização e Controlo" of Aveiro University for the support given to this research.

References

- [1] Arora, S., Lund, C., Motwani, R., Sudan, M, and Szegedy, M., Proof verification and hardness of approximation problems, *Proc. of 33rd IEEE Symp. on Foundations of Computer Science*, IEEE Computer Science Press, Los Alamitos, CA, 14–23, 1992.
- [2] Berge, C., Graphs, North-Holland, Amsterdam, 1991.
- [3] Bomze, I.M., Budinich, M., Pardalos, P.M. and Pelillo, M., *The Maximum Clique Problem*, in D.Z. Du and P. M. Pardalos (Eds.), Handbook of Combinatorial Optimization, suppl. Vol. A, 1–74, Kluwer Academic Publishers, 1999.
- [4] Cardoso, D. M., Convex Quadratic Programming Approach to the Maximum Matching Problem, Journal of Global Optimization, 19, 291–306, 2001.
- [5] Cvetkovic, D., Doob, M. and Sachs, H., Spectra of Graphs, Theory and Applications, VEB Deutscher Verlag der Wissenschaften, Berlin (D.D.R.), 1979.
- [6] Gibbons, L. E., Hearn, D. W., Pardalos, P. M. and Ramana, M. V., Continuous characterizations of the maximum clique problem, Mathematics of Operations Research, 22, 754–768, 1997.
- [7] Grötschel, M., Lovász, L. and Schrijver, A., The Ellipsoid Method and its consequences is Combinatorial Optimization, Combinatorica, 1, 169–197, 1981.
- [8] Grötschel, M., Lovász, L. and Schrijver, A., *Relaxations of Vertex Packing*, Journal of Combinatorial Theory, Series B 40, 330–343, 1986.
- [9] Grötschel, M., Lovász, L. and Schrijver, A., Geometric Algorithms and Combinatorial Optimization, Springer, Berlin, 1988.
- [10] Helmberg, C., Semidefinite Programming for Combinatorial Optimization, Konrad-Zuse-Zentrum für Informationstechnik, Berlin, 2000.
- [11] Knuth, D. E., The Sandwich Theorem, Electronic Journal of Combinatorics, 1: Article 1, 1994.
- [12] Lovász, L., On the Shannon capacity of a graph, IEEE Transactions on Information Theory, 25(2), 1–7, 1979.
- [13] Lovász, L. and Schrijver, A., Cones of matrices and set-functions and 0-1 optimization, SIAM Journal on Optimization, 1(2), 166–190, 1991.
- [14] Luz, Carlos J., An upper bound on the independence number of a graph computable in polynomial time, Operations Research Letters, 18, 139–145, 1995.

- [15] Luz, C. J. and Cardoso, D. M., A Generalization of the Hoffman-Lovász Upper Bound on the Independence Number of a Regular Graph, Annals of Operations Research, 81, 307–319, 1998.
- [16] Motzkin, T. S. and Straus, E. G., Maxima for graphs and a new proof of a theorem of Turán, Canadian Journal of Mathematics, 17, 533–540, 1965.