On the Lovász Theta Function and Some Variants

Laura Galli* Adam N. Letchford†

March 2017. To appear in *Discrete Optimization*.

Abstract

The Lovász theta function of a graph is a well-known upper bound on the stability number. It can be computed efficiently by solving a semidefinite program (SDP). Actually, one can solve either of two SDPs, one due to Lovász and the other to Grötschel *et al.* The former SDP is often thought to be preferable computationally, since it has fewer variables and constraints. We derive some new results on these two equivalent SDPs. The surprising result is that, if we weaken the SDPs by aggregating constraints, or strengthen them by adding cutting planes, the equivalence breaks down. In particular, the Grötschel *et al.* scheme typically yields a stronger bound than the Lovász one.

Keywords: stable set problem; semidefinite programming; Lovász theta function

1 Introduction

Consider an undirected graph G with vertex set V and edge set E. A set of pairwise non-adjacent (respectively, adjacent) vertices is called a *stable* or *independent* set (resp. *clique*). Given a graph G, the maximum cardinality of a stable set (resp. clique) in G is called the *stability number* (resp. *clique number*) and denoted by $\alpha(G)$ (resp. $\omega(G)$). The problem of determining $\alpha(G)$ (resp. $\omega(G)$) is called the *maximum stable set problem* or MSSP (resp. *maximum clique problem* or MCP).

The *complement* of G, denoted by $\bar{G}$, is the graph with vertex set V and edge set

$$\bar{E} = \big\{\{i,j\} \subset V : \{i,j\} \notin E\big\}.$$

Since $\alpha(G) = \omega(\bar{G})$, the MSSP and MCP are equivalent. They have a wide range of applications in operational research and computer science.

*Dipartimento di Informatica, Università di Pisa, Largo B. Pontecorvo 3, 56127 Pisa, Italy. E-mail: laura.galli@unipi.it

†Department of Management Science, Lancaster University, Lancaster LA1 4YW, United Kingdom. E-mail: A.N.Letchford@lancaster.ac.uk

Unfortunately, they are $\mathcal{NP}$-hard in the strong sense, and hard even to approximate (Håstad [15]).

In his seminal paper, Lovász [18] defined the so-called *theta function* of a graph G, denoted by $\vartheta(G)$. He then proved that $\alpha(G) \leq \vartheta(G) \leq \chi(\bar{G})$, where $\chi(\bar{G})$ denotes the chromatic number of $\bar{G}$. Grötschel *et al.* [12] showed that $\vartheta(G)$ can be computed in polynomial time (to arbitrary fixed precision) by solving a semidefinite program (SDP). This result is remarkable, given that computing the chromatic number of a graph, or even approximating it, is also $\mathcal{NP}$-hard in the strong sense (Feige & Killian [8]).

In fact, there are two alternative SDP characterisations of $\vartheta(G)$ [13, 19]. The one in [19], having fewer variables and constraints, seems preferable computationally. Gruber & Rendl [14] showed how to map any feasible solution to the larger SDP onto a feasible solution of the smaller SDP with no smaller profit. Yildirim & Fan [26] gave a partial mapping in the reverse direction. In this paper, we give a complete mapping in both directions, and then extend it to the weighted case, in which each node has a positive weight and the goal is to find a maximum weight stable set.

One can define stronger variants of the theta function by adding cutting planes to the SDPs (see, e.g., [6, 9, 14, 19, 25]). We show that, perhaps surprisingly, the two SDPs behave differently in this context. In particular, adding a cutting plane to the smaller SDP can lead to less bound improvement than adding the analogous inequality to the larger SDP.

Going in the opposite direction, Lieder *et al.* [17] recently proposed to *weaken* the theta function, by aggregating equations in the larger SDP. We show that here, too, the two SDPs behave differently.

The paper is structured as follows. Section 2 is a literature review. Section 3 presents the results on the theta function and its weighted version. Section 4 presents the results on strengthened and weakened variants of the theta function. Some computational results are presented in Section 5. Concluding remarks are made in Section 6.

Throughout the paper, n will denote $|V|$, m will denote $|E|$, e will denote the all-ones vector of dimension n and $J = ee^T$ will denote the all-ones square matrix of order n. The trace of a square matrix M will be denoted by $\mathrm{Tr}(M)$. Given two symmetric matrices A, B of the same order, $A \bullet B$ will denote $\mathrm{Tr}(AB)$, the usual inner product of A and B.

We will also use the following facts (see, e.g., [13]). A symmetric matrix $M \in \mathbb{R}^{n \times n}$ is *positive semidefinite* (psd) if $v^T M v \geq 0$ for all $v \in \mathbb{R}^n$. The set of psd matrices of order n forms a convex cone in $\mathbb{R}^{n \times n}$, which is denoted by S^n_+. If M is psd, there exists a matrix $Y \in \mathbb{R}^{n \times n}$ such that $M = Y^T Y$ (Cholesky factorisation). If we let y_i denote the ith column of Y, then $M_{ij} = y_i^T y_j$ for all i and j. The vectors $y_1, \ldots, y_n$ form the so-called *Gram* *representation* of M. Finally, an augmented matrix of the form

$$\begin{pmatrix} 1 & v^T \\ v & M \end{pmatrix},$$

where $v \in \mathbb{R}^n$ is psd if and only if the matrix $M - vv^T$ is psd (Schur complement).

2 Literature Review

Now we review the relevant literature. For simplicity of notation, we assume throughout this section that G contains no isolated nodes.

2.1 LP-based bounds

The standard 0-1 LP formulation of the MSSP is constructed as follows (e.g., [13, 21]). For each $i \in V$, let x_i be a binary variable, taking the value 1 if and only if i is in the stable set. The formulation is then:

$$\begin{array}{rll} \max & e^T x & \\ \text{s.t.} & x_i + x_j \leq 1 & (\forall \{i,j\} \in E) \\ & x_i \in \{0,1\} & (\forall i \in V). \end{array} \tag{1}$$

To strengthen the LP relaxation one can add valid linear inequalities, such as the following ones, due to Padberg [21]:

- *clique* inequalities, which take the form $\sum_{i \in C} x_i \leq 1$, where C is a maximal clique in G;
- *odd hole* inequalities, which take the form $\sum_{i \in H} x_i \leq \left\lfloor \frac{|H|}{2} \right\rfloor$, where H is a set of nodes that induce an 'odd hole' (chordless odd cycle) in G.

Other known inequalities are surveyed, e.g., in [3, 9, 13, 19].

The upper bound obtained if one uses all clique (and non-negativity) inequalities is called the *fractional clique covering number* and denoted by $\chi^f(\bar{G})$. By definition, $\alpha(G) \leq \chi^f(\bar{G}) \leq \chi(\bar{G})$. Unfortunately, computing $\chi^f(\bar{G})$ is also $\mathcal{NP}$-hard [13]. Nevertheless, reasonably good computational results have been obtained using LP relaxations (e.g., [1, 4, 10, 11, 20, 23, 24]).

2.2 The theta function

A rather different upper bound on $\alpha(G)$ was derived by Lovász [18]. Suppose that a vector $x \in \{0,1\}^n$ is feasible for the 0-1 LP formulation presented in the previous subsection. Define the matrix $Z = xx^T/(e^T x) \in [0,1]^{n \times n}$. By definition, Z is psd and $J \bullet Z = (e^T x)^2/(e^T x) = e^T x$. Moreover, $\mathrm{diag}(Z) = x/(e^T x)$, which implies that $\mathrm{Tr}(Z) = 1$. This leads naturally to the following SDP relaxation of the MSSP, which we will call '`SDP1`':

$$\begin{aligned} \max \quad & J \bullet Z && (2)\\ \text{s.t.} \quad & Z_{ij} = 0 \qquad (\forall \{i,j\} \in E) && (3)\\ & \mathrm{Tr}(Z) = 1 && (4)\\ & Z \in \mathcal{S}_+^n. && (5) \end{aligned}$$

Lovász called the resulting upper bound $\vartheta(G)$, and proved that $\alpha(G) \le \vartheta(G) \le \chi(\bar{G})$. Nowadays, $\vartheta(G)$ is called the *Lovász theta function*.

Shortly after the publication of [18], it was shown that SDPs can be solved in polynomial time (to arbitrary fixed precision) [12]. Thus, $\vartheta(G)$ can be computed efficiently, at least in theory.

Another characterisation of $\vartheta(G)$ was derived by Grötschel *et al.* [13]. Given a feasible vector $x \in \{0,1\}^n$ as before, consider the matrix $X = xx^T \in \{0,1\}^{n \times n}$. By definition, X is psd, its main diagonal is equal to x, and it has zero entries for all edges. Moreover, the augmented matrix

$$X^+ = \begin{pmatrix} 1 \\ x \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix}^T = \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix}$$

is also psd. This leads to the following SDP relaxation, which we call '`SDP2`':

$$\begin{aligned} \max \quad & e^T x && (6)\\ \text{s.t.} \quad & X_{ii} = x_i \qquad (\forall i \in V) && (7)\\ & X_{ij} = 0 \qquad (\forall \{i,j\} \in E) && (8)\\ & \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \in \mathcal{S}_+^{n+1}. && (9) \end{aligned}$$

The resulting upper bound is again $\vartheta(G)$.

Note that `SDP2` has $n+m$ linear constraints, whereas `SDP1` has only $m+1$. On the other hand, `SDP2` has a very nice property [13]: if one projects its feasible region into x-space, the resulting convex set satisfies all clique and non-negativity inequalities. This implies that

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

with equality if G is perfect.

The relationship between `SDP1` and `SDP2` is more complicated than it might appear. Gruber & Rendl [14] proved the following. Let (x^*, X^*) be any feasible solution to `SDP2`, and let $\gamma = e^T x^*$ denote its profit. If $\gamma > 0$, then one can obtain a feasible solution Z^* to `SDP1` whose profit $J \bullet Z^*$ is no smaller than γ by setting Z_{ij} to X^*_{ij}/γ for all i and j. Yildirim & Fan

[26] proved a partial result in the other direction. Let Z^* be any feasible solution to `SDP1`, again with positive profit γ. Then there exists a feasible solution (x^*, X^*) to `SDP2` whose profit is no smaller than γ, in which

$$\begin{aligned} x_i^* &= \frac{\left(\sum_{j=1}^n Z_{ij}^*\right)^2}{\gamma Z_{ii}^*} \text{ if } Z_{ii}^* > 0 \\ &= \quad 0 \qquad \text{otherwise.} \end{aligned}$$

We remark that a third SDP formulation of $\vartheta(G)$, with $n + |\bar{E}|$ constraints, was given by Dukanovic & Rendl [6]. This is preferable when G is very dense.

2.3 The weighted theta function

Now consider the weighted version of the stable set problem, mentioned in the introduction. For each vertex $i \in V$, let $w_i > 0$ be the associated weight. Adapting `SDP2` to this case is trivial: just change the objective function to $w^T x$. Grötschel *et al.* [12] showed that to adapt `SDP1`, it is necessary to replace the objective function with

$$\sum_{i \in V} \sum_{j \in V} \sqrt{w_i w_j} Z_{ij}. \tag{10}$$

The resulting bound, the weighted theta function, is denoted by $\vartheta(G, w)$. A proof that both SDPs yield the same bound can be found in [13].

2.4 Variants of the theta function

As mentioned in the introduction, it is possible to strengthen both `SDP1` and `SDP2` by adding cutting planes (i.e., valid linear inequalities). Schrijver [25] proposed to strengthen `SDP1` simply by adding the non-negativity inequalities

$$Z_{ij} \geq 0 \qquad (\forall \{i,j\} \in \bar{E}). \tag{11}$$

The resulting bound is denoted by $\vartheta'(G)$ [13].

Dukanovic & Rendl [6] proposed to add, in addition,

$$Z_{ik} + Z_{jk} \leq Z_{kk} \qquad (\forall \{i,j\} \in E,\, k \in V \setminus \{i,j\}) \tag{12}$$

$$Z_{ik} + Z_{jk} \leq Z_{ij} + Z_{kk} \quad (\forall \text{ stable } \{i,j,k\}). \tag{13}$$

As for `SDP2`, Lovász & Schrijver [19] proposed to add

$$X_{ij} \geq 0 \qquad (\forall \{i,j\} \in \bar{E}) \tag{14}$$

$$X_{ik} + X_{jk} \leq x_k \qquad (\forall \{i,j\} \in E,\, k \neq i,j) \tag{15}$$

$$x_i + x_j + x_k \leq 1 + X_{ik} + X_{jk} \quad (\forall \{i,j\} \in E,\, k \neq i,j). \tag{16}$$

Gruber & Rendl [14] proposed to add, in addition,

$$X_{ik} + X_{jk} \le x_k + X_{ij} \qquad (\forall \text{ stable } \{i,j,k\}) \tag{17}$$
$$x_i + x_j + x_k \le 1 + X_{ij} + X_{ik} + X_{jk} \qquad (\forall \text{ stable } \{i,j,k\}). \tag{18}$$

For further strengthenings of SDP2, see, e.g., [9, 10, 16, 19]. For computational results, see, e.g., [1, 4, 6, 14]. Generally speaking, adding cutting planes usually yields some improvement in the upper bound, but at the cost of dramatically increased computing times.

Going in the opposite direction, Lieder *et al.* [17] recently proposed to *weaken* the theta function, as follows. Take SDP2 and replace the equations (8) with the single equation

$$\sum_{\{i,j\} \in E} X_{ij} = 0. \tag{19}$$

The resulting bound can be computed much more quickly than $\vartheta(G)$, but, unfortunately, it is typically significantly weaker.

3 The Theta Function and its Weighted Version

In this section, we consider the theta function and the two associated SDPs. In Subsection 3.1, we show how to map a feasible Z^* onto a feasible X^*. In Subsection 3.2, we extend the mappings in both directions to the weighted case.

3.1 The theta function

Yildirim & Fan [26] showed how to map a feasible matrix Z^* to a feasible vector x^* without losing any profit, but they did not show how to construct the associated matrix X^*. The following theorem shows how to do this.

Theorem 1 *Let Z^* be a feasible solution to SDP1, and let $\gamma = J \bullet Z^* > 0$ be the associated profit. Let x^* be defined according to Yildirim and Fan (see the end of Subsection 2.2). If $Z^*_{ij} = 0$, then set X^*_{ij} to 0. Otherwise, set X^*_{ij} to*

$$Z^*_{ij} \sqrt{\frac{x^*_i x^*_j}{Z^*_{ii} Z^*_{jj}}} = Z^*_{ij} \frac{\left(\sum_{k=1}^n Z^*_{ik}\right)\left(\sum_{k=1}^n Z^*_{jk}\right)}{\gamma Z^*_{ii} Z^*_{jj}}.$$

Then the resulting pair (x^*, X^*) forms a feasible solution to SDP2, and its profit, $e^T x^*$, is at least as large as γ.

Proof. Since the profit of (x^*, X^*) is just $e^T x^*$, it does not depend on X^*. Therefore, the fact that the profit is at least as large as γ follows from the result in [26]. To complete the proof, it suffices to show that (x^*, X^*) satisfies (7) – (9).

To see that (7) is satisfied, simply note that $X^*_{ii} = Z^*_{ii} x^*_i / Z^*_{ii} = x^*_i$ by construction. To see that (8) is satisfied, simply recall that $Z^*_{ij} = 0$ for all $\{i, j\} \in E$, and note that $X^*_{ij} = 0$ whenever $Z^*_{ij} = 0$.

To show that (9) is satisfied, it suffices to construct a Gram representation of the augmented matrix

$$(X^+)^* = \begin{pmatrix} 1 & (x^*)^T \\ x^* & X^* \end{pmatrix}.$$

Let $Z^* = Y^T Y$ and let $y_1, \ldots, y_n$ be the columns of Y. Note that $||y_i|| = \sqrt{Z^*_{ii}}$ and that y_i is the zero vector whenever $Z^*_{ii} = 0$. We define vectors $v_0, \ldots, v_n$ as follows. We set v_0 to $Ye/\sqrt{\gamma}$. For all $i \in V$ such that $Z^*_{ii} > 0$, we set v_i to $y_i \sqrt{x^*_i}/||y_i||$. For all $i \in V$ such that $Z^*_{ii} = 0$, we set v_i to the zero vector.

Now we show that $v_0, \ldots, v_n$ form a Gram representation of $(X^+)^*$. We do this in three steps. First, we note that $v_0^T v_0 = 1$, since

$$||v_0||^2 = \frac{||Ye||^2}{\gamma} = \frac{(e^T Y^T)(Ye)}{\gamma} = \frac{e^T Z^* e}{\gamma} = \frac{J \bullet Z^*}{\gamma} = \frac{\gamma}{\gamma} = 1.$$

Second, we show that $v_0^T v_i = x^*_i$ for all $i \in V$. If $Z^*_{ii} = 0$, this is trivial, since v_i is the zero vector and $x^*_i = 0$ in this case. On the other hand, if $Z^*_{ii} > 0$, we have

$$v_0^T v_i = \left(\frac{Ye}{\sqrt{\gamma}}\right)^T \frac{y_i \sqrt{x^*_i}}{||y_i||} = (e^T Y^T y_i) \frac{\sqrt{x^*_i}}{\sqrt{\gamma}\,||y_i||} = \sum_{j=1}^{n} Z^*_{ij} \frac{\sqrt{x^*_i}}{\sqrt{\gamma}||y_i||}.$$

Then, since

$$x^*_i = \left(\sum_{j=1}^{n} Z^*_{ij}\right)^2 / \gamma Z^*_{ii}$$

when $Z^*_{ii} > 0$, we have that

$$v_0^T v_i = \sum_{j=1}^{n} Z^*_{ij} \frac{\sum_{j=1}^{n} Z^*_{ij}}{\gamma ||y_i|| \sqrt{Z^*_{ii}}} = x^*_i$$

as desired.

Third, we show that $v_i^T v_j = X^*_{ij}$ for all i and j. If $Z^*_{ij} = 0$, this is trivial, since v_i and v_j are orthogonal and $X^*_{ij} = 0$ in this case. On the other hand, if $Z^*_{ij} > 0$, we have

$$v_i^T v_j = \frac{(y_i^T y_j)\sqrt{x^*_i x^*_j}}{||y_i||\,||y_j||} = \frac{Z^*_{ij}\sqrt{x^*_i x^*_j}}{\sqrt{Z^*_{ii} Z^*_{jj}}} = X^*_{ij}$$

as desired. $\square$

We illustrate Theorem 1 on a small example:

Example 1: Let G be the graph on 3 nodes with $E = \big\{\{1,2\},\{2,3\}\big\}$. It holds trivially that $\alpha(G) = \vartheta(G) = 2$. One can check (e.g., with an Eigenvalue calculator) that the following matrix is a feasible solution for `SDP1`:

$$Z^* = \begin{pmatrix} 1/3 & 0 & 1/3 \\ 0 & 1/3 & 0 \\ 1/3 & 0 & 1/3 \end{pmatrix},$$

We have $\gamma = 5/3$. Applying the procedure of Yildirim and Fan, we obtain $x^* = (4/5, 1/5, 4/5)^T$, with profit $9/5 > \gamma$. Applying our procedure, we obtain

$$X^* = \begin{pmatrix} 4/5 & 0 & 4/5 \\ 0 & 1/5 & 0 \\ 4/5 & 0 & 4/5 \end{pmatrix}.$$

One can check that the corresponding augmented matrix $(X^+)^*$ is indeed psd. $\square$

The interesting thing about this example is that, if we take the matrix X^* and apply the mapping of Gruber and Rendl, we obtain the following feasible solution of `SDP1`:

$$\begin{pmatrix} 4/9 & 0 & 4/9 \\ 0 & 1/9 & 0 \\ 4/9 & 0 & 4/9 \end{pmatrix},$$

with an even better profit of 17/9. Further iterations of the mappings yield the profit values 33/17, 65/33 and 129/65. It is apparent that this sequence of profit values is rapidly converging to $\vartheta(G) = \alpha(G) = 2$.

An interesting question is to find conditions on the graph G and the starting matrix Z^* under which the sequence of profit values is guaranteed to converge to $\vartheta(G)$. Potentially, this could yield a new (primal) algorithm to compute $\vartheta(G)$. We do not examine this here, however. Instead, we give a second example, which illustrates how the mappings in both directions work when both Z^* and (x^*, X^*) are already optimal.

Example 2: Let G be the odd hole on 5 nodes. That is, let $n = 5$ and let

$$E = \big\{\{1,2\},\{1,5\},\{2,3\},\{3,4\},\{4,5\}\big\}.$$

It holds trivially that $\alpha(G) = 2$, and Lovász [18] showed that $\vartheta(G) = \sqrt{5} \approx 2.236$. One can check with an SDP solver that the optimal solution Z^* to `SDP1` has $Z^*_{ii} = 1/5$ for $i \in V$, $Z^*_{ij} = 0$ for $\{i,j\} \in E$, and $Z^*_{ij} = (\sqrt{5}-1)/10 \approx 0.1236$ for $\{i,j\} \in \bar{E}$. One can also check that the optimal solution (x^*, X^*) to `SDP2` has $x^*_i = X^*_{ii} = 1/\sqrt{5} \approx 0.4472$ for $i \in V$, $X^*_{ij} = 0$ for $\{i,j\} \in E$, and $X^*_{ij} = (1 - 5^{-1/2})/2 \approx 0.2764$ for $\{i,j\} \in \bar{E}$. Finally, one can check that the mappings between Z^* and (x^*, X^*) work in both directions. In particular, (i) $Z^*_{ij} = X^*_{ij}/\gamma$ for all i and j, as predicted by Gruber and Rendl, (ii) $x^*_i = \frac{\left(\sum_{j=1}^n Z^*_{ij}\right)^2}{\gamma Z^*_{ii}}$ for all i, as predicted by Yildirim and Fan, and (iii) $X^*_{ij} = Z^*_{ij}\sqrt{\frac{x^*_i x^*_j}{Z^*_{ii} Z^*_{jj}}}$ for all i and j, as predicted by Theorem 1. □

3.2 The weighted theta function

In this subsection, we extend the mappings to the weighted case. We start by giving an intuitive explanation for the weighted version of `SDP1`, due to Grötschel *et al.* [12], that was presented in Subsection 2.3.

Recall that $w \in \mathbb{R}^n_+$ is the vector of node weights. Let $\sqrt{w}$ denote the vector in $\mathbb{R}^n$ whose ith component is $\sqrt{w_i}$. Then, the objective function in the weighted version of `SDP1`, namely (10), can be written in the simpler form $\sqrt{w}^T Z \sqrt{w}$. Now, let $\bar{x} \in \{0,1\}^n$ be the incidence vector of a stable set, whose weight $\gamma = w^T \bar{x}$ is positive. Let $\bar{z} \in \mathbb{R}^n_+$ be the vector whose ith component is $\sqrt{w_i}\bar{x}_i/\sqrt{\gamma}$, and let $\bar{Z} = \bar{z}\bar{z}^T$. We then have

$$\bar{Z} = \frac{\mathrm{Diag}(\sqrt{w})\bar{x}\bar{x}^T \,\mathrm{Diag}(\sqrt{w})}{\gamma}.$$

One can check that $\bar{Z}$ satisfies (3)–(5), and is therefore feasible for `SDP1`. Moreover, we have

$$\sqrt{w}^T \bar{Z}\sqrt{w} = \frac{\sqrt{w}^T \,\mathrm{Diag}(\sqrt{w})\bar{x}\bar{x}^T \,\mathrm{Diag}(\sqrt{w})\sqrt{w}}{\gamma} = \frac{w^T \bar{x}\bar{x}^T w}{\gamma} = \frac{\gamma^2}{\gamma} = \gamma,$$

which explains the form of the objective function (10).

We are now in a position to extend the mappings to the weighted case. This is accomplished in the following two theorems.

Theorem 2 *Let (x^*, X^*) be a feasible solution to the weighted version of `SDP2`, whose profit $\gamma = w^T x^*$ is positive. We can obtain a feasible solution Z^* to the weighted version of `SDP1`, whose profit is at least as large, by setting Z^*_{ij} to $\sqrt{w_i w_j} X^*_{ij}/\gamma$ for all i and j.*

Proof. By construction, $Z^*_{ij} = 0$ whenever $X^*_{ij} = 0$, and therefore Z^* satisfies (3). Also, we have

$$\mathrm{Tr}(Z^*) = \sum_{i \in V} \frac{\sqrt{w_i w_i} X^*_{ii}}{\gamma} = \frac{1}{\gamma}\sum_{i \in V} w_i x^*_i = \gamma/\gamma = 1,$$

which shows that (4) is satisfied. Moreover, by construction, we have

$$Z^* = \frac{\mathrm{Diag}(\sqrt{w})\, X^* \,\mathrm{Diag}(\sqrt{w})}{\gamma}.$$

Since X^* is psd, it can be factorised as $Y^T Y$. This gives

$$Z^* = \frac{\mathrm{Diag}(\sqrt{w})\, Y^T Y \ \mathrm{Diag}(\sqrt{w})}{\gamma} = \left(\frac{Y \ \mathrm{Diag}(\sqrt{w})}{\sqrt{\gamma}}\right)^T \left(\frac{Y \ \mathrm{Diag}(\sqrt{w})}{\sqrt{\gamma}}\right),$$

which shows that (5) is satisfied.

Finally, we have to show that the profit of Z^* is at least γ. Due to (9) and the Schur complement, $X^* - x^*(x^*)^T$ must be psd. This implies that $w^T X^* w - (w^T x^*)^2 \geq 0$, and therefore $(w^T X^* w)/\gamma \geq w^T x^*$. But

$$\frac{w^T X^* w}{\gamma} = \frac{\sqrt{w}^T \, \mathrm{Diag}(\sqrt{w}) X^* \, \mathrm{Diag}(\sqrt{w}) \sqrt{w}}{\gamma} = \sqrt{w}^T Z^* \sqrt{w},$$

and this is the profit of Z^*. □

Theorem 3 *Let Z^* be a feasible solution to the weighted version of* `SDP1`*, whose profit $\gamma = \sqrt{w}^T Z^* \sqrt{w}$ is positive. We can obtain a feasible solution (x^*, X^*) to the weighted version of* `SDP2`*, whose profit is at least as large, as follows. If $Z^*_{ii} = 0$, set $x^*_i = 0$. Otherwise, set x^*_i to*

$$\frac{\left(\sum_{j=1}^n \sqrt{w_j} Z^*_{ij}\right)^2}{\gamma Z^*_{ii}}.$$

If $Z^*_{ij} = 0$, set $X^*_{ij} = 0$. Otherwise, set X^*_{ij} to

$$Z^*_{ij} \sqrt{\frac{x^*_i x^*_j}{Z^*_{ii} Z^*_{jj}}} \; = \; Z^*_{ij} \, \frac{\left(\sum_{k=1}^n \sqrt{w_k} Z^*_{ik}\right) \left(\sum_{k=1}^n \sqrt{w_k} Z^*_{jk}\right)}{\gamma Z^*_{ii} Z^*_{jj}}.$$

Proof. The proof that (x^*, X^*) is feasible is similar to the proof of Theorem 1. The only difference is that, in the Gram representation of the augmented matrix $(X^+)^*$, we need to set v_0 to $Y\sqrt{w}/\sqrt{\gamma}$.

The hard part is to show that the profit of (x^*, X^*) is at least as large as that of Z^*. We adapt the proof in [26]. Let $V' = \{i \in V : Z^*_{ii} > 0\}$. By construction, we have the following for all $i \in V'$:

$$x^*_i \; = \; v_0^T v_i \; = \; \frac{\sqrt{x^*_i}}{\sqrt{\gamma}||y_i||} \left(y_i^T Y \sqrt{w}\right) \; = \; \frac{\sqrt{x^*_i}}{\sqrt{\gamma}||y_i||} \sum_{j \in V} \sqrt{w_j} Z^*_{ij}.$$

Multiplying both sides by $||y_i||/\sqrt{x^*_i}$, we have

$$||y_i|| \sqrt{x^*_i} \; = \; \frac{1}{\sqrt{\gamma}} \sum_{j \in V} \sqrt{w_j} Z^*_{ij}.$$

Now, multiplying these equations by $\sqrt{w_i}$ and summing over all $i \in V'$, we obtain

$$\sum_{i \in V'} ||y_i|| \sqrt{w_i x^*_i} \; = \; \frac{\sqrt{w}^T Z^* \sqrt{w}}{\sqrt{\gamma}} \; = \; \gamma/\sqrt{\gamma} \; = \; \sqrt{\gamma}.$$

Equivalently,

$$\gamma = \left(\sum_{i \in V} ||y_i|| \sqrt{w_i x_i^*} \right)^2 .$$

By the Cauchy-Schwarz inequality, this cannot exceed

$$\left(\sum_{i \in V} ||y_i||^2 \right) \left(\sum_{i \in V} w_i x_i^* \right) \; = \; \mathrm{Tr}(Z^*)\, w^T x^*.$$

The result then follows from the fact that $\mathrm{Tr}(Z^*) = 1$. $\square$

Together, Theorems 2 and 3 provide an alternative proof that both SDPs yield the same upper bound, namely $\vartheta(G, w)$. They are illustrated on the following example.

Example 3: Let G be the odd hole on 5 nodes, as in Example 2, but now suppose that nodes $1, \ldots, 4$ have weight 2 and node 5 has weight 3. There are two optimal stable sets, each of weight 5. One can check with an SDP solver that the optimal solution to the weighted version of `SDP1` is as follows (to 3 d.p.):

$$Z^* = \begin{pmatrix} 0.074 & 0 & 0.057 & 0.057 & 0 \\ 0 & 0.194 & 0 & 0.057 & 0.198 \\ 0.057 & 0 & 0.194 & 0 & 0.198 \\ 0.057 & 0.057 & 0 & 0.074 & 0 \\ 0 & 0.198 & 0.198 & 0 & 0.464 \end{pmatrix}.$$

One can also check that the optimal solution to the weighted version of `SDP2` is

$$(X^+)^* = \begin{pmatrix} 1 & 0.188 & 0.494 & 0.494 & 0.188 & 0.788 \\ 0.188 & 0.188 & 0 & 0.145 & 0.145 & 0 \\ 0.494 & 0 & 0.494 & 0 & 0.145 & 0.412 \\ 0.494 & 0.145 & 0 & 0.494 & 0 & 0.412 \\ 0.188 & 0.145 & 0.145 & 0 & 0.188 & 0 \\ 0.788 & 0 & 0.412 & 0.412 & 0 & 0.788 \end{pmatrix}.$$

These solutions both yield $\vartheta(G, w) \approx 5.091 > 5$. Finally, one can check that the mappings described in Theorems 2 and 3 work in this case. $\square$

4 Variants of the Theta Function

Now we consider the variants of the theta function mentioned in Subsection 2.4. In Subsection 4.1, we consider stronger variants that are obtained by the addition of cutting planes. In Subsection 4.2, we consider the weakened variant of Lieder *et al.* [17].

4.1 Strengthening the function with cutting planes

It is interesting to observe that `SDP1` and `SDP2` behave rather differently when it comes to cutting planes. For example, whereas the inequalities (11), (12) and (13) are very similar to the inequalities (14), (15) and (17), respectively, the remaining inequalities for `SDP2`, i.e., (16) and (18), have not been adapted to `SDP1`. To explain this fact, we will need the following definition:

Definition 1 (Padberg [22]) *The Boolean quadric polytope of order n, denoted by BQP_n, is*

$$conv\left\{(x,y) \in \{0,1\}^{n+\binom{n}{2}} : \, y_{ij} = x_i x_j \, (1 \le i < j \le n)\right\}.$$

This family of polytopes has been studied in great depth, and many families of valid and facet-defining inequalities are known (see Deza & Laurent [5] for a survey). In particular, Padberg showed that the non-negativity inequalities $y_{ij} \ge 0$ define facets for $1 \le i < j \le n$.

We will also need the following notation:

- S denotes the set of all incidence vectors of stable sets. That is,

$$S = \{x \in \{0,1\}^n : (1) \text{ hold}\}.$$

- S^+ denotes the set of incidence vectors of *non-empty* stable sets. That is,

$$S^+ = \{x \in S : ||x|| \ne 0\}.$$

- $P(Z)$ denotes the convex hull of all feasible solutions to `SDP1` that represent stable sets. That is,

$$P(Z) = \text{conv}\left\{Z \in \mathbb{R}_+^{n\times n} : \, Z = \frac{xx^T}{e^T x} \text{ for some } \, x \in S^+\right\}.$$

- $P(x,X)$ denotes the convex hull of all feasible solutions to `SDP2` that represent stable sets. That is,

$$P(x,X) = \text{conv}\left\{(x,X) \in \mathbb{R}^{n+n^2} : \, x \in S, \, X = xx^T\right\}.$$

We have the following propositions:

Proposition 1 *If the inequality $\lambda^T x + \mu^T y \le 0$ is valid for BQP_n, then the inequality*

$$\sum_{i\in V} \lambda_i Z_{ii} + \sum_{\{i,j\}\in \bar{E}} \mu_{ij} Z_{ij} \le 0 \tag{20}$$

is valid for $P(Z)$. Moreover, the inequalities of this type, together with the equations (3), (4) and $Z_{ij} = Z_{ji}$ for $1 \le i < j \le n$, give a complete linear description of $P(Z)$.

Proof. For the sake of brevity, we only sketch the proof. Let Z^* be an extreme point of $P(Z)$. We construct a point $(x^*, y^*) \in \mathbb{R}^{n+\binom{n}{2}}$ by setting x_i^* to Z_{ii}^* for $i \in V$ and y_{ij}^* to Z_{ij}^* for $1 \le i < j \le n$. Let $P(x, y)$ denote the convex hull of all such points. Note that $P(Z)$ and $P(x, y)$ are affinely congruent.

Now, one shows that $P(x, y)$ can instead be obtained as follows. Take all extreme points of BQP_n that satisfy $y_{ij} = 0$ for $\{i, j\} \in E$, construct the cone of these points, and then intersect that cone with the hyperplane defined by the equation $e^T x = 1$. From this, it follows that $P(x, y)$ is completely described by the equation $e^T x = 1$, the equations $y_{ij} = 0$ for $\{i, j\} \in E$, and all valid homogeneous inequalities for BQP_n. The result follows easily. $\square$

Proposition 2 *If the inequality $\lambda^T x + \mu^T y \le \nu$ is valid for BQP_n, then the inequality*

$$\sum_{i \in V} \lambda_i x_i + \sum_{\{i,j\} \in \bar{E}} \mu_{ij} X_{ij} \le \nu$$

is valid for $P(x, X)$. Moreover, the inequalities of this type, together with the equations (7), (8) and $X_{ij} = X_{ji}$ for $1 \le i < j \le n$, give a complete linear description of $P(x, X)$.

Proof. Again, we only sketch the proof. Let (x^*, X^*) be an extreme point of $P(x, X)$. We construct a point (x^*, y^*) by setting y_{ij}^* to X_{ij}^* for $1 \le i < j \le n$. Let $P'(x, y)$ denote the convex hull of all such points. Note that $P(x, X)$ and $P'(x, y)$ are affinely congruent.

Now, one notes that $P'(x, y)$ is also equal to

$$\left\{ (x, y) \in \mathrm{BQP}_n : \ y_{ij} = 0 \ (\forall \{i, j\} \in E) \right\}.$$

From this, it follows that $P'(x, y)$ is completely described by the equations $y_{ij} = 0$ for $\{i, j\} \in E$, and all valid inequalities for BQP_n. The result follows easily. $\square$

As an illustration of these results, we recall that Padberg [22] showed that the following *triangle* inequalities define facets of BQP_n:

$$y_{ik} + y_{jk} \le x_k + y_{ij} \qquad (\forall \{i, j, k\} \subset \{1, \ldots, n\}) \qquad (21)$$

$$x_i + x_j + x_k \le 1 + y_{ij} + y_{ik} + y_{jk} \quad (1 \le i < j < k \le n). \qquad (22)$$

It is apparent that the inequalities (21) are the source of the inequalities (12) and (13) for `SDP1` and the inequalities (15) and (17) for `SDP2`. The inequalities (22) are the source of the inequalities (16) and (18) for `SDP2`, but, being inhomogeneous, do not yield any useful inequalities for `SDP1`.

Another issue that turns out to be worth considering is the effect that cutting planes have on the quality of the upper bound on the stability number. We start with the simplest cutting planes, the non-negativity inequalities (11) and (14). One can check that the mappings given by Gruber & Rendl [14], Yildirim & Fan [26] and ourselves all preserve non-negativity. Thus, the bound obtained by imposing non-negativity on X in `SDP2` is the same as the one obtained by imposing it on Z in `SDP1`, i.e., is nothing but Schrijver's bound $\vartheta'(G)$.

More generally, suppose we strengthen `SDP2` by adding an arbitrary collection of *homogeneous* valid inequalities, such as some or all of (14), (15) and (17). Let $(\tilde{x}, \tilde{X})$ be the optimal solution to the strengthened relaxation. The mapping of Gruber and Rendl, which merely divides $\tilde{X}$ by a constant, yields a feasible solution $\tilde{Z}$ to `SDP1` that satisfies the analogous homogeneous inequalities, and has at least as much profit. This implies that the bound obtained by adding the inequalities to `SDP2` is at least as strong as the one obtained by adding the analogous inequalities to `SDP1`. Surprisingly, it can be strictly stronger in some cases. This is shown in the following example.

Example 2 (cont.): Again, let G be the odd hole on 5 nodes. The optimal solution to `SDP1` violates the homogeneous inequality $Z_{13} + Z_{14} \leq Z_{11}$, which is of the form (12). One can check with an SDP solver that, if this inequality is added to `SDP1` as a cutting plane, the optimal solution becomes (to 3 d.p.):

$$\begin{pmatrix} 0.272 & 0 & 0.136 & 0.136 & 0 \\ 0 & 0.185 & 0 & 0.113 & 0.113 \\ 0.136 & 0 & 0.179 & 0 & 0.113 \\ 0.136 & 0.113 & 0 & 0.179 & 0 \\ 0 & 0.113 & 0.113 & 0 & 0.185 \end{pmatrix}.$$

The associated upper bound is 2.224.

The optimal solution to `SDP2` violates the analogous homogeneous inequality $X_{13} + X_{14} \leq x_1$, which is of the form (15). One can check that, if this is added to `SDP2`, the optimal matrix X^+ becomes

$$\begin{pmatrix} 1 & 0.586 & 0.293 & 1/2 & 1/2 & 0.293 \\ 0.586 & 0.586 & 0 & 0.293 & 0.293 & 0 \\ 0.293 & 0 & 0.293 & 0 & 0.293 & 0.121 \\ 1/2 & 0.293 & 0 & 1/2 & 0 & 0.293 \\ 1/2 & 0.293 & 0.293 & 0 & 1/2 & 0 \\ 0.293 & 0 & 0.121 & 0.293 & 0 & 0.293 \end{pmatrix}.$$

The associated upper bound is 2.172. □

Thus, if one wishes to add cutting planes, the choice between `SDP1` and `SDP2` is not at all clear-cut. It is true that we reduce the number of variables and constraints by using SDP1, but this can come at the cost of a weakened upper bound.

We remark that Proposition 1 is not valid in the weighted case. In that case, one has to modify the inequality (20) by dividing each λ_i by w_i and dividing each μ_{ij} by $\sqrt{w_i w_j}$. It remains true, however, that adding homogeneous cutting planes to SDP2 always leads to at least as much bound improvement as adding the analogous cutting planes to SDP1. (This can be shown by a modification of the proof of Theorem 2.)

4.2 The weakened theta function

Recall that the weakened theta function of Lieder *et al.* [17] is obtained by taking the constraints (8) in SDP2 and replacing them with the single weaker constraint (19). We will denote this bound by $\vartheta^{LRJ}(G)$. A natural question is whether one obtains the same bound if one weakens SDP1 by replacing the constraints (3) with the following single constraint:

$$\sum_{\{i,j\} \in E} Z_{ij} = 0. \tag{23}$$

The following proposition shows that the bound obtained in this way cannot be better than $\vartheta^{LRJ}(G)$.

Proposition 3 *If we weaken SDP1 by replacing (3) with (23), the resulting upper bound on $\alpha(G)$ is at least as large as $\vartheta^{LRJ}(G)$.*

Proof. We follow the strategy used in [14]. Suppose the pair (x^*, X^*) satisfies (7), (9) and (19), and let $\gamma = e^T x^* > 0$ be the associated objective value. Consider the matrix $Z^* = X^*/\gamma$. It satisfies (4), since

$$\mathrm{Tr}(Z^*) \,=\, \mathrm{Tr}(X^*)/\gamma \,=\, e^T x^*/\gamma \,=\, \gamma/\gamma \,=\, 1.$$

It satisfies (5), since X^* is psd; and it satisfies (23), since

$$\sum_{\{i,j\} \in E} Z^*_{ij} \;=\; \left(\sum_{\{i,j\} \in E} X^*_{ij} \right) /\gamma \;=\; 0/\gamma \;=\; 0.$$

Now, since $X^* - (x^*)(x^*)^T$ is psd, we have $e^T X^* e - e^T (x^*)(x^*)^T e \geq 0$. This implies that $J \bullet X^* \geq \gamma^2$, which in turn implies that $J \bullet Z^* \geq \gamma$. $\square$

The following example shows that, in fact, the bound obtained by weakening SDP1 can be larger than $\vartheta^{LRJ}(G)$. For this reason we will denote it by $\vartheta^{LRJ}_+(G)$.

Example 4: Let $n = 4$ and $E = \big\{\{1,2\}, \{1,3\}, \{2,3\}, \{3,4\}\big\}$. One can check with an SDP solver that $\alpha(G) = \vartheta(G) = \vartheta^{LRJ}(G) = 2$, while $\vartheta^{LRJ}_+(G) \approx 2.3855$. $\square$

5 Computational Results

We have seen that relaxations based on SDP1 can be weaker than the analogous relaxations based on SDP2. We conducted some computational experiments to see whether this is likely to occur in practice. To this end, we began by creating some random graphs according to the classical Erdös–Rényi process [7], in which each edge is present independently with probability p. For $n \in \{20, 30, 100\}$ and $p \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$, we constructed five graphs, making 135 graphs in total.

For each graph, we then solved several different SDP relaxations, using CSDP version 6.1.1 (see Borchers [2]). The experiments were conducted on a 2.299 GHz AMD Opteron 6376 with 16Gb RAM, under a 64 bit Linux operating system (Ubuntu 12.4). The C code was compiled with gcc 4.4.3, with -O3 optimisation.

The results for the 90 smaller instances, with 20 and 30 nodes, are displayed in Tables 1 to 3. Table 1 shows the mean gap (averaged over 5 instances) between the upper bound and the optimum, expressed as a percentage of the optimum, for eight SDP relaxations and nine different densities. The columns headed ϑ and ϑ' are self-explanatory, but note that $\vartheta'(G)$ can be obtained in two ways: by adding constraints (11) to SDP1 or adding constraints (14) to SDP2. The column headed "+(12)" was obtained by adding both constraints (11) and (12) to SDP1, and the column headed "+(13)" was obtained by adding, in addition, constraints (13). Similarly, the column headed "+(15)" was obtained by adding both constraints (14) and (15) to SDP2, and the column headed "+(17)" was obtained by adding, in addition, constraints (17). We did not experiment with adding constraints (16) or (18), since they have no analogue for SDP1.

We see that $\vartheta'(G)$ is only a little stronger than $\vartheta(G)$, a fact already observed in [4, 6]. On the other hand, adding further cutting planes to either SDP1 or SDP2 typically leads to significant improvements in the upper bound. In fact, for at least 50 instances out of 90, the additional cutting planes were able to close all of the gap. This accords with results given in [4, 6, 14]. As for the bounds based on weakening SDP1 or SDP2, their performance is disappointing.

The crucial point, however, is that the bounds based on SDP1 (in columns 5, 6 and 9) are frequently worse than their SDP2 counterparts (columns 7, 8 and 10). This indicates that the phenomenon mentioned in Section 4 is of practical importance.

Tables 2 and 3 show the times (in seconds) taken to compute the various bounds for the small instances. It is apparent that the addition of the cutting planes (12), (13) or (15), (17) slows down the SDP solver significantly, whereas weakening the SDP relaxations leads to a big saving in time. Strangely, however, the times for the variants of SDP2 are not much longer than for the variants of SDP1, even though the former have n fewer

n	Density	ϑ	ϑ'	+(12)	+(13)	+(15)	+(17)	ϑ_+^{LRJ}	ϑ^{LRJ}
	0.1	0.00	0.00	0.00	0.00	0.00	0.00	17.83	6.50
	0.2	0.94	0.94	0.09	0.07	0.00	0.00	24.81	9.64
	0.3	2.85	2.75	0.69	0.68	0.32	0.32	28.89	14.05
	0.4	1.91	1.90	0.51	0.50	0.16	0.15	37.70	19.34
20	0.5	0.79	0.79	0.00	0.00	0.00	0.00	32.65	15.01
	0.6	1.36	1.36	0.14	0.13	0.00	0.00	38.42	18.67
	0.7	6.22	6.01	4.00	4.00	3.36	3.34	61.86	40.24
	0.8	0.00	0.00	0.00	0.00	0.00	0.00	46.88	28.59
	0.9	2.36	2.36	2.02	2.02	1.34	1.34	44.73	23.56
	0.1	0.30	0.30	0.00	0.00	0.00	0.00	20.46	8.82
	0.2	0.28	0.21	0.00	0.00	0.00	0.00	27.83	11.09
	0.3	2.36	2.03	1.01	0.79	0.72	0.68	41.90	18.68
	0.4	5.89	5.57	4.25	4.22	3.49	3.47	54.21	28.94
30	0.5	0.31	0.24	0.00	0.00	0.00	0.00	41.64	21.80
	0.6	1.36	1.19	0.24	0.24	0.12	0.12	49.09	30.75
	0.7	2.34	2.28	1.49	1.45	1.08	1.05	66.23	41.92
	0.8	1.31	1.31	0.82	0.82	0.69	0.69	50.09	32.74
	0.9	0.00	0.00	0.00	0.00	0.00	0.00	40.88	28.63

Table 1: Mean percentage integrality gaps for various upper bounds.

constraints than the latter. It is also remarkable that "`SDP2`+(15)" is both stronger and faster to compute than "`SDP1`+(13)".

For the instances with 100 nodes, the SDP solver was unable to compute some of the stronger bounds in a reasonable amount of time. The results with the remaining bounds are shown in Tables 4 and 5. We see that, again, $\vartheta'(G)$ is only a little stronger than $\vartheta(G)$, $\vartheta^{LRJ}(G)$ is much weaker, and $\vartheta_+^{LRJ}(G)$ is much weaker still. As for running times, the effect of adding or aggregating constraints is even more marked. Moreover, we now see that the bounds based on `SDP2` now take about 50% longer to compute than those based on `SDP1`.

6 Concluding Remarks

It has been known since the 1980s [13] that the Lovász theta function can be computed by solving either of two equivalent SDPs, which we have labelled `SDP1` and `SDP2`. We have shown, however, that the equivalence between `SDP1` and `SDP2` breaks down when they are either strengthened (via cutting planes) or weakened (via constraint aggregation).

Our results have implications for the development of exact algorithms for the maximum stable set problem. In a branch-and-bound algorithm, a trade-off must be made between the quality of the upper bound and the time

n	Density	ϑ	ϑ'	+(12)	+(13)	ϑ_+^{LRJ}
	0.1	0.00	0.03	0.90	8.43	0.00
	0.2	0.01	0.03	2.66	9.97	0.00
	0.3	0.01	0.03	6.26	13.00	0.00
	0.4	0.01	0.03	12.79	18.03	0.00
20	0.5	0.01	0.02	18.37	22.20	0.00
	0.6	0.01	0.02	34.28	35.41	0.00
	0.7	0.02	0.02	61.24	59.58	0.00
	0.8	0.03	0.03	84.54	78.52	0.00
	0.9	0.04	0.02	95.45	90.67	0.00
	0.1	0.01	0.42	18.02	345.17	0.01
	0.2	0.02	0.24	58.52	316.61	0.01
	0.3	0.03	0.44	204.57	484.02	0.01
	0.4	0.04	0.23	405.88	665.35	0.01
30	0.5	0.08	0.42	843.04	1053.02	0.01
	0.6	0.10	0.23	1505.87	1673.03	0.01
	0.7	0.18	0.25	2618.84	2610.24	0.01
	0.8	0.16	0.19	3222.37	3166.29	0.01
	0.9	0.23	0.21	4124.27	3857.24	0.01

Table 2: Mean time taken to solve relaxations based on SDP1.

n	Density	ϑ	ϑ'	+(15)	+(17)	ϑ^{LRJ}
	0.1	0.01	0.04	0.88	8.89	0.01
	0.2	0.01	0.04	3.55	12.18	0.01
	0.3	0.01	0.05	7.73	15.23	0.00
	0.4	0.02	0.05	14.73	20.41	0.01
20	0.5	0.02	0.04	20.33	24.04	0.01
	0.6	0.02	0.04	39.82	41.08	0.00
	0.7	0.03	0.04	64.19	61.60	0.01
	0.8	0.04	0.04	82.08	77.82	0.01
	0.9	0.05	0.03	106.64	101.67	0.01
	0.1	0.02	0.63	18.38	367.98	0.01
	0.2	0.02	0.35	69.34	415.20	0.01
	0.3	0.04	0.63	214.68	572.30	0.01
	0.4	0.06	0.33	433.66	736.79	0.01
30	0.5	0.11	0.56	790.25	1041.33	0.01
	0.6	0.14	0.37	1537.81	1711.69	0.01
	0.7	0.20	0.35	2589.68	2673.10	0.02
	0.8	0.22	0.30	2948.88	2989.95	0.01
	0.9	0.31	0.26	3854.47	3850.89	0.01

Table 3: Mean time taken to solve relaxations based on SDP2.

Density	ϑ	ϑ'	ϑ_+^{LRJ}	ϑ^{LRJ}
0.1	7.71	7.00	45.20	23.20
0.2	11.60	10.67	65.19	40.41
0.3	13.69	12.64	79.08	52.91
0.4	15.49	14.55	92.90	68.45
0.5	16.80	16.03	107.30	81.78
0.6	10.88	10.27	103.94	83.85
0.7	8.28	7.94	109.61	87.54
0.8	2.55	2.38	107.78	89.60
0.9	0.00	0.00	88.35	73.55

Table 4: Mean percentage integrality gaps for instances with 100 nodes.

	SDP1			SDP2		
Density	ϑ	ϑ'	ϑ_+^{LRJ}	ϑ	ϑ'	ϑ^{LRJ}
0.1	0.45	249.19	0.12	0.85	483.40	0.20
0.2	2.11	242.63	0.10	2.39	430.05	0.13
0.3	4.84	233.45	0.13	6.53	381.38	0.21
0.4	10.14	233.33	0.09	13.81	367.74	0.12
0.5	18.80	237.10	0.12	25.30	478.52	0.20
0.6	32.32	224.19	0.09	42.78	332.43	0.13
0.7	51.99	214.54	0.12	65.34	320.85	0.28
0.8	83.10	232.80	0.09	106.57	298.75	0.13
0.9	122.62	213.00	0.09	147.30	312.65	0.13

Table 5: Mean time taken to solve SDP relaxations for instances with 100 nodes.

taken to compute it. If one wishes to use the theta function itself as the upper bound, then `SDP1` is to be preferred to `SDP2`. If, however, one wishes to use a strengthened or weakened variant of the theta number, then the choice between `SDP1` and `SDP2` is no longer obvious, and further experimentation will be needed.

References

References

[1] E. Balas, S. Ceria, G. Cornuéjols & G. Pataki (1996) Polyhedral methods for the maximum clique problem. In D.S. Johnson & M.A. Trick (eds.) *Cliques, Coloring and Satisfiability: The 2nd DIMACS Implementation Challenge*. Providence, RI: American Mathematical Society.

[2] B. Borchers (1999) CSDP, a C library for semidefinite programming. *Optim. Meth. & Software*, 11, 613–623.

[3] R. Borndörfer (1998) *Aspects of Set Packing, Partitioning and Covering*. Doctoral Thesis, Technical University of Berlin.

[4] S. Burer & D. Vandenbussche (2006) Solving lift-and-project relaxations of binary integer programs. *SIAM J. Opt.*, 16, 726–750.

[5] M.M. Deza & M. Laurent (1997) *Geometry of Cuts and Metrics*. Berlin: Springer-Verlag.

[6] I. Dukanovic & F. Rendl (2007) Semidefinite programming relaxations for graph coloring and maximal clique problems. *Math. Program.*, 109, 345–365.

[7] P. Erdös & A. Rényi (1959) On random graphs. *Publicationes Mathematicae*, 6, 290–297.

[8] U. Feige & J. Kilian (1998) Zero-knowledge and the chromatic number. *J. Comput. Syst. Sci.*, 57, 187–199.

[9] M. Giandomenico & A.N. Letchford (2006) Exploring the relationship between max-cut and stable set relaxations. *Math. Program.*, 106, 159-175.

[10] M. Giandomenico, A.N. Letchford, F. Rossi & S. Smriglio (2009) An application of the Lovász-Schrijver $M(K, K)$ operator to the stable set problem. *Math. Program.*, 120, 381–401.

[11] M. Giandomenico, F. Rossi & S. Smriglio (2013) Strong lift-and-project cutting planes for the stable set problem. *Math. Program.* , 141, 165–192.

[12] M. Grötschel, L. Lovász & A.J. Schrijver (1981) The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1, 169–197.

[13] M. Grötschel, L. Lovász & A.J. Schrijver (1988) *Geometric Algorithms and Combinatorial Optimization*. New York: Wiley.

[14] G. Gruber & F. Rendl (2003) Computational experience with stable set relaxations. *SIAM J. Opt.*, 13, 1014–1028.

[15] J. Håstad (1999) Clique is hard to approximate within $n^{1-\epsilon}$. *Acta Math.*, 182, 105–142.

[16] M. Laurent (2001) A comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre relaxations for 0-1 programming. *Math. Oper. Res.*, 28, 470–496.

[17] F. Lieder, F. Rad & F. Jarre (2015) Unifying semidefinite and set-copositive relaxations of binary problems and randomization techniques. *Comput. Optim. Appl.*, 61, 669–688.

[18] L. Lovász (1979) On the Shannon capacity of a graph. *IEEE Trans. Inform. Th.*, 25, 1–7.

[19] L. Lovász & A.J. Schrijver (1991) Cones of matrices and set-functions and 0-1 optimization. *SIAM J. Opt.*, 1, 166–190.

[20] G.L. Nemhauser & G. Sigismondi (1992) A strong cutting-plane / branch-and-bound algorithm for node packing. *J. Oper. Res. Soc.*, 43, 443–457.

[21] M.W. Padberg (1973) On the facial structure of set packing polyhedra. *Math. Program.*, 5, 199–215.

[22] M.W. Padberg (1989) The Boolean quadric polytope: some characteristics, facets and relatives. *Math. Program.*, 45, 139–172.

[23] S. Rebennack, M. Oswald, D.O. Theis, H. Seitz, G. Reinelt & P.M. Pardalos (2011) A branch and cut solver for the maximum stable set problem. *J. Combin. Optim.*, 21, 434–457.

[24] F. Rossi & S. Smriglio (2001) A branch-and-cut algorithm for the maximum cardinality stable set problem. *Oper. Res. Lett.*, 28, 63–74.

[25] A.J. Schrijver (1979) A comparison of the Delsarte and Lovász bounds. *IEEE Trans. Inf. Th.*, IT-25, 425–429.

[26] E.A. Yildirim & X. Fan–Orzechowski (2006) On extracting maximum stable sets in perfect graphs using Lovászs theta function. *Comput. Optim. Appl.*, 33, 229–247.