Semidefinite Relaxations, Multivariate Normal Distributions, and Order Statistics

Dimitris Bertsimas * and Yinyu Ye†

November 1997

Abstract

We propose two new dependent randomized rounding algorithms for approximating the global maximum of a quadratic program subject to linear as well as boolean constraints. We show that the first of these methods based on the multivariate normal distribution, provides (a) a 0.878 approximation algorithm for $s-t$ max-cut problems, (b) a 0.878 approximation algorithm for the max-cut problem that has the property that the expected number of nodes in one part of the cut is $n/2$, (c) a 0.878 approximation algorithm for $s-u-v$ max-cut problems with probability at least 0.91, and (d) a guarantee on the degree of infeasibility and suboptimality for quadratic optimization with linear constraints. The second of those methods based on order statistics provides near optimal solutions on maximum bisection problems on graphs with up to 3,000 nodes.

Key words. Quadratic optimization, graph bisection, $s-t$ max-cut, $s-u-v$ max-cut, approximation algorithms

*Boeing Professor of Operations Research, Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Mass. 02139. Research supported by NSF grant DMI-9610486.

†Department of Management Sciences, The University of Iowa, Iowa City, Iowa 52242. Research supported in part by NSF grants DMI-9522507 and DMS-9703490.

1 Introduction

Given the symmetric matrix $Q = \{q_{ij}\} \in \Re^{n\times n}$ and the constraint matrix $A = \{a_{ij}\} \in \Re^{m\times n}$, we consider the quadratic programming (QP) problem with linear and boolean constraints

$$\text{(QP)}\quad \begin{aligned} \text{Maximize} \quad & q(x) := x'Qx \\ \text{subject to} \quad & \left|\sum_{j=1}^{n} a_{ij}x_j\right| = b_i,\ i = 1,\ldots,m, \\ & x_j^2 = 1,\ j = 1,\ldots,n. \end{aligned}$$

Note that the constraint $x_j^2 = 1$ will force $x_j = 1$ or $x_j = -1$, making it a boolean variable.

Examples of problems that fit into our framework: The graph maximum bisection problem on an undirected graph $G = (V, E)$, which is the problem of partitioning the nodes in V into two sets S and $V \setminus S$ of equal cardinality so that $\sum_{(i,j)\in E,\, i\in S,\, j\in V\setminus S} c_{ij}$ is maximized, can be formulated as follows:

$$\text{(GB)}\quad \begin{aligned} \text{Maximize} \quad & \frac{1}{2}\sum_{i,j} c_{ij}(1 - x_i x_j) \\ \text{subject to} \quad & \sum_{j=1}^{n} x_j = 0, \\ & x_j^2 = 1,\ j = 1,\ldots,n. \end{aligned}$$

Another example is the $s - t$ max-cut problem

$$\text{(ST}^-\text{)}\quad \begin{aligned} \text{Maximize} \quad & \frac{1}{2}\sum_{i,j} c_{ij}(1 - x_i x_j) \\ \text{subject to} \quad & x_s + x_t = 0, \\ & x_j^2 = 1,\ j = 1,\ldots,n, \end{aligned}$$

meaning that nodes s and t should be separated in the optimal cut; and

$$\text{(ST}^+\text{)}\quad \begin{aligned} \text{Maximize} \quad & \frac{1}{2}\sum_{i,j} c_{ij}(1 - x_i x_j) \\ \text{subject to} \quad & |x_s + x_t| = 2, \\ & x_j^2 = 1,\ j = 1,\ldots,n, \end{aligned}$$

meaning that nodes s and t should be on the same side of the optimal cut. The third example is the $s-u-v$ max-cut problem

$$
\text{(SUV)} \qquad \begin{aligned} \text{Maximize} \quad & \frac{1}{2}\sum_{i,j} c_{ij}(1 - x_i x_j) \\ \text{subject to} \quad & |x_s + x_u + x_v| = 1, \\ & x_j^2 = 1, \; j = 1, \ldots, n, \end{aligned}
$$

meaning that nodes s, u and v cannot be sided together in the optimal cut. See Bertsimas, Teo and Vohra [4, 5] for more detailed descriptions of these problems; also see Gibbons, Hearn and Pardalos [7].

Recently, there were several results on approximating specific quadratic problems. Goemans and Williamson [8] provided a 0.878 approximation algorithm for the max-cut and $s-t$ max-cut problems. Frieze and Jerrum [6] generalized the Goemans and Williamson approach to the k-max-cut problem and provided a 0.65 approximation algorithm for the maximum graph bisection problem. Nesterov [12] provided a $\frac{2}{\pi}$ approximation algorithm for the boolean QP problem

$$
\begin{aligned} \text{Maximize} \quad & q(x) = x'Qx \\ \text{subject to} \quad & x_j^2 = 1, \; j = 1, \ldots, n, \end{aligned}
$$

where Q is positive semidefinite. Ye [17] further extended the $\frac{2}{\pi}$ result to solving the general QP problem

$$
\begin{aligned} \text{Maximize} \quad & q(x) = x'Qx \\ \text{subject to} \quad & \sum_{j=1}^{n} a_{ij} x_j^2 = b_i, \; i = 1, \ldots, m, \\ & -e \leq x \leq e, \end{aligned}
$$

where e is the vector of all ones. (Some negative results were given by Bellare and Rogaway [2].) The central idea for these approximation algorithms, due to Goemans and Williamson [8], is to use a semidefinite relaxation and then use randomized rounding to round the solution using the geometry of the semidefinite relaxation. To the best of our knowledge, there are no approximation guarantees that use the semidefinite relaxation for the $s-u-v$ max-cut problem.

Our contributions in this paper are as follows:

1. We introduce a new randomization technique based on the multivariate normal distribution to obtain feasible solutions from semidefinite relaxations. We show that the method provides (a) a 0.878 approximation algorithm for $s-t$ max-cut problems, (b) a 0.878 approximation algorithm for the max-cut problem, that shows the cut obtained from our technique partitions the graph into two equal sizes in expectation, (c) a 0.878 approximation algorithm for $s-u-v$ max-cut problems with probability at least 0.91, and (d) a guarantee on the degree of infeasibility and suboptimality for quadratic optimization with linear constraints. It is

interesting that the probability formulas we use is from Sheppard [15] from 1900. The method does not use the geometry of the semidefinite relaxation and it is based on interpreting the matrix X that arises from the semidefinite relaxation as a covariance matrix of a multivariate normal distribution.

2. We introduce another randomization scheme based on order statistics. A version of the heuristic that does not use a semidefinite relaxation provides a 0.5 approximation algorithm for the graph maximum bisection problem. We conjecture that another version that uses the semi-definite relaxation leads to a new algorithm for graph bisection, which is within close to 0.878 of the optimal. We report computational result on graphs with up to 3,000 nodes that show that the proposed method provides near optimal solutions.

2 Positive semidefinite relaxations

The semidefinite programming (SDP) relaxation for (QP) is as follows:

$$
\text{(SDP)} \qquad
\begin{aligned}
Z_{SD} = \quad & \text{Maximize} && \langle Q, X\rangle \\
& \text{subject to} && \langle A_i, X\rangle = b_i^2,\ i = 1,\ldots,m, \\
& && d(X) = e,\ X \succeq 0.
\end{aligned}
\tag{1}
$$

Here, $A_i = a_i' a_i \in \Re^{n\times n}$, $i = 1,\ldots,m$, (a_i is the ith row of the matrix A) and $X \in \Re^{n\times n}$ is a symmetric matrix. Furthermore, $\langle\cdot,\cdot\rangle$ is the matrix inner product $\langle Q, X\rangle = \text{trace}(QX)$, $d(X)$ is a column vector containing the diagonal components of X, $e \in \Re^n$ is the vector of all ones, and $X \succeq Z$ means that $X - Z$ is positive semidefinite. Obviously, (SDP) is feasible if (QP) is feasible, since $xx' \in \Re^{n\times n}$ is feasible for (SDP) if $x \in \Re^n$ is feasible for (QP). Thus, we assume that (SDP) is feasible throughout this paper (see Nesterov, Todd and Ye [14] for detecting the feasibility of (SDP)).

The dual of the (SDP) problem is

$$
\begin{aligned}
Z_{SD} = \quad & \text{Minimize} && \textstyle\sum_{i=1}^m b_i^2 \cdot \lambda_i + e'y \\
& \text{subject to} && \textstyle\sum_{i=1}^m \lambda_i A_i + D(y) \succeq Q,
\end{aligned}
\tag{2}
$$

where $D(y)$ is the diagonal matrix such that $d(D(y)) = y \in \Re^n$. Note that the dual is always feasible and it has an interior so that there is no duality gap between the primal and dual. Denote by $\bar{X}$ and $(\bar{\lambda}, \bar{y})$ an optimal solution pair for the primal (1) and dual (2).

The positive semidefinite relaxation was first proposed by Lovász and Shrijver [11]. This relaxation problem can be solved in polynomial time, e.g., see Nesterov and Nemirovskii [13] and Alizadeh [1].

Let Z_{QP} denote the optimal solution value of problem (QP), which we assume is feasible. Let Z_{SD} denote the optimal solution value of problem (SDP).

Proposition 1 *If Problem (QP) is feasible, then,*

$$Z_{QP} \leq Z_{SD}.$$

Proof. Let $\bar{x}$ be an optimal solution of Problem (QP). Let $X = \bar{x}\bar{x}' \in \Re^{n\times n}$. Then $X \succeq 0$, $d(X) = e$,

$$\langle A_i, X \rangle = \bar{x}' A_i \bar{x} = (a_i' \bar{x})^2 = 0, \; i = 1, \ldots, m,$$

and $\langle Q, X \rangle = \bar{x}' Q \bar{x} = Z_{QP}$. Thus, we have $Z_{QP} \leq Z_{SD}$. ■

3 A randomized rounding method based on the multivariate normal distribution

Let $\bar{X}$ be an optimal solution of Problem (SDP). We first review the Goemans and Williamson randomization, and then proceed with our new method.

The Goemans-Williamson randomization

Since $\bar{X}$ is positive semidefinite, there is a factorization matrix $\bar{V} = (\bar{v}_1, \ldots, \bar{v}_n) \in \Re^{n\times n}$, i.e., $\bar{v}_j$ is the jth column of $\bar{V}$, such that $\bar{X} = \bar{V}'\bar{V}$. The randomization method of Goemans and Williamson [8] generates a random vector u uniformly distributed on the n-dimensional unit ball and then assigns

$$\hat{x} = \text{sign}(\bar{V}'u), \tag{3}$$

and for any $x \in \Re^n$, sign(x) is the vector whose components are sign(x_j), $j = 1, \ldots, n$, that is,

$$\text{sign}(x_j) = \begin{cases} 1 & \text{if } x_j \geq 0 \\ -1 & \text{if } x_j < 0. \end{cases}$$

A new randomized rounding heuristic

We interpret the positive semidefinite solution $\bar{X}$ as the covariance matrix of a vector sample space.

A multivariate normal rounding heuristic (H1)

1. We generate a vector $\bar{x}$ from **a multivariate normal distribution** with 0 mean and covariance matrix $\bar{X}$, that is,

$$\bar{x} \in N(0, \bar{X}).$$

2. We assign

$$\hat{x} = \text{sign}(\bar{x}). \tag{4}$$

To generate $\bar{x}$, we simply generate a random vector u whose components are iid $N(0,1)$ and let $\bar{x} = \bar{V}'u$.

Proposition 2

$$\mathrm{E}[\hat{x}_j] = 0, \quad \mathrm{E}[\hat{x}_j^2] = 1, \quad j = 1, 2, \ldots, n,$$

$$\mathrm{E}[\hat{x}_i \hat{x}_j] = \frac{2}{\pi} \arcsin(\bar{x}_{ij}), \quad i, j = 1, 2, \ldots, n.$$

Proof. The marginal distribution of $\bar{x}_i$ is $N(0,1)$, and thus $\Pr(\hat{x}_i = 1) = \Pr(\hat{x}_i = -1) = 1/2$. Thus, $\mathrm{E}[\hat{x}_i] = 0$ and $\mathrm{E}[\hat{x}_i^2] = 1$. Furthermore,

$$\begin{aligned}
&\mathrm{E}[\hat{x}_i \hat{x}_j] \\
&= \Pr(\hat{x}_i = 1, \hat{x}_j = 1) + \Pr(\hat{x}_i = -1, \hat{x}_j = -1) - \Pr(\hat{x}_i = 1, \hat{x}_j = -1) - \Pr(\hat{x}_i = -1, \hat{x}_j = 1) \\
&= \Pr(x_i \geq 0, x_j \geq 0) + \Pr(x_i < 0, x_j < 0) - \Pr(x_i \geq 0, x_j < 0) - \Pr(x_i < 0, x_j \geq 0).
\end{aligned}$$

The tail probabilities of a multivariate normal distribution is a problem that has been studied in the last 100 years. Sheppard [15] shows (see Johnson and Kotz [10], p. 95) that

$$\begin{aligned}
\Pr(x_i \geq 0, x_j \geq 0) = \Pr(x_i < 0, x_j < 0) &= \frac{1}{4} + \frac{1}{2\pi} \arcsin(\bar{x}_{ij}) \\
\Pr(x_i \geq 0, x_j < 0) = \Pr(x_i < 0, x_j \geq 0) &= \frac{1}{4} - \frac{1}{2\pi} \arcsin(\bar{x}_{ij}).
\end{aligned}$$

This leads to

$$\mathrm{E}[\hat{x}_i \hat{x}_j] = \frac{2}{\pi} \arcsin(\bar{x}_{ij}). \tag{5}$$

■

We next apply this randomization method to several problems.

3.1 The $s - t$ max-cut problem

The relaxation of Problem (ST^-) that we solve in this case is

$$(\mathrm{RST}^-) \qquad \begin{aligned}
\text{Maximize} \quad & \frac{1}{2} \sum_{i,j} c_{ij} (1 - x_{ij}) \\
\text{subject to} \quad & x_{st} = -1, \\
& x_{jj} = 1, \; j = 1, \ldots, n, \\
& X \succeq 0.
\end{aligned}$$

Theorem 1 *Heuristic H1 provides a feasible solution for the (ST^-) max-cut problem with objective value Z_H:*

$$\mathrm{E}[Z_H] \geq 0.878 Z_{SD}.$$

Proof. First, we notice that $\mathrm{E}[\hat{x}_s + \hat{x}_t] = 0$, and

$$\mathrm{E}[(\hat{x}_s + \hat{x}_t)^2] = 2 + 2\mathrm{E}[\hat{x}_s \hat{x}_t] = 2 + \frac{4}{\pi}\arcsin(-1) = 2 - 2 = 0,$$

i.e., $\hat{x}_s + \hat{x}_t = 0$ with probability 1, i.e., the solution is feasible.

Moreover, the value of the heuristic solution is

$$\begin{aligned}
\mathrm{E}[Z_H] &= \sum_{ij} c_{ij}\Big(\Pr(\hat{x}_i = 1, \hat{x}_j = -1) + \Pr(\hat{x}_i = 1, \hat{x}_j = -1)\Big) \\
&= \sum_{ij} c_{ij}\left(\frac{1}{2} - \frac{1}{\pi}\arcsin(\bar{x}_{ij})\right) \\
&\geq 0.878\frac{1}{2}\sum_{ij} c_{ij}(1 - \bar{x}_{ij}) \\
&= 0.878 Z_{SD}.
\end{aligned}$$

■

The relaxation of Problem (ST^+) that we solve in this case is

$$(\mathrm{RST}^+) \qquad \begin{array}{ll}
\text{Maximize} & \frac{1}{2}\sum_{i,j} c_{ij}(1 - x_{ij}) \\
\text{subject to} & x_{st} = 1, \\
& x_{jj} = 1, \; j = 1, \ldots, n, \\
& X \succeq 0.
\end{array}$$

Theorem 2 *Heuristic H1 provides a feasible solution for the (ST^+) max-cut problem with objective value Z_H:*

$$\mathrm{E}[Z_H] \geq 0.878 Z_{SD}.$$

Proof. First, we notice that

$$\mathrm{E}[(\hat{x}_s + \hat{x}_t)^2] = 2 + 2\mathrm{E}[\hat{x}_s \hat{x}_t] = 2 + \frac{4}{\pi}\arcsin(1) = 4,$$

and

$$\mathrm{E}[(\hat{x}_s + \hat{x}_t)^4] = 8 + 8\mathrm{E}[\hat{x}_s \hat{x}_s] = 8 + \frac{16}{\pi}\arcsin(1) = 16.$$

Thus,

$$\mathrm{Var}[(\hat{x}_s + \hat{x}_t)^2] = 16 - 16 = 0,$$

i.e., $(\hat{x}_s + \hat{x}_t)^2 = 4$ or $|\hat{x}_s + \hat{x}_t| = 2$ with probability 1, i.e., the solution is feasible.

Again, the value of the heuristic solution satisfies $\mathrm{E}[Z_H] \geq 0.878 Z_{SD}$. ■

An identical analysis shows that Heuristic H1 provides the 0.878 bound for the max-cut problem and $\mathrm{E}[\sum_{j=1}^n \hat{x}_j] = 0$, i.e., the approximate cut is expected to be equally divided.

3.2 The $s-u-v$ max-cut problem

For simplicity, let $s=1$, $u=2$ and $v=3$. The relaxation of Problem (SUV) that we solve in this case is

$$
\text{(RSUV)} \qquad \begin{array}{ll}
\text{Maximize} & \frac{1}{2}\sum_{i,j} c_{ij}(1-x_{ij}) \\
\text{subject to} & x_{12}+x_{13}+x_{23}=-1, \\
& x_{jj}=1,\ j=1,\ldots,n, \\
& X \succeq 0.
\end{array}
$$

We now prove the following lemma.

Lemma 1 *Let $X \succeq 0$, $d(X)=e$ and $x_{12}+x_{13}+x_{23}=-1$. Then*

$$\arcsin(x_{12})+\arcsin(x_{12})+\arcsin(x_{12}) \le -3\arcsin(1/3).$$

Proof. From $X \succeq 0$ and $d(X)=e$, $x_{12}, x_{13}, x_{23} \in [-1,1]$. Thus, at least one of the three is negative. Now we have only two cases: the remaining two have the opposite sign; or the remaining two are both non-positive. In the former case, the maximum value of $\arcsin(x_{12})+\arcsin(x_{12})+\arcsin(x_{12})$ is $-2\arcsin(1/2)$, and in the latter its maximum value is $-3\arcsin(1/3)$. Since $-3\arcsin(1/3) > -2\arcsin(1/2)$, we derive the lemma. ■

Theorem 3 *Heuristic H1 applied to the $s-u-v$ max-cut problem generates a solution $\hat{x}$ with the following properties:*

$$
\begin{array}{ll}
\textbf{(a)} & \Pr((\hat{x}_1+\hat{x}_2+\hat{x}_3)^2=1) \ge \frac{3}{4}+\frac{3}{2\pi}\arcsin(1/3) > 0.912. \\
\textbf{(b)} & \mathrm{E}[Z_H] \ge 0.878 Z_{SD}.
\end{array}
$$

Proof.

$$
\begin{aligned}
\Pr((\hat{x}_1+\hat{x}_2+\hat{x}_3)^2=1) &= 1-\Pr(x_1\ge 0, x_2\ge 0, x_3\ge 0)-\Pr(x_1<0,x_2<0,x_3<0) \\
&= 1-2\Pr(x_1\ge 0, x_2\ge 0, x_3\ge 0) \\
&= 1-2\Big(\frac{1}{8}+\frac{1}{4\pi}(\arcsin(\bar{x}_{12})+\arcsin(\bar{x}_{13})+\arcsin(\bar{x}_{23}))\Big) \\
&\quad \text{(See Tong [16], p. 190.)} \\
&\ge \frac{3}{4}+\frac{3}{2\pi}\arcsin(1/3) > 0.912 \text{ (from the previous lemma).}
\end{aligned}
$$

As before, $\mathrm{E}[Z_H] \ge 0.878 Z_{SD}$. ■

Hence, we have

Corollary 1 *Heuristic H1 provides a feasible solution with probability 0.91 for the (SUV) max-cut problem and delivers an objective value Z_H:*

$$\mathrm{E}[Z_H] \ge 0.878 Z_{SD}.$$

3.3 Constrained quadratic optimization

In this section, we consider Problem (QP) with m linear homogeneous constraints, i.e., $b_i = 0$ for $i = 1, \ldots, m$. We assume that the problem is feasible and the matrix Q is positive semidefinite. Heuristic H1 does not necessarily generate a feasible solution. We show, however, that every constraint is satisfied in expectation. We also show an upper bound on the degree of violation of each constraint.

Theorem 4 *Heuristic H1 generates a solution $\hat{x}$ with the following properties:*

$$\begin{array}{ll} (\mathbf{a}) & \mathrm{E}[a_i'\hat{x}] = 0. \\ (\mathbf{b}) & \dfrac{\sqrt{\mathrm{E}[(a_i'\hat{x})^2]}}{\|a_i\|\|\hat{x}\|} \leq \sqrt{\dfrac{\pi-2}{\pi}}. \\ (\mathbf{c}) & \mathrm{E}[Z_H] \geq \dfrac{2}{\pi} Z_{SD}. \end{array}$$

Before proving the above theorem, we prove a lemma. For any function of one variable $f(t)$ and $X \in \Re^{n\times n}$, let $f[X] \in \Re^{n\times n}$ be the matrix with the components $f(x_{ij})$.

Lemma 2 *Let $X \succeq 0$ and $d(X) \leq e$. Then*

$$X + \frac{\pi-2}{2}(\sum_{j=1}^{n} x_{jj})I \succeq \arcsin[X] \succeq X,$$

where I is the identity matrix in $\Re^{n\times n}$.

Proof. The right-side inequality is proved by Nesterov [12]. We now prove the left. Note that $X \succeq 0$ and $|x_{ij}| \leq 1$ for all $i, j = 1, \ldots, n$, we have $[X]^t \succeq 0$ for all $t = 1, 2, \ldots$. Thus

$$\begin{aligned} \arcsin[X] &= X + \frac{1}{2\cdot 3}[X]^3 + \frac{1\cdot 3}{2\cdot 4\cdot 5}[X]^5 + \ldots \\ &\preceq X + \frac{1}{2\cdot 3}\Big(\sum_{j=1}^{n} x_{jj}^3\Big)I + \frac{1\cdot 3}{2\cdot 4\cdot 5}\Big(\sum_{j=1}^{n} x_{jj}^5\Big)I + \ldots \\ &\preceq X + \frac{1}{2\cdot 3}\Big(\sum_{j=1}^{n} x_{jj}\Big)I + \frac{1\cdot 3}{2\cdot 4\cdot 5}\Big(\sum_{j=1}^{n} x_{jj}\Big)I + \ldots \\ &= X + \Big(\sum_{j=1}^{n} x_{jj}\Big)I\Big(\frac{1}{2\cdot 3} + \frac{1\cdot 3}{2\cdot 4\cdot 5} + \ldots\Big) \\ &= X + \Big(\sum_{j=1}^{n} x_{jj}\Big)I\Big(\frac{\pi}{2} - 1\Big). \end{aligned}$$

■

Proof of the theorem
Since $\hat{x}$ has 0 mean,

$$\mathrm{E}[a_i'\hat{x}] = 0.$$

Recall that $\bar{X} \succeq 0$ and $d(\bar{X}) = e$. Since $A_i = a_i' a_i \succeq 0$, we have:

$$\begin{aligned}
\mathrm{E}[(a_i'\hat{x})^2] &= \mathrm{E}[\hat{x}' A_i \hat{x}] \\
&= \mathrm{E}[\langle A_i, \hat{x}\hat{x}' \rangle] \\
&= \langle A_i, \mathrm{E}[\hat{x}\hat{x}'] \rangle \\
&= \frac{2}{\pi}\langle A_i, \arcsin[\bar{X}] \rangle \\
&\leq \frac{2}{\pi}\langle A_i, \bar{X} + \frac{\pi - 2}{2}\left(\sum_{j=1}^{n} \bar{x}_{jj}\right) I \rangle \\
&= \frac{2}{\pi}\langle A_i, \frac{\pi - 2}{2}\left(\sum_{j=1}^{n} \bar{x}_{jj}\right) I \rangle \\
&= \frac{2}{\pi}\langle A_i, \frac{\pi - 2}{2} n I \rangle \\
&= \frac{\pi - 2}{\pi} n \langle A_i, I \rangle \\
&= \frac{\pi - 2}{\pi} n \|a_i\|^2,
\end{aligned}$$

proving the second part of the theorem.

The heuristic obtains a solution with expected value

$$\begin{aligned}
\mathrm{E}[Z_H] &= \sum_{i,j} q_{ij} \mathrm{E}[\hat{x}_i \hat{x}_j] \\
&= \frac{2}{\pi} \sum_{i,j} q_{ij} \arcsin(\bar{x}_{ij}) \\
&= \frac{2}{\pi} \langle Q, \arcsin[\bar{X}] \rangle \\
&\geq \frac{2}{\pi} \langle Q, \bar{X} \rangle \quad (\text{since } Q \succeq 0, \ \arcsin[\bar{X}] \succeq \bar{X}) \\
&= \frac{2}{\pi} Z_{SD},
\end{aligned}$$

proving the last part of the theorem. ■

The previous theorem shows that the solution produced by our method is feasible in expectation. It is not obvious that the Goemans-Williamson technique has this property. For the case without linear constraints, we obtain that heuristic H1 leads to a $\frac{2}{\pi}$ approximation guarantee. Nesterov [12] has showed that the randomization considered in Goemans and Williamson [8] leads to a $\frac{2}{\pi}$ approximation algorithm.

3.4 The maximum graph bisection problem

In this section, we show that the degree of violation of the constraint $\sum_{j=1}^n x_j = 0$ for the graph bisection problem (GB), after we apply heuristic H1 is indeed smaller than the general boolean, quadratic optimization problem. Similar results are obtained by Frieze and Jerrum [6]. Note that in solving (GB), we have $m = 1$, $A_1 = ee'$, $d(\bar{X}) = e$, and $e'\bar{X}e = \langle A_1, \bar{X}\rangle = 0$. Thus,

$$\mathrm{E}[(e'\hat{x})^2] = \frac{2}{\pi}\langle ee', \arcsin[\bar{X}]\rangle = \frac{2}{\pi} e' \arcsin[\bar{X}]e. \tag{6}$$

We now prove the following general lemma using the bound developed by Goemans and Williamson [8].

Lemma 3 *Let $X \succeq 0$, $d(X) = e$ and $a'Xa = 0$ where a is a nonnegative vector. Then*

$$\frac{2}{\pi} a' \arcsin[X] a \leq 0.122(e'a)^2.$$

Proof. Noting that $\sum_{ij} a_i a_j x_{ij} = 0$, we have

$$\begin{aligned}
\frac{2}{\pi} a' \arcsin[X] a &= \sum_{ij} a_i a_j \frac{2}{\pi} \arcsin(x_{ij}) \\
&= \sum_{ij} a_i a_j - \sum_{ij} a_i a_j (1 - \frac{2}{\pi}\arcsin(x_{ij})) \\
&\leq \sum_{ij} a_i a_j - 0.878 \sum_{ij} a_i a_j (1 - x_{ij}) \\
&= \sum_{ij} a_i a_j - 0.878 \sum_{ij} a_i a_j \\
&= 0.122(e'a)^2.
\end{aligned}$$

■

The following result is immediate:

Theorem 5 *Heuristic H1 applied to the graph bisection problem generates a solution $\hat{x}$ with the following properties:*

(a) $\mathrm{E}[e'\hat{x}] = 0.$

(b) $\dfrac{\sigma[e'\hat{x}]}{n} = \dfrac{\sqrt{\mathrm{E}[(e'\hat{x})^2]}}{n} \leq 0.3493.$

(c) $\mathrm{E}[Z_H] \geq 0.878 Z_{SD}.$

From the Chebyshev inequality

$$\Pr(|e'\hat{x}| \geq 0.36n) \leq \frac{\mathrm{E}[(e'\hat{x})^2]}{(0.36n)^2} = 0.943.$$

Thus, by repeating the randomization process, we can generate a cut in polynomial time that is within 0.878 from optimal such that at least 32% of the nodes are in one part of the cut and at least 32% of the nodes are in the other part.

4 A heuristic based on order statistics and its application to the graph bisection problem

In order to motivate Heuristic H3 below, we propose Heuristic H2 for the graph bisection problem that provides a bisection, which is guaranteed to be within 0.5 from optimal.

An order statistics randomized heuristic (H2)

1. For each i, $i = 1, \dots, n$, generate a uniform random variable U_i in $[0, 1]$.
2. Let $n = 2k$. Let $U_{(r)}$ be the order statistics of the numbers U_i. Let $U_{MED} = U_{(k)}$ be the median of the numbers U_i.
3. Nodes i, such that $U_i \geq U_{MED}$, belong to one part of the cut, and nodes with $U_i < U_{MED}$ belong in the other part of the cut.

Clearly, the above heuristic provides a feasible solution for the graph bisection problem. Let Z_{GB} be the optimal solution value of the graph bisection problem.

Theorem 6 *The order statistics randomized heuristic H2 provides a cut such that*

$$\mathrm{E}[Z_H] = \frac{k-1}{2k-1} \sum_{i,j} c_{ij} \geq \frac{k-1}{2k-1} Z_{GB}.$$

Proof. The expected value of the heuristic is:

$$E[Z_H] = \sum_{i,j} c_{ij} \Big(\Pr(U_i \leq U_{MED} < U_j) + \Pr(U_j \leq U_{MED} < U_i) \Big).$$

The event $U_j < U_{MED} < U_i$ is equivalent to the event that $k-1$ of the U_l $(l \neq i, j)$ are less than or equal to z, U_j is less than or equal to z, one of the U_l $(l \neq i, j)$ is within $(z, z + dz)$, $k-2$ of the U_l are greater than z, and U_i is greater than z, for some $z \in [0, 1]$. Thus,

$$\Pr(U_j < U_{MED} < U_i) = \frac{(2k-2)!}{(k-1)!\,1!\,(k-2)!} \int_0^1 z^k (1-z)^{k-1} dz.$$

The integral in the rhs is a beta integral $B(k+1, k)$, where

$$B(r, s) = \int_0^1 z^{r-1} (1-z)^{s-1} dz.$$

Since

$$B(k+1,k) = \frac{k!(k-1)!}{(2k)!},$$

we obtain that

$$\Pr(U_j < U_{MED} < U_i) = \frac{(2k-2)!}{(k-1)!\,1!\,(k-2)!}\frac{k!(k-1)!}{(2k)!} = \frac{k-1}{2(2k-1)}.$$

Symmetrically,

$$\Pr(U_i < U_{MED} < U_j) = \frac{k-1}{2(2k-1)}.$$

Thus, we obtain

$$E[Z_H] = \frac{k-1}{2k-1}\sum_{i,j} c_{ij} \geq \frac{k-1}{2k-1} Z_{GB}.$$

■

Notice that the above order statistics randomized heuristic is within $\frac{1}{2}(1-\frac{1}{n-1})$ from the optimal graph bisection.

The following algorithm is a natural generalization of Heuristics H1 and H2.

Multivariate normal order statistic heuristic (H3)

1. Solve the semidefinite relaxation of the graph bisection problem. Let $\bar{X}$ be an optimal solution.
2. Generate a random vector x distributed according to the multivariate normal with mean 0, and covariance matrix $\bar{X}$.
3. Compute the median x_{MED} of the numbers x_i.
4. Nodes i with $x_i \geq x_{MED}$, are in one part of the cut, and nodes i with $x_i < x_{MED}$ are in the other part of the cut.

Clearly, the above heuristic provides a feasible solution to the graph bisection problem. However, its analysis appears to be difficult as it involves the characterization of the median of correlated random variables. Note that since $\mathrm{Var}[\sum_i x_i] = E[e'xx'e] = e'\bar{X}e = 0$. Therefore, the sample mean $\sum_i x_i/n$ is equal to 0 deterministically. We then give a heuristic argument that leads us to believe that our randomized method produces a feasible cut which is within close to 0.878 from optimum. Heuristically speaking, one would expect that the median and the sample mean would be close to each other, i.e., we expect that $x_{MED} \approx 0$. Thus,

$$\begin{aligned} \mathrm{E}[Z_H] &= \sum_{i,j} c_{ij}\Big(\Pr(x_i \leq x_{MED} < x_j) + \Pr(x_j \leq x_{MED} < x_i)\Big) \\ &\approx \textstyle\sum_{i,j} c_{ij}\Big(\Pr(x_i \leq 0 < x_j) + \Pr(x_j \leq 0 < x_i)\Big) \\ &\geq 0.878 Z_{SD}, \end{aligned}$$

from our previous analysis. In the next section we show computational results that show that the above heuristic provides near optimal solutions empirically.

5 Computational results

We have applied the randomization methods H1 and H3 to approximating the maximum graph bisection of the well-known $G - set$ graph problems of Helmberg and Rendl [9]. These problems are generated by **rudy**, a machine independent graph generator written by G. Rinaldi, and become standard test problems for graph optimization. The positive semidefinite relaxation solution $\bar{X}$ are computed by the program of Benson, Ye and Zhang [3].

Graph	Dim.	Z_{SD}	$H1: \ |e'\hat{x}|$	H3: Z_H	Z_H/Z_{SD}
G1	800	48331.86	11.70	45760	0.9467
G2	800	48357.05	11.93	45596	0.9429
G3	800	48334.91	11.48	45788	0.9473
G11	800	2516.62	12.58	2136	0.8487
G12	800	2495.33	13.87	2112	0.8463
G13	800	2588.35	12.30	2224	0.8592
G14	800	12759.36	12.13	11908	0.9332
G15	800	12684.49	12.64	11872	0.9359
G16	800	12698.21	11.99	11864	0.9343
G17	800	12683.56	12.35	11880	0.9366
G18	800	4662.46	13.52	3696	0.7927
G19	800	4322.78	13.39	3380	0.7819
G20	800	4442.65	13.56	3532	0.7950
G22	2000	56543.19	18.03	51848	0.9169
G24	2000	56561.79	17.76	51796	0.9157
G32	2000	6270.51	20.76	5208	0.8305
G34	2000	6186.67	21.15	5184	0.8379
G43	1000	28128.19	12.71	26004	0.9244
G44	1000	28110.77	12.48	25952	0.9232
G45	1000	28098.26	12.44	25928	0.9227
G50	3000	23952.69	3.38	23520	0.9819
G51	1000	16022.41	13.64	14932	0.9319
G52	1000	16036.39	14.24	14964	0.9331
G53	1000	16036.85	13.96	14944	0.9318
G54	1000	16021.85	12.94	14956	0.9334

Table 1: Quality of new randomization algorithms for the maximum graph bisection problem.

We have solved 54 problems. Table 1 includes some representative problems. The second column of the Table includes the number of nodes in each graph. The third column is the positive semidefinite relaxation objective value Z_{SD} of the maximum graph bisection (GB). The fourth column is the average $|e'\hat{x}|$ of 2,000 vectors $\hat{x}$ generated by randomization method H1 for each graph, respectively. We see that this value is rather small relative to the number of nodes, indicating that the graph is close to being equally divided by $\hat{x}$. The fifth column is the largest objective value of the 2,000 vectors $\hat{x}$ generated by randomization method H3. Note that the vectors $\hat{x}$ generated by H3 are always feasible (i.e., $e'\hat{x} = 0$) for the graph bisection problem. The last column depicts the ratio Z_H/Z_{SD}.

Acknowledgement We would like to thank Steve Benson and Xiong Zhang for helping us to conduct the computational test.

References

[1] F. Alizadeh, *Combinatorial optimization with interior point methods and semi–definite matrices*, PhD thesis, University of Minnesota, Minneapolis, MN, 1991.

[2] M. Bellare and P. Rogaway, "The complexity of approximating a nonlinear program," *Mathematical Programming* 69 (1995) 429-442.

[3] S. Benson Y. Ye and X. Zhang, "Solving large-scale sparse semidefinite programs for combinatorial optimization," Working Paper, Computational and Applied Mathematics, The University of Iowa, Iowa City, IA 52242,1997.

[4] D. Bertsimas, C. Teo and R. Vohra, "On dependent randomized rounding algorithms," *Proc. 5th IPCO Conference* (1996) 330-344.

[5] D. Bertsimas, C. Teo and R. Vohra, "Nonlinear relaxations and improved randomized approximation algorithms for multicut problems," *Proc. 4th IPCO Conference* (1995) 29-39.

[6] A. Frieze and M. Jerrum, "Improved approximation algorithms for max k-cut and max bisection," *Proc. 4th IPCO Conference* (1995) 1-13.

[7] L. E. Gibbons, D. W. Hearn and P. M. Pardalos, "A continuous based heuristic for the maximum clique problem," *DIMACS Series in Discrete Mathematics and Theoretical Computer Science* 26 (1996) 103-124.

[8] M. X. Goemans and D. P. Williamson, " Improved approximation algorithms for Maximum Cut and Satisfiability problems using semidefinite programming," *Journal of ACM* 42 (1995) 1115–1145.

[9] C. Helmberg and F. Rendl, "A spectral bundle method for semidefinite programming," ZIB Preprint SC 97-37, Konrad-Zuse-Zentrum fuer Informationstechnik Berlin, Takustrasse 7, D-14195 Berlin, Germany, August 1997.

[10] N. Johnson and S. Kotz, *Distributions in Statistics: Continuous Multivariate Distributions*, John Wiley & Sons, 1972.

[11] L. Lovász and A. Shrijver, "Cones of matrices and setfunctions, and $0-1$ optimization," *SIAM Journal on Optimization* 1 (1990) 166-190.

[12] Yu. E. Nesterov, "Quality of semidefinite relaxation for nonconvex quadratic optimization," CORE Discussion Paper, #9719, Belgium, March 1997.

[13] Yu. E. Nesterov and A. S. Nemirovskii, *Interior Point Polynomial Methods in Convex Programming : Theory and Algorithms*, SIAM Publications, SIAM, Philadelphia, 1993.

[14] Yu. E. Nesterov, M. J. Todd, and Y. Ye, "Infeasible-start primal-dual methods and infeasibility detectors for nonlinear programming problems," Technical Report No. 1156, School of Operations Research and and Industrial Engineering, Cornell University, Ithaca, NY 14853-3801, 1996.

[15] W. F. Sheppard, "On the calculation of the double integral expressing normal correlation," *Transactions of the Cambridge Philosophical Society* 19 (1900) 23-66.

[16] Y. L. Tong, *The Multivariate Normal Distribution*, Springer-Verlag, New York, 1990.

[17] Y. Ye, "Approximating quadratic programming with quadratic constraints," Working Paper, Department of Management Science, The University of Iowa, Iowa City, IA 52242, 1997.