# Relaxations and Randomized Methods for Nonconvex QCQPs

Alexandre d'Aspremont, Stephen Boyd
EE392o, Stanford University

Autumn, 2003

## 1 Introduction

While some special classes of nonconvex problems can be efficiently solved, most nonconvex problems are very difficult to solve (at least, globally). In this set of notes we show how convex optimization can be used to find bounds on the optimal value of a hard problem, and can also be used to find good (but not necessarily optimal) feasible points. We first focus on Lagrangian relaxations, *i.e.*, using weak duality and the convexity of duals to get bounds on the optimal value of nonconvex problems. In a second section, we show how randomization techniques provide near optimal feasible points with, in some cases, bounds on their suboptimality.

### 1.1 Nonconvex QCQPs

In this note, we will focus on a specific class of problems: nonconvex quadratically constrained quadratic programs, or nonconvex QCQP (see also §4.4 in [BV03]). We will see that the range of problems that can be formulated as nonconvex QCQP is vast, and we will focus on some specific examples throughout the notes. We write a nonconvex QCQP as:

$$\begin{array}{ll} \mbox{minimize} & x^T P_0 x + q_0^T x + r_0 \\ \mbox{subject to} & x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m, \end{array} \tag{1}$$

with variable $x \in \mathbf{R}^n$, and parameters $P_i \in \mathbf{S}^n$, $q_i \in \mathbf{R}^n$, and $r_i \in \mathbf{R}$. In the case where all the matrices $P_i$ are positive semidefinite, the problem is convex and can be solved efficiently. Here we will focus on the case where at least one of the $P_i$ is not positive semidefinite. Note that the formulation above implicitly includes problems with equality constraints, which are equivalent to two opposing inequalities.

The nonconvex QCQP is NP-hard: it is at least as hard as a large number of other problems that also seem to be hard. While no one has proved that these problems really are hard, it is widely suspected that they are, and as a practical matter, all known algorithms to solve them have a complexity that grows exponentially with problem dimensions. So it's reasonable to consider them hard to solve (globally).

## 1.2 Examples and applications

We list here some examples of nonconvex QCQPs.

### 1.2.1 Boolean least squares

The problem is:

$$\begin{array}{ll} \text{minimize} & \|Ax-b\|^2 \\ \text{subject to} & x_i \in \{-1,1\}, \quad i=1,\ldots,n, \end{array} \tag{2}$$

in the variable $x \in \mathbf{R}^n$. This is a basic problem in digital communications (maximum likelihood estimation for digital signals). A brute force solution is to check all $2^n$ possible values of $x$. The problem can be expressed as a nonconvex QCQP:

$$\begin{array}{ll} \text{minimize} & x^TA^TAx - 2b^TAx + b^Tb \\ \text{subject to} & x_i^2 - 1 = 0, \quad i=1,\ldots,n. \end{array} \tag{3}$$

### 1.2.2 Minimum cardinality problems

The problem is to find a minimum cardinality solution to a set of linear inequalities:

$$\begin{array}{ll} \text{minimize} & \mathbf{Card}(x) \\ \text{subject to} & Ax \preceq b, \end{array} \tag{4}$$

in the variable $x \in \mathbf{R}^n$, with $\mathbf{Card}(x)$ the cardinality of the set $\{i \mid x_i \neq 0\}$. We assume that the feasible set $Ax \preceq b$ is included in the $\ell_\infty$ ball centered at zero with radius $R > 0$. We reformulate this problem as:

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^Tv \\ \text{subject to} & Ax \preceq b \\ & -Rv \preceq x \preceq Rv \\ & v \in \{0,1\}^n, \end{array} \tag{5}$$

in the variables $x, v \in \mathbf{R}^n$, and we then turn this into a nonconvex QCQP by replacing the constraints $v_i \in \{0,1\}$ by $v_i^2 - v_i = 0$. The problem then becomes:

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^Tv \\ \text{subject to} & Ax \preceq b \\ & -Rv \preceq x \preceq Rv \\ & v_i^2 - v_i = 0, \quad i=1,\ldots,n. \end{array} \tag{6}$$

This problem has many applications in engineering and finance, including for example low-order controller design and portfolio optimization with fixed transaction costs.

### 1.2.3 Partitioning problems

We consider here the two-way partitioning problem described in §5.1.4 and exercise 5.39 of [BV03]:

$$\begin{array}{ll} \text{minimize} & x^TWx \\ \text{subject to} & x_i^2 = 1, \quad i=1,\ldots,n, \end{array} \tag{7}$$

with variable $x \in \mathbf{R}^n$, where $W \in \mathbf{S}^n$. This problem is directly a nonconvex QCQP of the form (1). A feasible $x$ corresponds to the partition

$$\{1, \ldots, n\} \;=\; \{i \mid x_i = -1\} \;\cup\; \{i \mid x_i = 1\},$$

and the matrix coefficient $W_{ij}$ can be interpreted as the cost of having the elements $i$ and $j$ in the same partition, with $-W_{ij}$ the cost of having $i$ and $j$ in different partitions. The objective in (7) is the total cost, over all pairs of elements, and problem (7) seeks to find the partition with least total cost.

### 1.2.4 MAXCUT

On a $n$ node graph $G$, we define nonnegative weights $a_{ij}$ associated with each arc $(i,j)$, where $a_{ij} = 0$ if no arc connects node $i$ and $j$. The MAXCUT problem seeks to find a cut of the graph with the largest possible weight, *i.e.* a partition of the set of nodes in two parts $G_1$, $G_2$ such that the total weight of all arcs linking these parts is maximized. MAXCUT is a classic problem in network optimization and a particular case of the partitioning problem above. The weight of a particular cut $x_i$ is computed as

$$\frac{1}{2} \sum_{\{i \;|\; x_i x_j = -1\}} a_{ij}$$

which is also equal to

$$\frac{1}{4} \sum_{i,j=1}^{n} a_{ij}(1 - x_i x_j),$$

hence we let $W \in \mathbf{S}^n$ be a matrix defined by $W_{ij} = -a_{ij}$ if $i \neq j$ and $W_{ii} = \sum_{j=1,\ldots,n} a_{ij}$. Note that the matrix $W$ is here positive semidefinite. The problem is then formulated as:

$$\begin{array}{ll} \text{maximize} & x^T W x \\ \text{subject to} & x_i^2 = 1, \quad i = 1, \ldots, n, \end{array} \tag{8}$$

with variable $x \in \mathbf{R}^n$. Thus, MAXCUT is a special case of the partitioning problem.

### 1.2.5 Polynomial problems

A polynomial problem seeks to minimize a polynomial over a set defined by polynomial inequalities:

$$\begin{array}{ll} \text{minimize} & p_0(x) \\ \text{subject to} & p_i(x) \leq 0, \quad i = 1, \ldots, m. \end{array}$$

While seemingly much more general than simple nonconvex QCQPs, all polynomial problems can be turned into nonconvex QCQPs. Let us briefly detail how. First, we notice that we can reduce the maximum degree of an equation by adding variables. For example, we can turn the constraint

$$y^{2n} + (\ldots) \leq 0$$

into

$$u^n + (\ldots) \leq 0, \qquad u = y^2.$$

We have reduced the maximum degree of the original inequality by introducing a new variable and a quadratic equality constraint. We can also get rid of product terms; this time

$$xyz + (\ldots) \leq 0$$

becomes

$$ux + (\ldots) \leq 0, \qquad u = yz.$$

Here, we have replaced a product of three variables by a product of two variables (quadratic) plus an additional quadratic equality constraint. By applying these transformations iteratively, we can transform the original polynomials into quadratic objective and constraints, thus turning the original polynomial problem into a nonconvex QCQP, with additional variables.

**Example.** Let's work out a specific example. Suppose that we want to solve the following polynomial problem:

$$\begin{array}{ll} \text{minimize} & x^3 - 2xyz + y + 2 \\ \text{subject to} & x^2 + y^2 + z^2 - 1 = 0, \end{array}$$

in the variables $x, y, z \in \mathbf{R}$. We introduce two new variables $u, v \in \mathbf{R}$ with

$$u = x^2, \quad v = yz.$$

The problem then becomes:

$$\begin{array}{ll} \text{minimize} & xu - 2xv + y + 2 \\ \text{subject to} & x^2 + y^2 + z^2 - 1 = 0 \\ & u - x^2 = 0 \\ & v - yz = 0, \end{array}$$

which is a nonconvex QCQP of the form (1), in the variables $x, y, z, u, v \in \mathbf{R}$.

# 2 Convex relaxations

In this section, we begin by describing some direct relaxations of (1) using semidefinite programming (cf. [VB96] or [BV97]). We then detail how Lagrangian duality can be used as an "automatic" procedure to get lower bounds on the optimal value of the nonconvex QCQP described in (1). Note that both techniques provide lower bounds on the optimal value of the problem but give only a minimal hint on how to find an approximate solution (or even a feasible point ...), this will be the object of the next section.

## 2.1 Semidefinite relaxations

Starting from the original nonconvex QCQP:

$$\begin{array}{ll} \text{minimize} & x^T P_0 x + q_0^T x + r_0 \\ \text{subject to} & x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m, \end{array}$$

using $x^T P x = \mathbf{Tr}(P(xx^T))$, we can rewrite it:

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}(XP_0) + q_0^T x + r_0 \\ \text{subject to} & \mathbf{Tr}(XP_i) + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m, \\ & X = xx^T. \end{array} \tag{9}$$

We can directly relax this problem into a convex problem by replacing the last nonconvex equality constraint $X = xx^T$ with a (convex) positive semidefiniteness constraint $X - xx^T \succeq 0$. We then get a lower bound on the optimal value of (1) by solving the following convex problem:

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}(XP_0) + q_0^T x + r_0 \\ \text{subject to} & \mathbf{Tr}(XP_i) + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m, \\ & X \succeq xx^T. \end{array}$$

The last constraint $X \succeq xx^T$ is convex and can be formulated as a Schur complement (see §A.5.5 in [BV03]):

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}(XP_0) + q_0^T x + r_0 \\ \text{subject to} & \mathbf{Tr}(XP_i) + q_i^T x + r_i \leq 0, \quad \text{i=1,...,m}, \\ & \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \succeq 0, \end{array} \tag{10}$$

which is an SDP. This is called the *SDP relaxation* of the original nonconvex QCQP. Its optimal value is a lower bound on the optimal value of the nonconvex QCQP. Since it's an SDP, it's easy to solve, so we have a cheaply computable lower bound on the optimal value of the original nonconvex QCQP.

## 2.2 Lagrangian relaxations

We now study another method for getting a cheaply computable lower bound on the optimal value of the nonconvex QCQP. We take advantage of the fact that the dual of a problem is always convex, hence efficiently solvable. Again, starting from the original nonconvex QCQP:

$$\begin{array}{ll} \text{minimize} & x^T P_0 x + q_0^T x + r_0 \\ \text{subject to} & x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m, \end{array}$$

we form the Lagrangian,

$$L(x, \lambda) = x^T \left( P_0 + \sum_{i=1}^m \lambda_i P_i \right) x + \left( q_0 + \sum_{i=1}^m \lambda_i q_i \right)^T x + r_0 + \sum_{i=1}^m \lambda_i r_i.$$

To find the dual function, we minimize over $x$, using the general formula (see example 4.5 in [BV03]):

$$\inf_{x \in \mathbf{R}} x^T P x + q^T x + r = \begin{cases} r - \frac{1}{4} q^T P^{\dagger} q, & \text{if } P \succeq 0 \text{ and } q \in \mathcal{R}(P) \\ -\infty, & \text{otherwise.} \end{cases}$$

The dual function is then:

$$
\begin{aligned}
g(\lambda) &= \inf_{x\in\mathbf{R}^n} L(x,\lambda) \\
&= -\frac{1}{4}\left(q_0+\sum_{i=1}^m \lambda_i q_i\right)^T \left(P_0+\sum_{i=1}^m \lambda_i P_i\right)^\dagger \left(q_0+\sum_{i=1}^m \lambda_i q_i\right) + \sum_{i=1}^m \lambda_i r_i + r_0.
\end{aligned}
$$

We can form the dual of (1), using Schur complements (cf. §A.5.5):

$$
\begin{array}{ll}
\text{maximize} & \gamma + \sum_{i=1}^m \lambda_i r_i + r_0 \\
\text{subject to} & \left[\begin{array}{cc} \left(P_0+\sum_{i=1}^m \lambda_i P_i\right) & \left(q_0+\sum_{i=1}^m \lambda_i q_i\right)/2 \\ \left(q_0+\sum_{i=1}^m \lambda_i q_i\right)^T/2 & -\gamma \end{array}\right] \succeq 0 \\
& \lambda_i \geq 0, \quad i=1,\ldots,m,
\end{array}
\tag{11}
$$

in the variable $\lambda \in \mathbf{R}^m$. As the dual to (1), this is a convex program, it is in fact a semidefinite program. This SDP is called the *Lagrangian relaxation* of the nonconvex QCQP. It's easy to solve, and gives a lower bound on the optimal value of the nonconvex QCQP.

An interesting question is, what is the relation between the Lagrangian relaxation and the SDP relaxation? They are both SDPs, and they both provide lower bounds on the optimal value of the nonconvex QCQP. In particular, is one of the bounds better than the other? The answer turns out to be simple: (10) and (11) are duals of each other, and so (assuming a constraint qualification holds) the bounds are exactly the same.

## 2.3 Perfect duality

Weak duality implies that the optimal value of the Lagrangian relaxation is a lower bound on that of the original program. In some particular cases, even though the original program is not convex, this duality gap is zero and the convex relaxation produces the optimal value.

QCQP with only one constraint is a classic example (see Appendix B in [BV03], or [Fer99] for others), based on the fact that the numerical range of two quadratic forms is a convex set. This means that, under some technical conditions, the programs:

$$
\begin{array}{ll}
\text{minimize} & x^T P_0 x + q_0^T x + r_0 \\
\text{subject to} & x^T P_1 x + q_1^T x + r_1 \leq 0,
\end{array}
\tag{12}
$$

and

$$
\begin{array}{ll}
\text{maximize} & \gamma + \lambda r_1 + r_0 \\
\text{subject to} & \left[\begin{array}{cc} \left(P_0+\lambda P_1\right) & \left(q_0+\lambda q_1\right)/2 \\ \left(q_0+\lambda q_1\right)^T/2 & -\gamma \end{array}\right] \succeq 0 \\
& \lambda \geq 0,
\end{array}
\tag{13}
$$

in the variables $x \in \mathbf{R}^n$ and $\lambda \in \mathbf{R}$ respectively, produce the same optimal value, even if the first one is nonconvex. This result is also known as the $S$-procedure in control theory. The key implication here of course is that while the original program is possibly nonconvex and numerically hard, its dual is a semidefinite program and is easy to solve.

## 2.4 Examples

Let us now work out the Lagrangian relaxations of the examples detailed above.

### 2.4.1 MINCARD relaxation

Let's first consider the MINCARD problem detailed in (4):

$$\begin{array}{ll}\text{minimize} & \mathbf{Card}(x)\\ \text{subject to} & Ax \preceq b.\end{array}$$

Using the problem formulation in (6), the relaxation given by (10) is then:

$$\begin{array}{ll}\text{minimize} & \mathbf{1}^T v\\ \text{subject to} & Ax \preceq b\\ & -Rv \preceq x \preceq Rv\\ & X_{ii} - x_i = 0, \quad i = 1, \ldots, n\\ & \left[\begin{array}{cc} X & x\\ x^T & 1\end{array}\right] \succeq 0.\end{array}$$

It's easy to show that the optimal $X$ is simply $\mathbf{diag}(x)$ (see [BFH00] and [LO99, Th. 5.2]); this implies that this relaxation produces the same lower bound as the direct linear programming relaxation:

$$\begin{array}{ll}\text{minimize} & \mathbf{1}^T v\\ \text{subject to} & Ax \preceq b\\ & -Rv \preceq x \preceq Rv\\ & v \in [0,1]^n.\end{array} \tag{14}$$

It is also related to the classical $\ell_1$ heuristic described in [BFH00], which replaces the function $\mathbf{Card}(x)$ with its largest convex lower bound (over a ball in $\ell_\infty$) $\|x\|_1$:

$$\begin{array}{ll}\text{minimize} & \|x\|_1\\ \text{subject to} & Ax \preceq b.\end{array} \tag{15}$$

### 2.4.2 Boolean least squares

The original boolean least squares problem in (2) is written:

$$\begin{array}{ll}\text{minimize} & \|Ax - b\|^2\\ \text{subject to} & x_i^2 = 1, \quad i = 1, \ldots, n,\end{array}$$

We can relax its QCQP formulation (3) as an SDP:

$$\begin{array}{ll}\text{minimize} & \mathbf{Tr}(A^T A X) - 2b^T Ax + b^T b\\ \text{subject to} & \left[\begin{array}{cc} X & x\\ x^T & 1\end{array}\right] \succeq 0\\ & X_{ii} = 1, \quad i = 1, \ldots, n,\end{array} \tag{16}$$

in the variables $x \in \mathbf{R}^n$ and $X \in \mathbf{S}^n_+$. This program then produces a lower bound on the optimal value of the original problem.

### 2.4.3 Partitioning and MAXCUT

The partitioning problem defined above reads:

$$\begin{array}{ll} \text{minimize} & x^T W x \\ \text{subject to} & x_i^2 = 1, \quad i = 1, \ldots, n. \end{array} \tag{17}$$

Here, the problem is directly formulated as a nonconvex QCQP and the variable $x$ disappears from the relaxation, which becomes:

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}(WX) \\ \text{subject to} & X \succeq 0 \\ & X_{ii} = 1, \quad i = 1, \ldots, n. \end{array} \tag{18}$$

MAXCUT corresponds to a particular choice of matrix $W$.

# 3 Randomization

The Lagrangian relaxation techniques developed in §2 provided lower bounds on the optimal value of the program in (1), but did not however give any particular hint on how to compute good feasible points. The semidefinite relaxation in (10) produces a positive semidefinite or covariance matrix together with the lower bound on the objective. In this section, we exploit this additional output to compute good approximate solutions with, in some cases, hard bounds on their suboptimality.

## 3.1 Randomization

In the last section, the original nonconvex QCQP:

$$\begin{array}{ll} \text{minimize} & x^T P_0 x + q_0^T x + r_0 \\ \text{subject to} & x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m, \end{array}$$

was relaxed into:

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}(XP_0) + q_0^T x + r_0 \\ \text{subject to} & \mathbf{Tr}(XP_i) + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m, \\ & \left[ \begin{array}{cc} X & x \\ x^T & 1 \end{array} \right] \succeq 0. \end{array} \tag{19}$$

The last (Schur complement) constraint being equivalent to $X - xx^T \succeq 0$, if we suppose $x$ and $X$ are the solution to the relaxed program in (19), then $X - xx^T$ is a covariance matrix.

If we pick $x$ as a Gaussian variable with $x \sim \mathcal{N}(x, X - xx^T)$, $x$ will solve the nonconvex QCQP in (1) "on average" over this distribution, meaning:

$$\begin{array}{ll} \text{minimize} & \mathbf{E}(x^T P_0 x + q_0^T x + r_0) \\ \text{subject to} & \mathbf{E}(x^T P_i x + q_i^T x + r_i) \leq 0, \quad i = 1, \ldots, m, \end{array}$$

and a "good" feasible point can then be obtained by sampling $x$ a sufficient number of times, then simply keeping the best feasible point.

### 3.2 Feasible points

Of course the direct sampling technique above does not guarantee that a feasible point will be found. In particular, if the program includes an equality constraint, then this method will certainly fail. However, it is sometimes possible to directly project the random samples onto the feasible set. This is the case, for example, in the partitioning problem, where we can discretize the samples by taking the $\mathbf{sgn}(x)$ function. In this case, the randomization procedure then looks like this. First, sample points $x_i$ with a normal distribution $\mathcal{N}(0, X)$, where $X$ is the optimal solution of (18). Then get feasible points by taking $\hat{x}_i = \mathbf{sgn}(x_i)$.

### 3.3 Bounds on suboptimality

In certain particular cases, it is also possible to get a hard bound on the gap between the optimal value and the relaxation result. A classic example is that of the MAXCUT bound described in [GW95] or [BTN01, Th. 4.3.2]. The MAXCUT problem (8) reads:

$$\begin{array}{ll} \text{maximize} & x^T W x \\ \text{subject to} & x_i^2 = 1, \quad i = 1, \ldots, n, \end{array} \tag{20}$$

its Lagrangian relaxation was computed in (18):

$$\begin{array}{ll} \text{maximize} & \mathbf{Tr}(WX) \\ \text{subject to} & X \succeq 0 \\ & X_{ii} = 1, \quad i = 1, \ldots, n. \end{array} \tag{21}$$

We then sample feasible points $\hat{x}_i$ using the procedure described above. Crucially, when $\hat{x}$ is sampled using that procedure, the expected value of the objective $\mathbf{E}(\hat{x}^T W x)$ can be computed explicitly:

$$\mathbf{E}(\hat{x}^T W x) = \frac{2}{\pi} \sum_{i,j=1}^{n} W_{ij} \arcsin(X_{ij}) = \frac{2}{\pi} \mathbf{Tr}(W \arcsin(X)).$$

We are guaranteed to reach this expected value $2/\pi\, \mathbf{Tr}(W \arcsin(X))$ after sampling a few (feasible) points $\hat{x}$, hence we know that the optimal value $OPT$ of the MAXCUT problem is between $2/\pi\, \mathbf{Tr}(W \arcsin(X))$ and $\mathbf{Tr}(WX)$.

Furthermore, with $\arcsin(X) \succeq X$ (see [BTN01, p. 174]), we can simplify (and relax) the above expression to get:

$$\frac{2}{\pi} \mathbf{Tr}(WX) \leq OPT \leq \mathbf{Tr}(WX).$$

This means that the procedure detailed above guarantees that we can find a feasible point that is at most $2/\pi$ suboptimal (after taking a certain number of samples from a Gaussian distribution).

# 4 Linearization and convex restriction

The relaxation techniques detailed in §2 produce lower bounds on the optimal value but no feasible points. Here, we work on the complementary approach and try to find "good" feasible points corresponding to a local minimum. Let $x^{(0)}$ be an initial feasible point which might be found using the results in the last section or by a phase I procedure; see the discussion on phase I problems in §11.4 of [BV03]).

## 4.1 Linearization

We start by leaving all convex constraints unchanged, linearizing the nonconvex ones around the original feasible point $x^{(0)}$. Consider for example the constraint:

$$x^T P x + q^T x + r \leq 0,$$

we decompose the matrix $P$ into its positive and negative parts:

$$P = P_+ - P_-, \qquad P_+,\ P_- \succeq 0.$$

The original constraint can be rewritten as

$$x^T P_+ x + q_0^T x + r_0 \leq x^T P_- x,$$

and both sides of the inequality are now convex quadratic functions. We linearize the right hand side around the point $x_0$ to obtain

$$x^T P_+ x + q_0^T x + r_0 \leq x^{(0)T} P_- x^{(0)} + 2x^{(0)T} P_-(x - x^{(0)}).$$

The right hand side is now an affine lower bound on the original function $x^T P_- x$ (see §3.1.3 in [BV03]). This means that the resulting constraint is convex and more conservative than the original one, hence the feasible set of the new problem will be a convex subset of the original feasible set. Thus, by linearizing the concave parts of all constraints, we obtain a set of convex constraints that are tighter than the original nonconvex ones. In other words, we form a *convex restriction* of the problem.

## 4.2 Iteration

The new problem, formed by linearizing all the nonconvex constraints using the method described above, is a convex QCQP and can be solved efficiently to produce a new feasible point $x^{(1)}$ with a lower objective value. If we linearize again the problem around $x^{(1)}$ and repeat the procedure, we get a sequence of feasible points with decreasing objective values.

This simple idea has been discovered and rediscovered several times. It is sometimes called the *convex-concave procedure*. It doesn't work for Boolean problems, since the only convex subsets of the feasible set are singletons.

# 5 Numerical Examples

In this section, we work out some numerical examples.

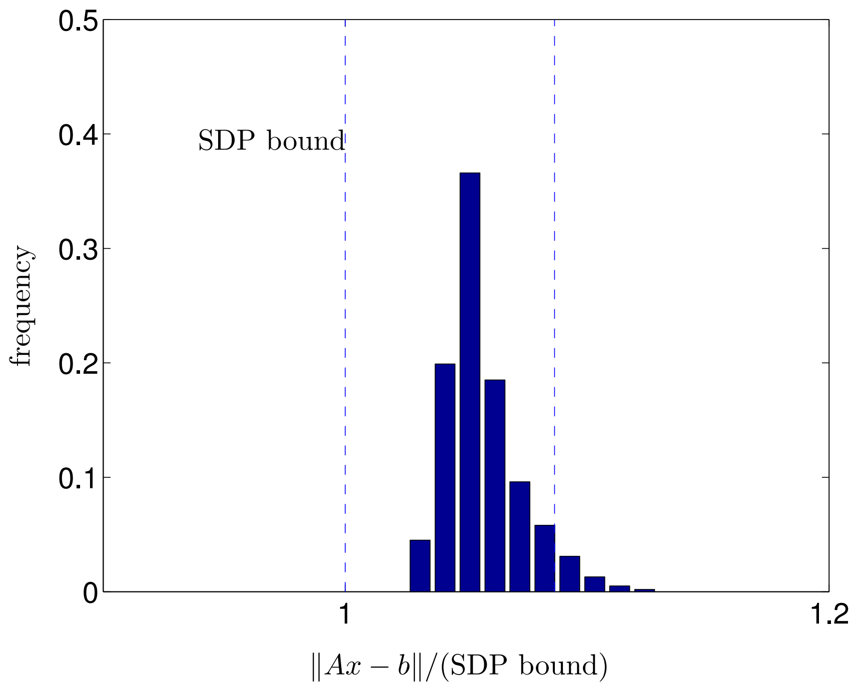


**Figure 1:** Distribution of objective values for points sampled using the randomization technique in §3.

## 5.1 Boolean least-squares

We use (randomly chosen) parameters $A \in \mathbf{R}^{150\times 100}$, $b \in \mathbf{R}^{150}$ and $x \in \mathbf{R}^{100}$, the feasible set has $2^{100} \approx 10^{30}$ points. In figure (5.1), we plot the distribution of the objective values reached by the feasible points found using the randomized procedure above. Our best solution comes within 2.6% of the SDP lower bound.

## 5.2 Partitioning

We consider here the two-way partitioning problem described in (7) and compare various methods.

### 5.2.1 A simple heuristic for partitioning

One simple heuristic for finding a good partition is to solve the SDPs above, to find $X^\star$ (and the bound $d^\star$). Let $v$ denote an eigenvector of $X^\star$ associated with its largest eigenvalue, and let $\hat{x} = \mathbf{sgn}(v)$. The vector $\hat{x}$ is our guess for a good partition.

### 5.2.2 A randomized method

We generate independent samples using the procedure described in §3.

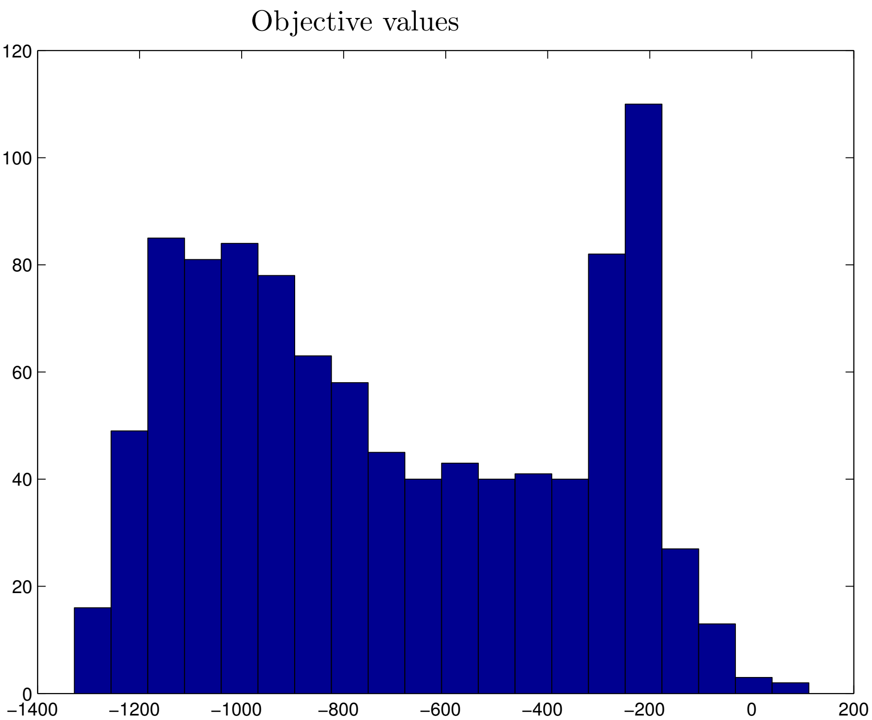


**Figure 2:** Histogram of the objective values attained by the random sample partitions.

### 5.2.3 Greedy method

We can improve these results a little bit using the following simple greedy heuristic. Suppose the matrix $Y = \hat{x}^T W \hat{x}$ has a column $j$ whose sum $\sum_{i=1}^n y_{ij}$ is positive. Switching $\hat{x}_j$ to $-\hat{x}_j$ will decrease the objective by $2\sum_{i=1}^n y_{ij}$. If we pick the column $y_{j_0}$ with largest sum, switch $\hat{x}_{j_0}$ and repeat until all column sums $\sum_{i=1}^n y_{ij}$ are negative, we further decrease the objective.

### 5.2.4 Numerical Example

For our example, the optimal SDP lower bound $d^\star$ is equal to $-1641$ and the $\mathbf{sgn}(x)$ heuristic gives a point (partition) with total cost $-1348$. Extracting a solution from the SDP solution using the simple heuristic above gives a solution with cost $-1280$, while applying the greedy method pushes that cost down to $-1372$. Exactly what the optimal value is, we can't say; all we can say at this point is that it is between $-1641$ and $-1372$.

We then try the randomized method, applying the greedy method to each sample, and plot in figure (5.2.4) a histogram of the objective obtained over 1000 samples. Many of these samples have an objective value larger than the original one above, but some have a lower cost. For our implementation, we found the minimum value $-1392$. The evolution of the minimum value found as a function of the sample size is shown in figure (5.2.4). Note that our best partition was found in around 100 samples. We're not sure what the optimal cost is, but now we know it's between $-1641$ and $-1392$.

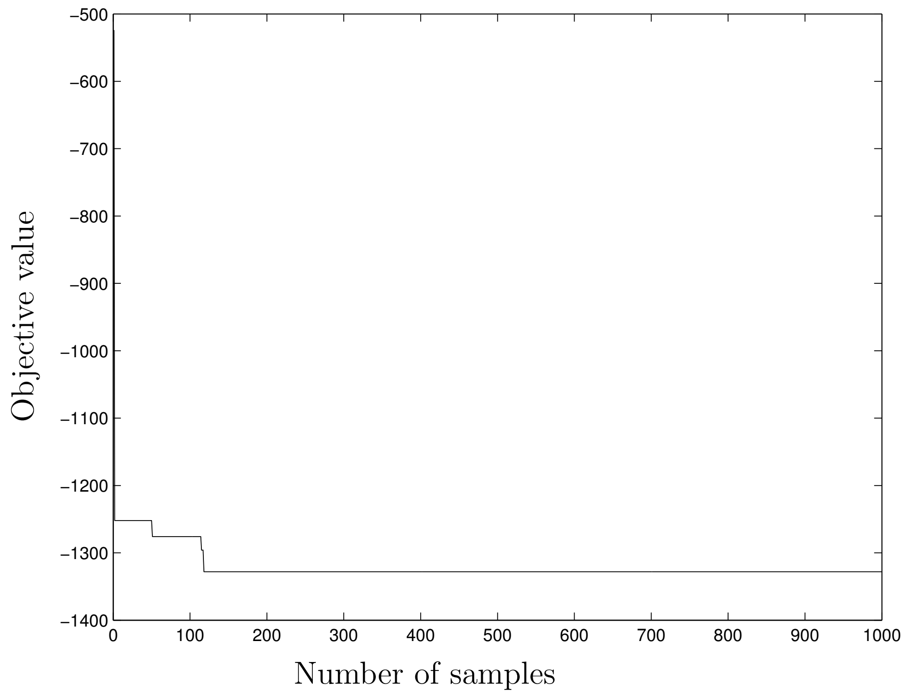


**Figure 3:** Best objective value versus number of sample points.

# References


[BFH00] S. P. Boyd, M. Fazel, and H. Hindi. A rank minimization heuristic with application to minimum order system approximation. *American Control Conference*, 2000.

[BTN01] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization. Analysis, Algorithms, and Engineering Applications*. Society for Industrial and Applied Mathematics, 2001.

[BV97] S. Boyd and L. Vandenberghe. Semidefinite programming relaxations of non-convex problems in control and combinatorial optimization. In A. Paulraj, V. Roychowdhuri, and C. Schaper, editors, *Communications, Computation, Control and Signal Processing: a Tribute to Thomas Kailath*, pages 279–288. Kluwer, 1997.

[BV03] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2003.

[Fer99] Eric Feron. Nonconvex quadratic programming, semidefinite relaxations and randomization algorithms in information and decision systems. In T. E. Djaferis and I. Schick, editors, *System theory: modeling analysis and control*. Kluwer, 1999.

[GW95] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the Association for Computing Machinery*, 42(6):1115–1145, 1995.

[LO99] Claude Lemaréchal and François Oustry. Semidefinite relaxations and lagrangian duality with application to combinatorial optimization. *INRIA, Rapport de recherche*, 3710, 1999.

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