Generating and Measuring Instances of Hard Semidefinite Programs, SDP

Hua Wei * Henry Wolkowicz †

January 23, 2006

University of Waterloo
Department of Combinatorics & Optimization
Waterloo, Ontario N2L 3G1, Canada
Research Report CORR 2006-01

Key words: Semidefinite Programming, Strict Complementarity, Max-Cut, primal-dual interior-point methods.

Abstract

Linear Programming, LP, problems with finite optimal value have a zero duality gap and a primal-dual strictly complementary optimal solution pair. On the other hand, there exists Semidefinite Programming, SDP, problems which have a nonzero duality gap (different primal and dual optimal values; not both infinite). The duality gap is assured to be zero if a constraint qualification, e.g Slater's condition (strict feasibility) holds. And, there exist SDP problems which have a zero duality gap but no strict complementary primal-dual optimal solution. We refer to these problems as *hard instances* of SDP.

In this paper, we introduce a procedure for generating hard instances of SDP. We then introduce two *measures of hardness* and illustrate empirically that these measures correlate well with the size of the gap in strict complementarity as well as with the asymptotic local convergence rate, and also with the number of iterations required to obtain optimal solutions to a specified accuracy. In addition, our numerical tests show that no correlation exists between the complementarity gaps and recently introduced geometrical measures or with Renegar's condition number. We include tests on the SDPLIB problem set.

*Research supported by The Natural Sciences and Engineering Research Council of Canada. E-mail h3wei@math.uwaterloo.ca

†Research supported by The Natural Sciences and Engineering Research Council of Canada. E-mail hwolkowicz@uwaterloo.ca

Contents

1 Introduction

Linear Programming, LP , problems with finite optimal value have a zero duality gap and a primal-dual strictly complementary optimal solution pair. On the other hand, there exists Semidefinite Programming, SDP , problems which have a nonzero duality gap (different primal and dual optimal values; not both infinite). The duality gap is assured to be zero if a constraint qualification, e.g Slater's condition (strict feasibility) holds. Measures of strict feasibility, also called *distance to infeasibility*, have been used in complexity analysis, e.g. [17, 5, 6, 7].

In addition, there exist SDP problems which have a zero duality gap but no strict complementary primal-dual optimal solution pair. We refer to these problems as *hard instances* of SDP . Similar to the lack of strict feasibility, the lack of strict complementarity can result in both theoretical and numerical difficulties. For example, many of the local superlinear and quadratic convergence results for interior-point methods depend on the strict complementarity assumption, e.g. [16, 11, 1, 14, 13]. Also, the convergence of the central path to the analytic center of the optimal face relies on strict complementarity, see [9].

In this paper we present an algorithm for generating *hard instances* of Semidefinite Programming, SDP , i.e. by hard we mean problems where strict complementarity fails. We use this set of hard problems to study the correlation between the loss of strict complementarity and the number of iterations needed to obtain optimality to a desired accuracy by interior-point algorithms. We compare and contrast our results to recent work by Ordóñez, Freund, and Toh [6], who found that the number of iterations needed by practical interior-point methods correlated well with the their aggregated geometrical measure as well as with

Renegar's condition number.

The primal SDP we consider is

$$\text{(PSDP)} \qquad \begin{array}{rll} p^* := & \min & \operatorname{trace} CX \\ & \text{s.t.} & \mathcal{A}(X) = b \\ & & X \succeq 0 \end{array} \tag{1.1}$$

and its dual

$$\text{(DSDP)} \qquad \begin{array}{rll} d^* := & \max & b^T y \\ & \text{s.t.} & \mathcal{A}^*(y) + Z = C \\ & & Z \succeq 0, \end{array} \tag{1.2}$$

where $C, X, Z \in \mathcal{S}^n$, $\mathcal{S}^n$ denotes the space of $n \times n$ real symmetric matrices, $y, b \in \Re^m$, and $\succeq$ ($\succ$) denotes positive semidefiniteness (resp. positive definiteness), known as the Löwner partial order; $\mathcal{A} : \mathcal{S}^n \rightarrow \Re^m$ is a (onto) linear transformation and $\mathcal{A}^*$ is the adjoint transformation. The set of optimal primal (resp. dual) solutions is denoted $\mathcal{P}^*$ (resp. $\mathcal{D}^*$).

The SDP model has important applications, elegant theory, and efficient solution techniques, see [20]. Moderate sized problems can be solved to near optimality using primal-dual interior-point (p-d i-p) methods. These methods are based on Newton's method with path following, i.e. the (Newton) search direction is found using a linearization of the (perturbed, symmetrized) optimality conditions. The iterates follow the central path, i.e. primal-dual feasible solutions with $ZX - \mu I = 0$, $\mu > 0$. On the central path, X and Z are mutually orthogonally diagonalizable, $X = QD_XQ^T, Z = QD_ZQ^T$; and their corresponding vectors of eigenvalues, $\lambda_X = \operatorname{diag}(D_X), \lambda_Z = \operatorname{diag}(D_Z)$, satisfy

$$\lambda_X \circ \lambda_Z = \mu e, \tag{1.3}$$

where $\circ$ denotes the *Hadamard* or elementwise product of the vectors, and $\operatorname{diag}(W)$ is the vector formed from the diagonal of W. The optimum dual pair of SDP is attained in the limit as $\mu \downarrow 0$; strict complementarity is indicated at $\mu = 0$ if $X + Z \succ 0$, i.e. strict positive definiteness. Therefore, as in linear programming, either $(\lambda_X)_i \downarrow 0, (\lambda_Z)_i \rightarrow \mathrm{O}(1)$ holds, or $(\lambda_Z)_i \downarrow 0, (\lambda_X)_i \rightarrow \mathrm{O}(1)$ holds. However, examples exist where the optimal X, Z have a nontrivial nullspace vector in common, i.e. strict complementarity fails. From (1.3), this means there exists i with both $(\lambda_Z)_i \downarrow 0, (\lambda_X)_i \downarrow 0$ but $(\lambda_Z)_i(\lambda_X)_i \cong \mu$, i.e. the value of each eigenvalue is order $\sqrt{\mu}$. For example, if the p-d i-p algorithm stops with a near optimal solution with duality gap $\mu = \operatorname{trace} ZX/n = \mathrm{O}(10^{-12})$, then we can expect the value of both eigenvalues to be as large as $\sqrt{\mu} = 10^{-6}$. In addition, the Jacobian of the optimality conditions at an optimum is singular, raising the question of slowed convergence. (See Remark 4.1.) These problems result in *hard instances* of SDP . P-d i-p methods typically run into difficulties such as slow (linear rate) convergence and low accuracy of the optimum.

1.1 Outline of Results and Main Contributions

In this paper we outline a procedure for generating hard instances of SDP . We then introduce two *measures of hardness*. We empirically show that: (i) these measures can be evaluated accurately; (ii) the size of the complementarity gaps correlate well with the number of iteration for the SDPT3 [19] solver, as well as with the local asymptotic convergence rate; and (iii) larger complementarity gaps correlate with loss of accuracy in the solutions. In addition, the numerical tests show that there is *no* correlation between the complementary gaps and the geometrical measure used in [6], or with Renegar's condition number.

We include tests on the SDPLIB problem set. Here we only found weak correlations due to lack of accuracy in the optimal solutions.

The procedure for generating hard problems has been submitted to the Decision Tree for Optimization Software, URL: plato.la.asu.edu/guide.html. See also SDPLIB e.g. [2], URL: www.nmt.edu/˜sdplib/. The MATLAB programs are available with URL: orion.math.uwaterloo.ca:80/˜hwolkowi/henry/software/readme.html

2 Generating Hard SDP Instances

In this section we show how to generate the *hard* SDP instances; i.e. the problems where strict complementarity fails.

Definition 2.1 *A primal-dual pair of optimal solutions* $(\bar{X}, \bar{Z}) \in \mathcal{P}^* \times \mathcal{D}^*$ *is called a* maximal complementary solution pair *to the problems (P) and (D), if the pair maximizes the sum* $\text{rank}\,(X) + \text{rank}\,(Z)$ *over all primal-dual optimal solution pairs* (X, Z).

A primal-dual pair of optimal solutions $(\bar{X}, \bar{S})$ is maximal complementary if and only if

$$\mathcal{R}(\hat{X}) \subseteq \mathcal{R}(\bar{X}),\ \forall \hat{X} \in \mathcal{P}^*, \qquad \mathcal{R}(\hat{S}) \subseteq (\bar{S}),\ \forall \hat{S} \in \mathcal{D}^*, \tag{2.1}$$

where $\mathcal{R}$ denotes range space. This follows from the fact that

$$\hat{X}\bar{S} = \bar{X}\hat{S} = \hat{X}\hat{S} = 0,\ \forall \hat{X} \in \mathcal{P}^*, \forall \hat{S} \in \mathcal{D}^*,$$

i.e. all optimal solution pairs are mutually orthogonally diagonalizable.

Definition 2.2 *The* strict complementarity gap *is defined as* $g = n - \text{rank}\,(\bar{X}) - \text{rank}\,(\bar{Z})$, *where* $(\bar{X}, \bar{Z})$ *is a maximal complementary solution pair.*

Note that g is equal to the minimum of the number of zero eigenvalues of $X + Z$, where the minimum is taken over all optimal solution pairs (X, Z).

For more details and proofs of these characterizations see [4], [8] and the references therein.

Algorithm 2.3 *Constructing* Hard SDP *Instances with gap g*

1. Given: positive integers $r > 0$ and $m > 1$ are the rank of an optimum X and the number of constraints, respectively.

2. Let $Q = [Q_P|Q_N|Q_D]$ be an orthogonal matrix, where the dimensions of Q_P, Q_N, Q_D are $n \times r$, $n \times g$, $n \times (n - r - g)$, respectively, and $r > 0$. Construct positive semidefinite matrices X and Z as follows:

$$X := Q_P D_X Q_P^T, \qquad Z := Q_D D_Z Q_D^T,$$

where D_X and D_Z are diagonal positive definite.

3. Define

$$A_1 = [Q_P|Q_N|Q_D] \begin{bmatrix} 0 & 0 & Y_2^T \\ 0 & Y_1 & Y_3^T \\ Y_2 & Y_3 & Y_4 \end{bmatrix} [Q_P|Q_N|Q_D]^T, \tag{2.2}$$

where Y_1, Y_2, Y_3, and Y_4 are block matrices of appropriate dimensions, $Y_1 \succ 0$, and $Q_D Y_2 \neq 0$.

4. Choose $A_i \in \mathcal{S}^n$, $i = 2, \ldots, m$, such that $\{A_1 Q_P, A_2 Q_P, \ldots, A_m Q_P\}$ is a linearly independent set. (Note that $A_1 Q_P = Q_D Y_2 \neq 0$.)

5. Set

$$b := \mathcal{A}(X), \qquad C := \mathcal{A}^*(y) + Z, \quad \text{with} \quad y \in \mathbb{R}^m \text{ randomly generated.}$$

Theorem 2.4 *The data* $(\mathcal{A}, b, C)$ *constructed in Algorithm 2.3 gives a* hard SDP *instance with a strict complementarity gap g.*

Proof. Step 2 guarantees that X, Z are positive semidefinite and $ZX = 0$ (complementary slackness holds). Step 5 guarantees that X, y, Z are primal-dual feasible. Therefore, our construction implies that $X, (y, Z)$ are a primal-dual optimal pair.

Choose any $\bar{X}, \bar{Z} \in \mathcal{P}^* \times \mathcal{D}^*$ with $\mathcal{R}(X) \subset \mathcal{R}(\bar{X})$ and $\mathcal{R}(Z) \subset \mathcal{R}(\bar{Z})$. We now show that $\mathcal{R}(X) = \mathcal{R}(\bar{X})$ and $\mathcal{R}(Z) = \mathcal{R}(\bar{Z})$, i.e. by (2.1) X, Z are a maximal complementary pair.

Since $\bar{X}$ and Z must also be an optimal pair, i.e $\bar{X}Z = 0$, we get that $\mathcal{R}(\bar{X}) \subseteq \mathcal{R}(Z)^\perp = \mathcal{R}([Q_P|Q_N])$. So, we can write

$$\bar{X} = [Q_P|Q_N] \begin{bmatrix} D_{P,\bar{X}} & W_{\bar{X}}^T \\ W_{\bar{X}} & D_{N,\bar{X}} \end{bmatrix} [Q_P|Q_N]^T,$$

where, in particular, $D_{N,\bar{X}} \succeq 0$. Let

$$\Delta X = \bar{X} - X = [Q_P|Q_N] \begin{bmatrix} D_{P,\bar{X}} - D_X & W_{\bar{X}}^T \\ W_{\bar{X}} & D_{N,\bar{X}} \end{bmatrix} [Q_P|Q_N]^T.$$

Since

$$\operatorname{trace}(A_1 \Delta X) = 0,$$

We have

$$\text{trace}\left(\left[Q_P|Q_N\right]^T A_1 \left[Q_P|Q_N\right] \begin{bmatrix} D_{P,\bar{X}} - D_X & W_{\bar{X}}^T \\ W_{\bar{X}} & D_{N,\bar{X}} \end{bmatrix}\right) = 0$$

From the structure of A_1, we see that

$$[Q_P|Q_N]^T A_1 [Q_P|Q_N] = \begin{bmatrix} 0 & 0 \\ 0 & Y_1 \end{bmatrix}.$$

So

$$0 = \text{trace}\,(A_1 \Delta X) = \text{trace}\left(\begin{bmatrix} 0 & 0 \\ 0 & Y_1 \end{bmatrix} \begin{bmatrix} D_{P,\bar{X}} - D_X & W_{\bar{X}}^T \\ W_{\bar{X}} & D_{N,\bar{X}} \end{bmatrix}\right) = \text{trace}\,\left(Y_1 D_{N,\bar{X}}\right).$$

By $Y_1 \succ 0$ and $D_{N,\bar{X}} \succeq 0$, we have that $D_{N,\bar{X}} = 0$. Since $\bar{X}$ is positive semidefinite, we have $W_{\bar{X}} = 0$ and $\mathcal{R}(\bar{X}) = \mathcal{R}(Q_P) = \mathcal{R}(X)$.

Similarly, we see that $\mathcal{R}(\bar{Z}) \subseteq \mathcal{R}(Q_N, Q_D)$. Let $\Delta Z = \bar{Z} - Z$ and $\Delta y = \bar{y} - y$, where $\mathcal{A}^*(\bar{y}) + \bar{Z} = C$. Then we have $\mathcal{A}^*(\Delta y) = -\Delta Z$. Since Q_P is a subspace of the the null space of both $\bar{Z}$ and Z, we have $-\Delta Z Q_P = 0$, i.e. $\mathcal{A}^*(\Delta y) Q_P = 0$. We write it in matrix form,

$$\sum_{i=1}^m A_i Q_P \Delta y_i = 0.$$

Since $\{A_i Q_P\}$ are nonzero and linearly independent, we see that $\Delta y_i = 0$ for all i. Thus, $\bar{Z} = Z$.

Therefore X, Z is also a maximal complementary pair. Since, by construction, $\text{rank}\,(X) + \text{rank}\,(Z) = n - g$, we have shown that the SDP is a hard instance with complementarity gap g. ∎

To avoid conflicts between loss of strict complementarity and loss of strict feasibility, we can use the following additional condition.

Corollary 2.5 *Suppose that the data $(\mathcal{A}, b, C)$ is constructed using Algorithm 2.3 with the additional condition that A_2 satisfies*

$$[Q_P|Q_N]^T A_2 [Q_P|Q_N] \succ 0. \tag{2.3}$$

Then Slater's condition holds for the dual program (DSDP).

Proof. Suppose that X, y, Z are as constructed by the algorithm. Then $Z = C - \mathcal{A}^*(y) = Q_D D_Z Q_D^T \succeq 0$. From [3, Lemma 7.1], we get that Slater's condition fails for (DSDP) if and only if

$$\exists R \succeq 0 \text{ with } R \neq 0, RZ = 0, \nabla_y \text{trace}\, R(\mathcal{A}^* y - C) = \mathcal{A}(R) = 0.$$

Now $RZ = 0$ implies that $R = [Q_P|Q_N] D_R [Q_P|Q_N]^T$, for some symmetric D_R of appropriate size. Therefore, $\mathcal{A}(R) = 0$ implies that

$$0 = \text{trace}\, A_2 R = \text{trace}\, A_2 [Q_P|Q_N] D_R [Q_P|Q_N]^T = \text{trace}\,([Q_P|Q_N]^T A_2 [Q_P|Q_N]) D_R.$$

The assumption (2.3) now implies that $D_R = 0$ and so also $R = 0$. Therefore, Slater's condition holds. ∎

3 Measures for Strict Complementarity Gaps

In [6], the authors indicate the following difficulties in measuring the existence and size of the strict complementarity gap.

> "Furthermore, in interior-point methods for either linear or semidefinite programming, we terminate the algorithm with a primal-dual solution that is almost optimal but not actually optimal. Hence there are genuine conceptual difficulties in trying to quantify and compute the extent of near-non-strict-complementarity for an SDP instance."

The measure, κ in (3.6), is proposed in [6]. However, this measure does not distinguish between a small or large complementarity gap g. But, as our numerical tests in Section 4 indicate, large values of g are well correlated with large iterations numbers. This motivates the introduction of our following two new measures.

3.1 Complementarity Gap Measures g_t and g_s

3.1.1 Measure g_t

For barrier parameter $\mu > 0$, $\mu \downarrow 0$, and corresponding feasible pairs $X = X_\mu, Z = Z_\mu$ on the central path, let the orthogonal eigenvalue decomposition be $X = Q\Lambda_X Q^T$ and $Z = Q\Lambda_Z Q^T$. Consider the eigenvalue ratios $w_i^d := \Lambda_{Z i}/\Lambda_{X i}$. Then

$$XZ = Q\Lambda_X Q^T Q\Lambda_Z Q^T = \Lambda_X\Lambda_Z = \mu I, \qquad w_i^d = \frac{\mu}{(\Lambda_X)_i^2}.$$

Suppose that $X \to \bar{X}$, $Z \to \bar{Z}$. We then expect the following behaviour.

$$w_i^d \to \begin{cases} \infty & \text{if } \Lambda_{\bar{X}i} + \Lambda_{\bar{Z}i} > 0 \text{ (no gap) and } \Lambda_{Xi} \to 0 \\ 0 & \text{if } \Lambda_{\bar{X}i} + \Lambda_{\bar{Z}i} > 0 \text{ (no gap) and } \Lambda_{Zi} \to 0 \\ O(1) & \text{if } \Lambda_{\bar{X}i} + \Lambda_{\bar{Z}i} = 0 \text{ (a gap).} \end{cases}$$

Empirical evidence suggests that the sequence $\{w_i^d\}$ converges when there is a complementarity gap. The measure we define exploits this behaviour. In practice, we use the vector of eigenvalues

$$w^d = \frac{1}{2}\lambda\left(X^{-1}Z + ZX^{-1}\right). \tag{3.1}$$

(Note that the eigenvalues of $X^{-1}Z$ interlace the eigenvalues of $\frac{1}{2}(X^{-1}Z + ZX^{-1})$, e.g. [15].) For given tolerances T_u and T_l, we estimate the strict complementarity gap using the cardinality

$$g_t := |\{w_i^d : T_l < w_i^d < T_u\}|. \tag{3.2}$$

3.1.2 Measure g_s

The second measure exploits the idea from [6]. As in Section 3.1.1, we let X and Z to be a solution pair on the central path corresponding to $\mu > 0$. The eigenvalue decompositions of X and Z are $X = Q\Lambda_X Q^T$ and $Z = Q\Lambda_Z Q^T$. The measure uses the numerical rank (e.g. [18, 10]) of $X + Z$. We compute

$$w^s := \frac{1}{2\sqrt{\mu}}\lambda\left(X + Z\right), \tag{3.3}$$

where $\mu = \operatorname{trace} ZX/n$. Given a tolerance $T > 0$, we estimate the strict complementarity gap g using the cardinality

$$g_s := |\{w_j^s : w_j^s \leq T\}|. \tag{3.4}$$

Remark 3.1 *Note that on the central path, X, Z are mutually diagonalizable. Therefore, the eigenvalues of the sum $X + Z$ is the same as the sum of the eigenvalues. However, this is not necessarily true off the central path, see the recent paper [12]. In this remarkable paper the author solves a classical problem about the eigenvalues of sums of Hermitian operators, connecting it to the Schubert calculus for the homology of Grassmannians and the moduli of vector bundles.*

However, we should point out that this measure g_s may incorrectly include some indices which do not belong to the complementarity gap when the solution estimates X, Z are not accurate enough. Consider the following results from a randomly generated problem instance with strict complementarity gap 1. The first 7 eigenvalues from the solution estimates X and Z (obtained using SDPT3) are

$$\lambda_X = \begin{bmatrix} 7.3 \times 10^{-7} \\ 1.8 \times 10^{-6} \\ 2.0 \times 10^{-6} \\ 1.4 \times 10^{-5} \\ 6.2 \\ 8.6 \times 10^{3} \\ 9.2 \times 10^{3} \end{bmatrix}, \quad \lambda_Z = \begin{bmatrix} 65 \\ 59 \\ 54 \\ 2.4 \\ 2.1 \times 10^{-5} \\ 4.7 \times 10^{-9} \\ 4.5 \times 10^{-9} \end{bmatrix}. \tag{3.5}$$

Note that

$$(\lambda_X)_4 << (\lambda_X)_5 << (\lambda_X)_6, \qquad (\lambda_Z)_6 << (\lambda_Z)_5 << (\lambda_Z)_4,$$

i.e. the fifth elements are relatively small/large compared to the next/previous larger/smaller elements. This indicates that there is a strict complementarity gap $g = 1$. However, the sum of these two eigenvalues fails to correctly estimate the size of the gap,

$$\lambda_X + \lambda_Z = \begin{bmatrix} 65 \\ 59 \\ 54 \\ 2.4 \\ 6.2 \\ 8.6 \times 10^{3} \\ 9.2 \times 10^{3} \end{bmatrix}.$$

Higher accuracy in the approximate optimal solutions X, Z often corrects this issue, see the numerics in Section 4.

3.2 Measure κ

The last measure we introduce is κ used in [6]. For a given tolerance T, define the following index set $T^s := \{j : w_j^s \leq T\}$, where w^s is defined in (3.3). Then

$$\kappa := -\sum_{j \in T^s} \ln(w_j^s)/|T^s|. \tag{3.6}$$

When strict complementarity holds (resp. fails), we expect to see a relatively large (resp. small) κ.

4 Numerics

We now compare the various measures on randomly generated instances with guaranteed strict complementarity gaps as well as on problems from the SDPLIB test set, [2].

4.1 Randomly Generated Instances

We use Algorithm 2.3 to generate the hard instances. To implement the algorithm, we generate a random orthogonal matrix Q and random diagonal D_X and D_Z. The elements of D_X and D_Z are uniformly distributed in the range of $[0.1, 100.1]$ to ensure positivity of D_X and D_Z. The optimal solution of this hard SDP instance is then determined from step 2 in Algorithm 2.3. For the special matrix A_1, we construct Y_1 according to:

1. generate the random symmetric matrix Y_1 with uniformly distributed elements in $[-10000, 10000]$;
2. add rI to the above matrix Y_1, where r is a random number in $[0, 20000]$;
3. if Y_1 is *not* sufficiently positive definite, repeat the process from step 1.

All the elements of the random matrices Y_2, Y_3, and Y_4, are uniformly distributed in the interval $[-10000, 10000]$. If necessary, we symmetrize the matrices. If $Q_D Y_2$ is close to a zero matrix, we repeat the process for Y_2. Our special matrix A_1 is then constructed from Step 3 in Algorithm 2.3. Once we have such a special matrix, we generate random symmetric uniformly distributed matrices A_j. If one of the $A_j Q_p$ is not properly linearly independent, then we add a new A_j to the list. To guarantee that Slater's condition holds, we apply the condition in Corollary 2.5.

We present the average of results from 100 groups of tests. Each group consists of SDP instances with 26 different gap values. We set the following parameters: $m = 10, n = 30, \text{gap} = 0, \ldots, 25$. The rank for the dual optimal solution is fixed at 4. The name of the instance shows how large the gap is, e.g. gap5. The accuracy of solutions is given by the err term:

$$\text{err} := \max\left\{ \frac{\|\mathcal{A}(X) - b\| + |\min(\text{eig}(X), 0)|}{1 + \|b\|_\infty}, \frac{\|\mathcal{A}^*(y) + Z - C\| + |\min(\text{eig}(Z), 0)|}{1 + \|C\|_\infty}, \frac{|C \cdot X - b^T y|}{1 + |b^T y|} \right\} \tag{4.1}$$

When computing g_t (3.2), we set the tolerances T_u, T_l dynamically. More precisely, we sort the w_i^d (3.1) in ascending order. We then use the ratios $\bar{w}_i^d := w_i^d / w_{i+1}^d$, to measure how fast the w_i^d are changing. If there is only one small (< 0.02) $\bar{w}_i^d$, we assume that there is no gap. Otherwise, we find the two smallest valued $\bar{w}_i^d$ and set the two indices to be j and k ($j < k$). Then $T_l := w_j + \epsilon$ and $T_u := w_k + \epsilon$. In practice, once we have found the indices j and k, the estimated gap g_t is returned by using the value of $k - j$.

When computing g_s in (3.4), we set the tolerance $T = \max\{100, \min_i(w_i^s)\}$, where w^s is defined in (3.3). The tolerance T for the measure κ is the same as the one used for g_s. This is the same tolerance as that used in [6].

4.2 Plots for Randomly Generated Instances

To illustrate the relationships among the various measures we consider three groups of figures. To illustrate the influence of accuracy in the solutions, each group consists of three figures with decreasing stop tolerances 10^{-8}, 10^{-10}, and 10^{-12}, respectively.

The x-axis of each figure represents the complementarity gap ranging from 0 to 24. The y-axes, from left to right, represent, respectively:

iteration numbers,
negative log (base 10) of errors (4.1),
measure g_t,
measure g_s,
measure κ,
local convergence rate (discussed in Item 4 on page 12).

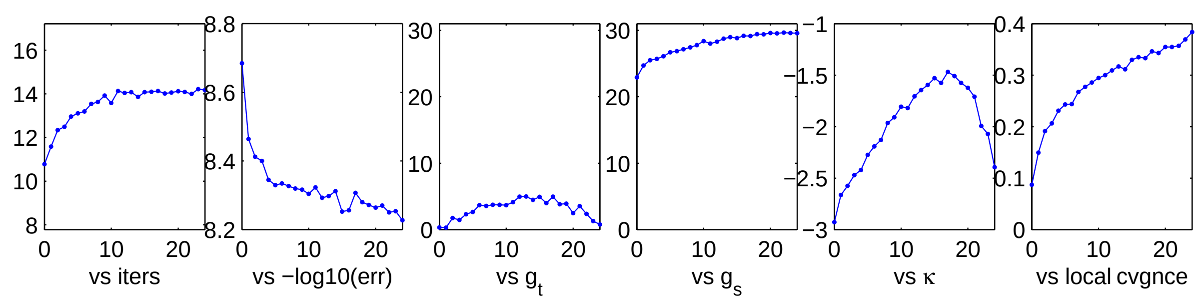

Figure 4.1: Slater's holds; stop tolerance 10^{-8}; complementarity gaps from 0 to 24 **versus**: iterations, $-\log_{10}$ err, g_t, g_s, κ, local convergence; 100 instances.

- The first three Figures 4.1, 4.2, 4.3, are average results from 100 instances. We apply Corollary 2.5 to guarantee that Slater's condition holds.
- The next three figures 4.4, 4.5, 4.6, show the behaviour of a typical single instance without applying Corollary 2.5. We see in Table 4.2 that Slater's condition generally holds for the primal but generally fails for the dual.

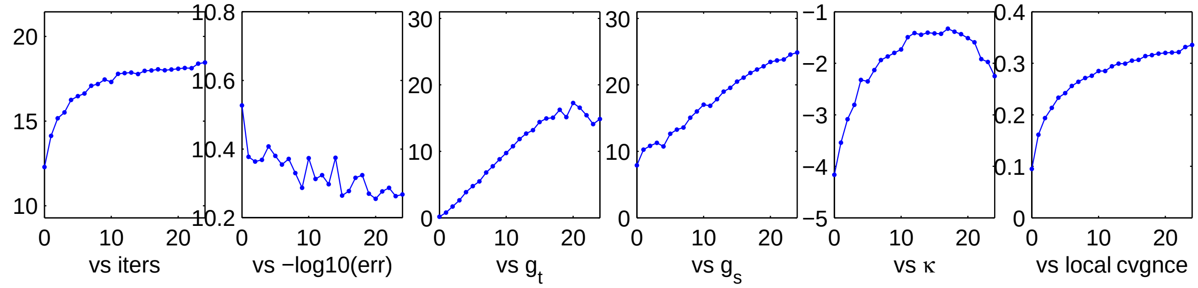

Figure 4.2: Slater's holds; stop tolerance 10^{-10}; complementarity gaps from 0 to 24 **versus**: iterations, $-\log_{10}$ err, g_t, g_s, κ, local convergence; 100 instances.

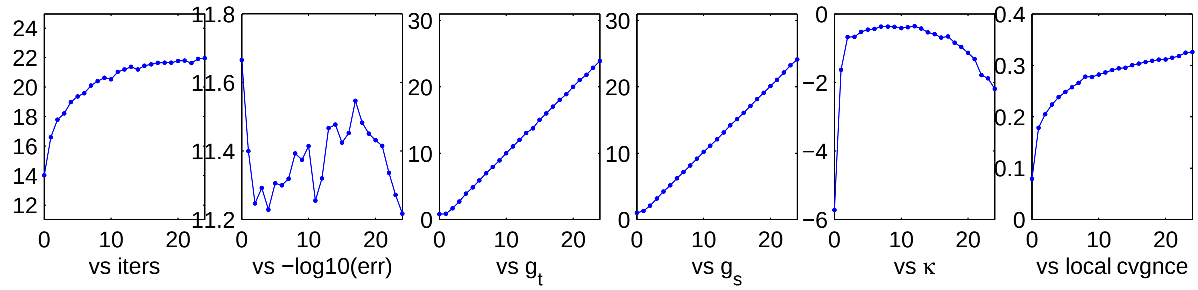

Figure 4.3: Slater's holds; stop tolerance 10^{-12}; complementarity gaps from 0 to 24 **versus**: iterations, $-\log_{10}$ err, g_t, g_s, κ, local convergence; 100 instances.

- The last three Figures 4.7,4.8,4.9 consider the average behaviour on 100 instances. Again, we do not apply Corollary 2.5.

Observations from the nine figures 4.1 to 4.9:

1. There is a *strong correlation* between the iteration number to achieve the desired stopping tolerance and the size of the complementarity gap. The correlation is stronger when the desired stopping tolerance is smaller.

2. The measures g_t, g_s both improve dramatically as the accuracy increases in Figures 4.1,

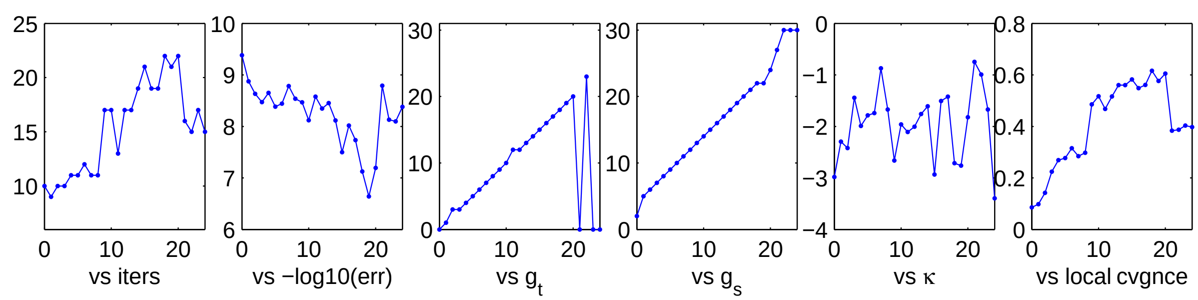

Figure 4.4: Slater's generally fails; stop tolerance 10^{-8}; complementarity gaps from 0 to 24 **versus**: iterations, $-\log_{10}$ err, g_t, g_s, κ, local convergence; single instance.

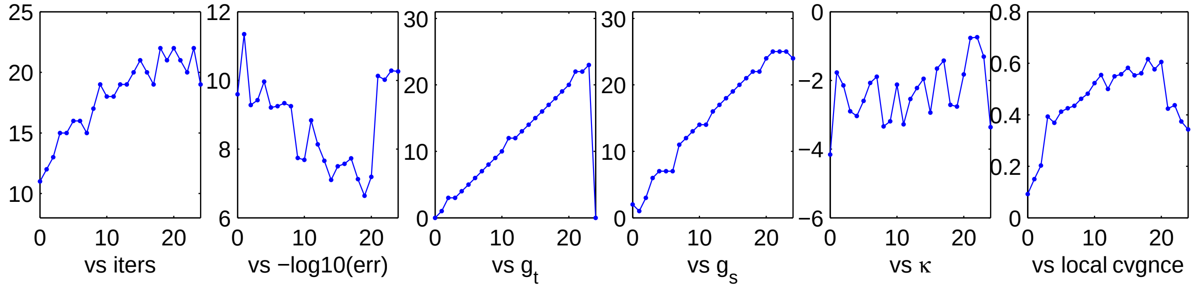

Figure 4.5: Slater's generally fails; stop tolerance 10^{-10}; complementarity gaps from 0 to 24 **versus**: iterations, $-\log_{10}$ err, g_t, g_s, κ, local convergence; single instance.

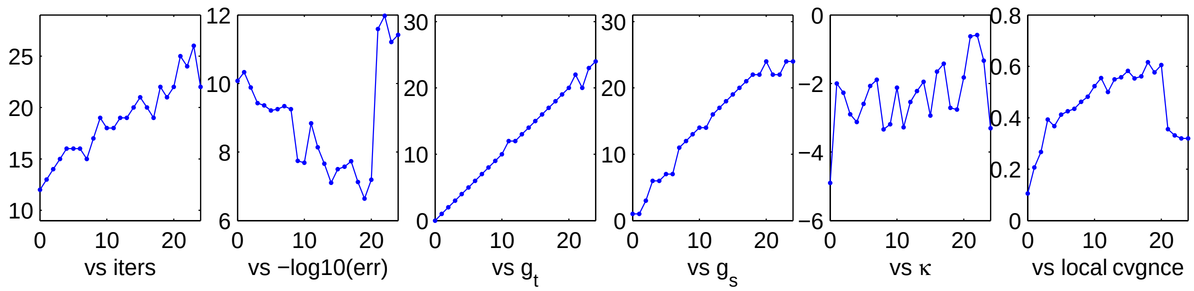

Figure 4.6: Slater's generally fails; stop tolerance 10^{-12}; complementarity gaps from 0 to 24 **versus**: iterations, $-\log_{10}$ err, g_t, g_s, κ, local convergence; single instance.

4.2, 4.3. We see this same phenomenon in the other two groups of figures.

3. The measure κ also improves with smaller stopping tolerances.

4. Local Asymptotic Convergence Rate vs Complementarity Gap In the literature, e.g. [16] [11] [1] [14] [13], local superlinear or quadratic convergence results depend on the assumption of strict complementarity. Thus it is intuitive to expect this in practice as well. Our numerical results confirm this conjecture. The convergence rate is defined by the ratio of the relative duality gap at successive iterations. We list the geometrical

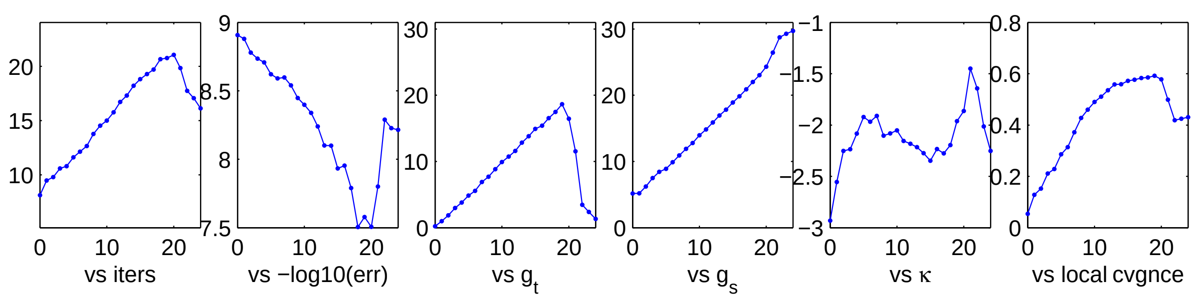

Figure 4.7: Slater's fails; stop tolerance 10^{-8}; complementarity gaps from 0 to 24 **versus average of**: iterations, error, g_t, g_s, κ, local convergence; 100 instances.

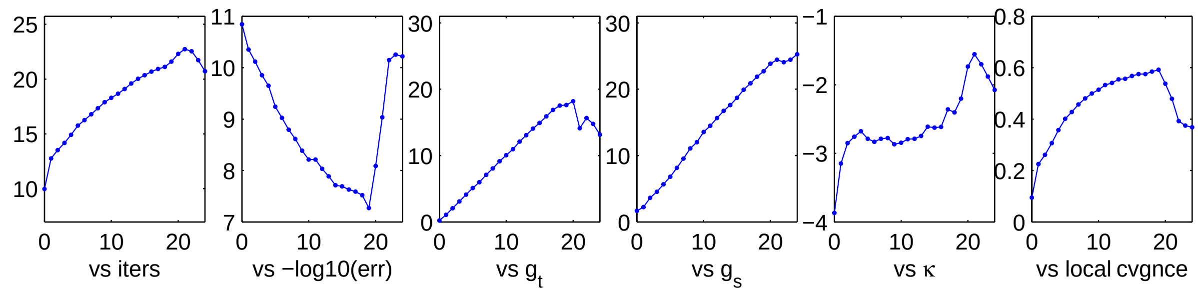

Figure 4.8: Slater's fails; stop tolerance 10^{-10}; complementarity gaps from 0 to 24 **versus average of**: iterations, error, g_t, g_s, κ, local convergence; 100 instances.

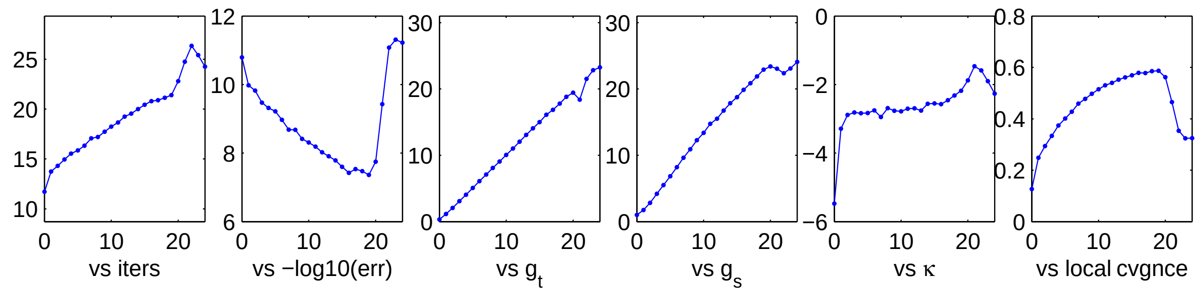

Figure 4.9: Slater's fails; stop tolerance 10^{-12}; complementarity gaps from 0 to 24 **versus average of**: iterations, error, g_t, g_s, κ, local convergence; 100 instances.

mean of the convergence rate for the last five iterations. This is illustrated in the rightmost picture in the figures.

Remark 4.1 *The slow convergence rates can be partially explained by the singularity of the Jacobian, which occurs in the presence of a complementarity gap.*

Suppose that strict complementarity fails for the optimum pair estimate X, Z. Then we can assume the joint diagonalization structure

$$X = Q \begin{bmatrix} D_X & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} Q^T, \qquad Z = Q \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & D_Z \end{bmatrix} Q^T$$

for some orthogonal matrix Q and positive definite diagonal matrices D_X, D_Z. Then we can rewrite the Jacobian of the SDP *optimality conditions as*

$$\begin{bmatrix} 0 & \bar{\mathcal{A}}^* & I \\ \bar{\mathcal{A}} & 0 & 0 \\ \begin{bmatrix} D_X & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} & 0 & \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & D_Z \end{bmatrix} \end{bmatrix} \begin{bmatrix} \Delta \bar{X} \\ \Delta \bar{y} \\ \Delta \bar{Z} \end{bmatrix} = 0,$$

where $\Delta \bar{X} = Q^T \Delta X Q, \Delta \bar{Z} = Q^T \Delta Z Q$ and the symmetric matrices A_i defining the linear transformation $\mathcal{A}$ are changed to $Q^T A_i Q$ for $\bar{\mathcal{A}}$. If we assume that both $\Delta \bar{X}, \Delta \bar{Z}$ are diagonal, then this reduces the problem to an ordinary square system and the resulting Jacobian

Problem	its	$-\log_{10}$(err)	g_t	g_s	κ	local convergence
gap0	10	10.7	0	2	-3.9	0.09
gap1	13	9.8	1	2	-3.1	0.22
gap2	14	8.8	2	4	-2.9	0.26
gap3	14	9.2	3	5	-2.8	0.31
gap4	15	9.2	4	6	-2.7	0.36
gap5	16	7.7	5	7	-2.8	0.40
gap6	16	7.0	6	8	-2.8	0.43
gap7	17	8.1	7	10	-2.8	0.46
gap8	17	6.7	8	11	-2.8	0.48
gap9	18	6.5	9	12	-2.9	0.50
gap10	18	7.2	10	14	-2.8	0.51
gap11	19	7.7	11	14	-2.8	0.53
gap12	19	6.7	12	16	-2.8	0.54
gap13	20	6.4	13	17	-2.7	0.55
gap14	20	6.7	14	18	-2.6	0.56
gap15	20	6.9	15	19	-2.6	0.57
gap16	21	7.1	16	20	-2.6	0.58
gap17	21	6.9	17	21	-2.4	0.57
gap18	21	7.1	18	22	-2.4	0.58
gap19	22	4.9	18	23	-2.2	0.59
gap20	22	6.8	18	24	-1.7	0.54
gap21	23	7.1	14	24	-1.6	0.48
gap22	23	8.7	16	24	-1.7	0.39
gap23	22	9.9	15	24	-1.9	0.37
gap24	21	10.2	13	25	-2.1	0.37

Table 4.1: The data corresponding to Figure 4.9.

is singular due to the zero row. The diagonal assumption does not change the feasibility of the first and third blocks of equations. We can then modify the off-diagonal part of $\Delta\bar{Z}$ to guarantee the feasibility of the second block of equations.

The singularity of the Jacobian means that we should expect loss of both quadratic and superlinear convergence for Newton type methods.

4.2.1 Geometrical Measure vs Large Complementarity Gaps

In [6], the authors use SDPT3 and the SDPLIB test set. They show that the aggregate geometrical measure g^m, i.e. the geometric mean of the four geometric measures D_p, g_p, D_d, g_d, in Table 4.2, is (generally) well correlated with the iteration number. They also show that the correlation holds for Renegar's condition number, see also Table 4.3. The values for these measures for the SDP instance in Figures 4.7,4.8,4.9 are given in Tables 4.2, 4.3. For details on the geometrical measure and Renegar's condition number and their computation, please see [6] and the references therein. We use the same code used in [6] to compute the geometrical measure g^m and Renegar's condition number. [1]

As pointed out in [6], the strict complementarity gap might not be theoretically related to the geometrical measures or Renegar's condition number. In fact, our numerical computations on our generated instances confirm this, see Tables 4.2, 4.3, i.e. though these geometric measures correlate well with the iteration number, they do not correlate well with the size of the complementarity gap.

4.3 SDPLIB Instances

Our results in Section 4 show that, generally, measure g_t can accurately measure the gap g, though it can give large errors when the solution estimates are not accurate enough. The measure g_s is more consistent in measuring the strict complementarity gap, g. The measure κ is also sensitive to the accuracy of the solution.

We applied these measures g_t, g_s, and κ to the SDPLIB [2] problem set. Though we used 10^{-10} as the stop tolerance in SDPT3, it was rarely attained. For some of the problems, there were big discrepancies between the two measures g_t and g_s. There was also no significant correlation between the iteration numbers and the three measures:

$$\text{corr}\,(g_t, \text{its}\,) = -0.01, \;\; \text{corr}\,(g_s, \text{its}\,) = -0.067, \;\; \text{and corr}\,(\kappa, \text{its}\,) = 0.2856.$$

However, if we only consider those SDP instances (47 such instances), where the error obtained was less than 10^{-7}, then we see a significant increase in the correlations between the measures and the iteration numbers:

$$\text{corr}\,(g_t, \text{its}\,) = 0.1472, \;\; \text{corr}\,(g_s, \text{its}\,) = 0.4509, \;\; \text{and corr}\,(\kappa, \text{its}\,) = 0.4371.$$

Their plots are showed in Figure 4.10.

[1] **Acknowledgement:** The authors thank Professor Ordóñez, University of Southern California, for providing the software for the measure evaluations.

[2] In [6] it is shown that $g_d = \infty \iff \rho_D(d) = 0$. However, due to inaccuracy from SDPT3, we get inconsistencies here.

Problem	D_p	g_p	D_d	g_d	g^m
gap0	Inf	1.9e+02	1.5e+02	Inf	Inf
gap1	1.2e+12	1.5e+02	2.4e+02	Inf	Inf
gap2	2.9e+08	1.6e+02	2.3e+02	Inf	Inf
gap3	3.1e+08	1.4e+02	1.7e+02	Inf	Inf
gap4	1.1e+11	1.8e+02	1.9e+02	MAXIT	N/A
gap5	1.6e+08	1.1e+02	2.4e+02	Inf	Inf
gap6	4.5e+08	9.9e+01	2.8e+02	Inf	Inf
gap7	1.9e+08	1.6e+02	1.1e+02	Inf	Inf
gap8	1.4e+09	2.1e+02	1.3e+02	Inf	Inf
gap9	1.4e+09	1.8e+02	1.8e+02	Inf	Inf
gap10	2.1e+09	1.3e+02	4.0e+02	6.2e+04	4.7e+00
gap11	1.0e+09	1.7e+02	1.4e+02	Inf	Inf
gap12	Inf	1.3e+02	3.2e+02	Inf	Inf
gap13	Inf	1.6e+02	2.6e+01	Inf	Inf
gap14	Inf	1.6e+02	Nacc	Inf	N/A
gap15	Inf	1.7e+02	6.9e+01	Inf	Inf
gap16	3.0e+10	2.5e+02	2.2e+02	Inf	Inf
gap17	Inf	2.6e+02	2.1e+02	Inf	Inf
gap18	Inf	1.5e+02	2.6e+02	Inf	Inf
gap19	1.2e+10	1.1e+02	2.6e+02	Inf	Inf
gap20	6.3e+10	1.8e+02	2.3e+02	Inf	Inf
gap21	1.2e+10	2.2e+02	1.4e+02	Inf	Inf
gap22	2.7e+02	1.2e+02	2.3e+02	MAXIT	N/A
gap23	1.8e+02	2.5e+02	1.4e+02	Nacc	N/A
gap24	3.6e+01	2.5e+02	1.1e+02	MAXIT	N/A

Table 4.2: Notation from [6]: (D_p, g_p) - primal geometrical measure; (D_d, g_d) - dual geometrical measure; (g^m) - aggregate geometrical measure, i.e. geometrical mean of D_p, g_p, D_d, and g_d. MAXIT - max iteration limit reached; Nacc - no accurate/meaningful solution.

Problem	$\rho_P(d)$	$\rho_D(d)$	$\|d\|_l$	$\|d\|_u$	$C(d)_l$	$C(d)_u$
gap0	2.8e+04	7.4e-04	1.1e+09	1.1e+09	1.5e+12	1.5e+12
gap1	3.1e+04	9.9e-04	2.2e+09	2.2e+09	2.2e+12	2.2e+12
gap2	2.9e+04	1.3e-03	2.5e+09	2.5e+09	2.0e+12	2.0e+12
gap3	3.2e+04	2.5e-04	7.3e+08	7.3e+08	2.9e+12	2.9e+12
gap4	3.4e+04	1.1e-03	8.0e+08	8.0e+08	7.3e+11	7.3e+11
gap5	2.9e+04	2.4e-03	8.1e+08	8.1e+08	3.4e+11	3.4e+11
gap6	3.0e+04	1.9e-04	1.0e+09	1.0e+09	5.3e+12	5.3e+12
gap7	3.0e+04	1.4e-03	4.3e+09	4.3e+09	3.0e+12	3.0e+12
gap8	3.1e+04	2.4e-04	1.1e+09	1.1e+09	4.6e+12	4.6e+12
gap9	2.7e+04	2.6e-03	3.2e+09	3.2e+09	1.3e+12	1.3e+12
gap10	3.1e+04	4.2e-03	8.5e+08	8.5e+08	2.0e+11	2.0e+11
gap11	3.2e+04	2.6e-04	4.3e+09	4.3e+09	1.7e+13	1.7e+13
gap12	2.8e+04	6.7e-03	1.9e+09	1.9e+09	2.9e+11	2.9e+11
gap13	2.5e+04	1.1e-03	6.9e+08	6.9e+08	6.1e+11	6.1e+11
gap14	2.4e+04	6.4e-03	9.8e+08	9.8e+08	1.5e+11	1.5e+11
gap15	2.5e+04	2.8e-04	2.1e+09	2.1e+09	7.2e+12	7.2e+12
gap16	2.4e+04	3.1e-03	5.0e+09	5.0e+09	1.6e+12	1.6e+12
gap17	2.4e+04	2.4e-04	7.1e+08	7.1e+08	3.0e+12	3.0e+12
gap18	2.1e+04	3.0e-04	7.1e+08	7.1e+08	2.3e+12	2.3e+12
gap19	2.5e+04	5.1e-03	1.9e+09	1.9e+09	3.7e+11	3.7e+11
gap20	2.0e+04	4.2e-03	1.4e+09	1.4e+09	3.3e+11	3.3e+11
gap21	1.6e+04	1.1e-03	4.1e+09	4.1e+09	3.7e+12	3.7e+12
gap22	2.3e+04	4.0e-03	7.0e+08	7.0e+08	1.7e+11	1.7e+11
gap23	1.5e+04	1.9e-03	4.5e+09	4.5e+09	2.3e+12	2.3e+12
gap24	1.5e+04	8.0e-03 [2]	4.4e+09	4.4e+09	5.4e+11	5.4e+11

Table 4.3: Renegar's condition number on SDPs with complementarity gaps. Notation from [6]: ($\rho_P(d)$) - distance to primal infeasibility; ($\rho_D(d)$) - distance to dual infeasibility; ($\|d\|_l, \|d\|_u$) - lower and upper bounds of the norm of the data; ($C(d)_l, C(d)_u$) - lower and upper bounds on Renegar's condition number, $C(d) = \frac{\|d\|}{\min\{\rho_P(d), \rho_D(d)\}}$.

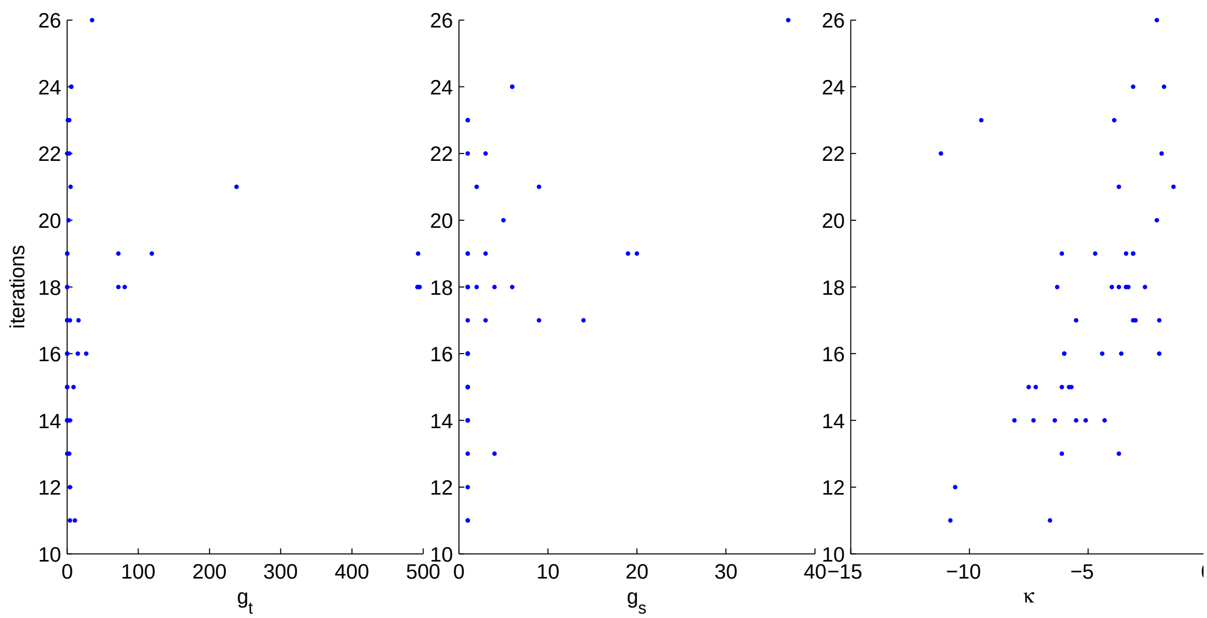

Figure 4.10: Scatter plots of g_t, g_s, κ versus # iterations for SDPLIB instances with attained tolerance $< 10^{-7}$.

5 Conclusion

We have presented an algorithm for generating hard SDP instances, i.e. problem instances where we can control the strict complementarity gap, g. We then tested several measures on randomly generated instances. The tests confirm the intuitive expectation: *The number of iterations for interior-point methods are closely related to the size of the complementarity gaps.* In addition, we tested three measures g_t, g_s, and κ on the generated hard SDP instances. These measures g_t, g_s generally provide accurate measurement of the strict complementarity gaps; with the measure g_s being more consistent. All three measures are negatively affected by inaccurate solution estimates.

We also tested the aggregated geometrical measure and Renegar's condition number on the generated hard SDP instances; and, we did not find any correlation between them and the size of the complementarity gaps. It appears that these geometric measures are more closely related to distance to infeasibility, i.e. strict feasiblity. One class of generated hard SDP instances have consistently large aggregate geometrical measure and large Renegar condition number, despite having different complementarity gap values.

Finally, we used the SDPLIB test set but had trouble coming to any concrete conclusions since the approximate solutions we found were not accurate enough. We hope to obtain improved solutions and redo these tests in the future.

Acknowledgement: The authors thank Professor Robert Freund from MIT - Massachusetts Institute of Technology, for helpful talks that improved the paper. In particular, the authors thank Professor Freund for pointing out the sudden drop in values for D_p when the complementarity gap is large; see column 2 in Table 4.2. This led to the addition of Corollary 2.5, as well as the addition of the numerical tests that satisfied Slater's condition. The authors also thank Professor Franz Rendl, from Universitaet Klagenfurt, for his collaboration on a preliminary version of this work.

References

[1] F. ALIZADEH, J-P.A. HAEBERLY, and M.L. OVERTON. Primal-dual interior-point methods for semidefinite programming: Convergence rates, stability and numerical results. *SIAM J. Optim.*, 8:746–768, 1998.

[2] B. BORCHERS. SDPLIB 1.2, a library of semidefinite programming test problems. *Optim. Methods Softw.*, 11(1):683–690, 1999. Interior point methods.

[3] J.M. BORWEIN and H. WOLKOWICZ. Regularizing the abstract convex program. *J. Math. Anal. Appl.*, 83(2):495–530, 1981.

[4] E. de KLERK, C. ROOS, and T. TERLAKY. Initialization in semidefinite programming via a self-dual skew-symmetric embedding. *Oper. Res. Lett.*, 20(5):213–221, 1997.

[5] R.M. FREUND. Complexity of convex optimization using geometry-based measures and a reference point. *Math. Program.*, 99(2, Ser. A):197–221, 2004.

[6] R.M. FREUND, F. ORDÓÑEZ, and K.C. TOH. Behavioral measures and their correlation with IPM iteration counts on semi-definite programming problems. USC-ISE working paper #2005-02, MIT, 2005. url: http://www-rcf.usc.edu/˜fordon/.

[7] R.M. FREUND and J.R. VERA. Condition-based complexity of convex optimization in conic linear form via the ellipsoid algorithm. *SIAM J. Optim.*, 10(1):155–176 (electronic), 1999.

[8] D. GOLDFARB and K. SCHEINBERG. Interior point trajectories in semidefinite programming. *SIAM J. Optim.*, 8(4):871–886, 1998.

[9] M. HALICKÁ, E. de KLERK, and C. ROOS. On the convergence of the central path in semidefinite optimization. *SIAM J. Optim.*, 12(4):1090–1099 (electronic), 2002.

[10] P.C. HANSEN and P.Y. YALAMOV. Computing symmetric rank-revealing decompositions via triangular factorization. *SIAM J. Matrix Anal. Appl.*, 23(2):443–458 (electronic), 2001.

[11] J. JI, F.A. POTRA, and R. SHENG. On the local convergence of a predictor-corrector method for semidefinite programming. *SIAM J. Optim.*, 10(1):195–210 (electronic), 1999.

[12] A.A. KLYACHKO. Stable bundles, representation theory and Hermitian operators. *Selecta Math. (N.S.)*, 4(3):419–445, 1998.

[13] M. KOJIMA, M. SHIDA, and S. SHINDOH. Local convergence of predictor-corrector infeasible-interior-point algorithms for SDPs and SDLCPs. Technical report, Dept. of Information Sciences, Tokyo Institute of Technology, Tokyo, Japan, 1996.

[14] Z-Q. LUO, J. F. STURM, and S. ZHANG. Superlinear convergence of a symmetric primal-dual path-following algorithm for semidefinite programming. *SIAM J. Optim.*, 8:59–81, 1998.

[15] A.W. MARSHALL and I. OLKIN. *Inequalities: Theory of Majorization and its Applications.* Academic Press, New York, NY, 1979.

[16] F.A. POTRA and R. SHENG. Superlinear convergence of a predictor-corrector method for semidefinite programming without shrinking central path neighborhood. *Bull. Math. Soc. Sci. Math. Roumanie (N.S.)*, 43(91)(2):107–124, 2000.

[17] J. RENEGAR. Linear programming, complexity theory and elementary functional analysis. *Math. Programming*, 70(3, Ser. A):279–351, 1995.

[18] G.W. STEWART. Updating a rank-revealing ULV decomposition. *SIAM J. Matrix Anal. Appl.*, 14(2):494–499, 1993.

[19] R.H. TÜTÜNCÜ, K.C. TOH, and M.J. TODD. Solving semidefinite-quadratic-linear programs using SDPT3. *Math. Program.*, 95(2, Ser. B):189–217, 2003. Computational semidefinite and second order cone programming: the state of the art.

[20] H. WOLKOWICZ, R. SAIGAL, and L. VANDENBERGHE, editors. *HANDBOOK OF SEMIDEFINITE PROGRAMMING: Theory, Algorithms, and Applications.* Kluwer Academic Publishers, Boston, MA, 2000. xxvi+654 pages.