

AN INTERIOR-POINT PERSPECTIVE ON
SENSITIVITY ANALYSIS IN LINEAR PROGRAMMING
AND SEMIDEFINITE PROGRAMMING

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Emre Alper Yıldırım

August 2001

© Emre Alper Yıldırım 2001

ALL RIGHTS RESERVED

AN INTERIOR-POINT PERSPECTIVE ON SENSITIVITY ANALYSIS IN
LINEAR PROGRAMMING AND SEMIDEFINITE PROGRAMMING

Emre Alper Yıldırım, Ph.D.

Cornell University 2001

This study is concerned with sensitivity analysis on perturbations of the right-hand side and the cost parameters in linear programming (LP) and semidefinite programming (SDP). Motivated by the desirable theoretical and practical properties of the interior-point methods, we investigate the possibility of performing sensitivity analysis relying entirely on interior-point methods and with very modest computational effort.

In Chapter 1, we discuss the advantages of the recent optimal partition approach over the traditional optimal basis approach to sensitivity analysis.

In Chapter 2, we start the theoretical development of our interior-point approach to sensitivity analysis. We seek tight bounds (interior-point bounds) on perturbations of the right-hand side and cost vectors which allow interior-point methods to recover a feasible and near-optimal solution in a single interior-point iteration from a near-optimal solution for the original problem. For LP, under a nondegeneracy assumption, we prove that the interior-point bounds evaluated at any near-optimal

strictly feasible iterates asymptotically coincide with the (symmetrized) optimal partition bounds. We also formulate explicit bounds for SDP using the Monteiro-Zhang (MZ) family of search directions.

We investigate the quality of the interior-point bounds for degenerate linear programs in Chapter 3. For a specific kind of degeneracy, we show that the asymptotic results of Chapter 2 continue to hold. For the remaining degeneracy situations, we show that the interior-point bounds still provide provably useful information in comparison with the symmetrized partition bounds.

Chapter 4 deals with the asymptotic behavior of the interior-point bounds for SDP. Under appropriate assumptions on nondegeneracy and on the convergence behavior of the central path, we prove that the interior-point bounds evaluated on the central path using the MZ family of search directions asymptotically coincide with the symmetrized optimal partition bounds. We extend these results to iterates lying in an appropriate (very narrow) central path neighborhood if the Nesterov-Todd direction is used to evaluate the interior-point bounds.

Computational results illustrating the behavior of the interior-point bounds on random LP and SDP instances are reported in Chapter 5.

Biographical Sketch

Emre Alper Yıldırım was born in 1975 in Erzurum, an eastern town of Turkey. After several relocations, his parents ended up settling in the Aegean port city of İzmir. In 1997, he earned his Bachelor of Science degree in Industrial Engineering from Bilkent University at Ankara. He then joined the Ph.D. program at Cornell University. During his studies, he was affiliated with the Cornell Computational Optimization Project. He spent the summer of 1999 as a Givens Research Associate at Argonne National Laboratory's Mathematics and Computer Science Division. He received a Master of Science degree and a Doctor of Philosophy degree in Operations Research from Cornell University in May 2000 and August 2001, respectively.

Starting from the fall of 2001, he will be working as an Assistant Professor at the Department of Applied Mathematics and Statistics at State University of New York at Stony Brook.

To my loving mother,
my sisters,
and
the memory of my father

Acknowledgements

I consider myself truly privileged and very fortunate to have worked under the supervision of Professor Michael J. Todd. He has always been available and ready to provide help, guidance, advice, and most importantly, constant support and encouragement especially during times of slow progress. His incredibly quick thinking, extensive knowledge, and always positive mood made him a great advisor. I would like to thank him for everything he has done for me.

Professor James Renegar has always been willing to help me even before I came to Cornell. I felt that I gained something useful every time I walked out of his office. I am grateful to him for his great support, his valuable feedback on my work, and his inspiring lectures that led me to the field of continuous optimization.

Professor Charles Van Loan was an extremely impressive lecturer. “Matrix Computations” was a lot of fun with him. I will always remember to “vectorize” my MATLAB codes. I thank him for serving on my committee.

I would also like to thank Steve Wright of Argonne National Laboratory for giving me the chance to spend a summer with him that led to a joint paper and for his continuing support ever since.

The School of Operations Research and Industrial Engineering at Cornell University provided an excellent atmosphere for a productive study. Among faculty members, I would especially like to thank Professors Gennady Samorodnitsky, David Shmoys, Narahari Prabhhu, Ann Trenk, and Sid Resnick for their help and encouragement.

Among the Ph.D. students, I am grateful to Kevin Wayne, Didi Chen, Pan Chen, Metin Çakanyıldırım, Mike Freimer, Michael Wagner, Vardges Melkonian, Tim Huh, Catalina Stefanescu, and Ganesh Janakiraman for their intimate friendship.

I thank Donna Moore, Kathy King, Jan Post, and Cindy Herter of OR & IE for efficiently handling all the administrative work and Bill Martin for his help with my computer problems.

I would also like to mention a few Turkish students at Cornell who went through a similar experience and helped make me feel at home. I extend my warmest thanks to Umut Çetin, Gizem Saka, Tuğkan Batu, İlker Tütüncü, Gökhan Hacısalihoğlu, Nil Elgün, Kıvanç Emiroğlu, and Ali Osman Öztürk.

During my studies at Cornell, I was supported by research assistantships funded by NSF through grant DMS-9805602 and ONR through grant N00014-96-1-0050. I gratefully acknowledge their support.

My father passed away during my first year at Cornell and did not see me earn my degree. I have always been proud to be his son.

Finally, I am indebted to my mother and my two sisters for their constant, unconditional love and emotional support.

Table of Contents

1	Introduction	1
1.1	Introduction and Brief History of Interior-Point Methods	1
1.2	Sensitivity Analysis	7
1.2.1	Optimal Basis Approach	8
1.2.2	Optimal Partition Approach	10
1.2.3	Comparison	11
1.2.4	Sensitivity in SDP	15
1.3	Motivation and Overview	17
2	Sensitivity Analysis in Linear Programming and Semidefinite Programming Using Interior-Point Methods*	22
2.1	Introduction	22
2.2	Linear Programming	26
2.2.1	Interior-Point Approach	27
2.2.2	Simplex Approach	33
2.2.3	Comparison of the Simplex and Interior-Point Approaches	35
2.3	Semidefinite Programming	44
2.3.1	General Results	49
2.3.2	The AHO Direction	57
2.3.3	The H..K..M Direction	62
2.3.4	The NT Direction	66
2.3.5	Comparison with LP	71
2.4	Discussion	74
3	An Interior-Point Approach to Sensitivity Analysis in Degenerate Linear Programs	76
3.1	Introduction	76
3.2	Preliminaries	79
3.2.1	Symmetrized Bounds	84
3.2.2	Interior-Point Approach and Central Path Neighborhoods	88

3.3	Equivalence	93
3.4	Unique Primal Solution	96
3.4.1	Perturbations of b	97
3.5	General Case	103
3.5.1	Perturbations of b	104
3.6	Conclusion	112
4	An Interior-Point Perspective on Sensitivity Analysis in Semidefinite Programming	114
4.1	Introduction	114
4.1.1	Notation	117
4.2	Semidefinite Programming	118
4.2.1	Sensitivity Analysis in SDP	127
4.2.2	Interior-Point Approach	130
4.3	Primal-Dual Symmetry	133
4.4	The Central Path	136
4.5	Central Path Neighborhood	156
4.6	Conclusion	162
5	Computational Experiments	164
5.1	Linear Programming	164
5.1.1	Generating Random LP Instances	165
5.1.2	Results	169
5.2	Semidefinite Programming	174
5.2.1	Generating Random SDP Instances	174
5.2.2	Results	179
5.3	Concluding Remarks	184
A	MATLAB Code for Generating a Random LP Instance	187
B	MATLAB Code for Generating a Random SDP Instance	192
	Bibliography	197

List of Tables

5.1	Computational Results for Linear Programming – Part 1 ($m = 200$, $n = 400$)	171
5.2	Computational Results for Linear Programming – Part 2 ($m = 200$, $n = 400$)	172
5.3	Computational Results for Linear Programming – Part 3 ($m = 200$, $n = 400$)	173
5.4	Computational Results for Semidefinite Programming – Perturbations of b ($m = 100$, $n = 20$)	180
5.5	Computational Results for Semidefinite Programming – Perturbations of C ($m = 100$, $n = 20$)	181

List of Figures

1.1	Feasible region of Example 1.2.1	13
1.2	Optimal value function of Example 1.2.1	14

Chapter 1

Introduction

1.1 Introduction and Brief History of Interior-Point Methods

Linear programming (LP) is the study of optimizing a linear functional (objective function) over the set of solutions to finitely many linear equalities and/or inequalities (constraints), which is called a polyhedron. Since its “first” formulation by Kantorovich in 1939 [33], LP has found applications in a wide variety of areas – including manufacturing, telecommunications, scheduling, finance, economics, and transportation.

Formally, an LP is formulated in its standard form as

$$\begin{aligned}
 \text{(P)} \quad & \min_x \quad c^T x \\
 & \text{subject to} \\
 & Ax = b, \\
 & x \geq 0,
 \end{aligned}$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$ constitute the data and $x \in \mathbb{R}^n$ is the decision variable. Here, \mathbb{R}^m and $\mathbb{R}^{m \times n}$ denote m -dimensional Euclidean space and the space of $m \times n$ real matrices, respectively. The matrix A is called the coefficient matrix. A closely related problem to (P) is given by

$$\begin{aligned}
 \text{(D)} \quad & \max_{y,s} \quad b^T y \\
 & \text{subject to} \\
 & A^T y + s = c, \\
 & s \geq 0,
 \end{aligned}$$

where $y \in \mathbb{R}^m$ and $s \in \mathbb{R}^n$ are the decision variables. Typically, (P) is called the *primal problem* and (D) the *associated dual problem*. There is a very nice duality theory for LP. The objective function value evaluated at any point satisfying the constraints of (P) – such a point is called a *feasible point* – provides an upper bound for the objective function value evaluated at any feasible point of (D). This is known as *weak duality*. Moreover, if the optimal solution is attained in (P), then (D) also has an optimal solution and the objective function values agree. This is the *strong duality theorem* for LP.

Inarguably, LP owes its significance to the existence of efficient and practical solution methods. The earliest efficient solution method was the simplex method

developed by Dantzig [8, 9]. Geometrically, the simplex method moves from one vertex of the feasible polyhedron to an adjacent one while improving on the objective function value. In practice, the simplex method has been very successful. On the theoretical side, Klee and Minty [38] gave instances of LPs where the simplex method had to go through exponentially many vertices.

The question of polynomial-time solvability of LP was answered affirmatively by Khachiyan's ellipsoid algorithm [35, 36] in 1979. However, the implementations of this algorithm were far away from fulfilling the expectations of a practical and efficient solution method.

Karmarkar's discovery of interior-point methods [34] for solving LPs in 1984 was probably the most significant achievement in mathematical programming since the development of the simplex method. In addition to the worst-case polynomial complexity, Karmarkar claimed that his algorithm enjoyed excellent practical performance on large LPs. Since then, interior-point methods have received intense consideration from researchers – turning into a new field itself. The field has become much more interesting after Nesterov and Nemirovski [47] unified the theory of interior-point methods in an elegant framework and extended it to general convex programming without sacrificing polynomial complexity. This seminal work made it possible to develop interior-point methods for a much larger class of optimization problems – including convex quadratic programming, convex programming, second-order cone programming and semidefinite programming.

In particular, semidefinite programming (SDP) has received considerable attention in the last decade. SDP is the study of optimizing a linear functional defined

on the space of symmetric matrices over the intersection of an affine set with the cone of symmetric and positive semidefinite matrices. Formally, SDP in its standard form is given by

$$(SP) \quad \min_X \quad C \bullet X$$

subject to

$$A_i \bullet X = b_i, \quad i = 1, \dots, m,$$

$$X \succeq 0,$$

where all $A_i \in \mathcal{S}^n$, $b \in \mathbb{R}^m$, $C \in \mathcal{S}^n$ are given, and $X \in \mathcal{S}^n$. Here, \mathcal{S}^n denotes the space of $n \times n$ symmetric matrices, and $X \succeq 0$ indicates that X is symmetric positive semidefinite. $P \bullet Q$ represents the trace inner product $\text{Trace}(P^T Q) = \sum_{ij} P_{ij} Q_{ij}$ on $n \times n$ matrices. The dual problem associated with (SP) is given by

$$(SD) \quad \max_{y,S} \quad b^T y$$

subject to

$$\sum_{i=1}^m y_i A_i + S = C,$$

$$S \succeq 0,$$

where $y \in \mathbb{R}^m$ and $S \in \mathcal{S}^n$. Due to the fundamental differences between the geometric structures of LP and SDP, the duality theory for SDP is not as strong as that for LP. The weak duality theorem carries over. However, strong duality might fail even when both (SP) and (SD) attain their optimal solutions. Under some regularity assumptions, however, SDP also enjoys strong duality.

SDP is a generalization of LP in the sense that n -dimensional Euclidean space is replaced by the space of $n \times n$ symmetric matrices and the nonnegative orthant by the cone of symmetric and positive semidefinite matrices. Furthermore, if all A_i

and C are restricted to be diagonal, then S is also diagonal and X can be replaced by its diagonal restriction. Therefore, we recover an LP, where A_i is the diagonal matrix whose entries on the diagonal are given by the i th row of A in (P), and C , X and S are the diagonal matrices obtained from c , x and s , respectively, in a similar manner. Consequently, LP is a special case of SDP.

SDP also enjoys applications in areas such as eigenvalue optimization, control theory, statistics, robust convex programming and approximation of combinatorial optimization problems. The reader is referred to [64] and to [56] for an extensive treatment of the theory and applications of SDP.

Unlike the simplex method, interior-point methods make progress by moving through the interior of the feasible region. The idea goes back to 1950s when Frisch [14] first suggested the use of a logarithmic barrier function. Fiacco and McCormick [13] developed barrier and penalty methods for nonlinear programming. The affine-scaling method was discovered and analyzed by Dikin [10, 11].

The survey paper by Todd [61] contains a more detailed treatment of the history of LP and the alternative solution techniques.

Currently, the most popular (and efficient) interior-point methods are the primal-dual path-following interior-point methods, the origins of which can be traced back to Renegar's primal path-following algorithm [53] that introduced the use of Newton's method. In fact, Renegar's algorithm achieves still the best known complexity bound. Path-following methods are motivated by perturbing the objective function of (P) by adding a logarithmic barrier function to ensure that the iterates stay away from the boundary of the feasible region – hence the term interior-point. More

precisely, they are based on the barrier problem

$$\begin{aligned}
 \text{(BP)} \quad \min_x \quad & c^T x - \mu \sum_{i=1}^n \ln x_i \\
 \text{subject to} \quad & \\
 & Ax = b, \\
 & (x > 0),
 \end{aligned}$$

where μ is a positive parameter. Assuming that (BP) has a feasible solution, the optimality conditions are given by

$$\begin{aligned}
 A^T y + s &= c, \\
 Ax &= b, \\
 XSe &= \mu e,
 \end{aligned} \tag{1.1}$$

with $x > 0$ and $s > 0$. Here, X and S denote the diagonal matrices whose entries are given by x and s , respectively, and e denotes the vector of ones. As $\mu \downarrow 0$, the nonlinear system (1.1) precisely converges to the optimality conditions for (P) and (D). The set of solutions $(x(\mu), y(\mu), s(\mu))$ of (1.1) is called the central path. Under some regularity assumptions, Adler and Monteiro [1] show that the system (1.1) indeed defines a continuous and differentiable path converging to the so-called analytic center of the primal-dual optimal solution set as $\mu \downarrow 0$. Roughly, interior-point methods work as follows: One uses Newton's method to approximately solve (1.1), decreases μ towards 0 appropriately, and iterates. Ironically, transforming a linear problem into a nonlinear one leads to a polynomial-time algorithm. A good reference for primal-dual path-following interior-point methods is the book by Wright [67]. The book by Ye [68] is another good reference for the interior-point

theory. The general theory of interior-point methods for convex optimization is treated in a form that is rather easy to digest in the book by Renegar [54].

Interior-point methods for SDP mimic those for LP. Certainly, the difference in the problem structures of LP and SDP necessitates the definition of an appropriate logarithmic barrier function. It turns out that the function $f : \mathcal{S}^n \rightarrow \mathbb{R}$ defined as $f(X) = -\ln \det(X)$ possesses the necessary properties.

The field of interior-point methods changes quite rapidly. Fortunately, there are excellent web sites that maintain links to the most recent technical reports. These web sites are extremely helpful to keep abreast of the developments in the area. Interior-Point Methods Online [66] is maintained by Wright and contains an archive of papers on interior-point methods and relevant links. A similar web site that maintains more general papers on optimization is Optimization Online [23], developed by Goux. For SDP, the web pages run by Alizadeh [3] and Helmberg [28] are invaluable resources.

1.2 Sensitivity Analysis

Sensitivity analysis (or post-optimality analysis) is the study of the behavior of the optimal solution of an optimization problem with respect to changes in the data of the original problem. Apart from being an interesting theoretical question in itself, sensitivity analysis is often very important in practice. Typically, the numerical data arising from real-life problems might not be precise due to measurement errors or inadequate estimates. Consequently, one might be interested in obtaining optimal

solutions under different scenarios resulting from small perturbations of the data. Even if the data is precise, one might be concerned with the behavior of the optimal solution with respect to specific perturbations. This might arise, for instance, in the context of capacity expansion. A company might want to determine the most critical resource whose per-unit increase would yield the most significant increase in production or profit.

In this dissertation, we study sensitivity analysis in LP and SDP and we restrict ourselves to perturbations in the right-hand side and cost parameters.

Next, we describe two different ways to perform sensitivity analysis in LP. We only discuss perturbations in the right-hand side vector in (P) due to the fact that perturbations in the cost vector is equivalent to perturbations of the right-hand side vector in (D).

1.2.1 Optimal Basis Approach

In (P), the coefficient matrix A can be assumed to have full row rank without loss of generality. Otherwise, a row reduction will either lead to an infeasible problem (i.e., a problem with no feasible points) or to the identification of redundant rows which can be eliminated from the problem. After a possible rearrangement of the columns of A , let $A = [B \ N]$, where $B \in \mathbb{R}^{m \times m}$ is nonsingular. The solution

$$x = \begin{bmatrix} x_B \\ x_N \end{bmatrix} = \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix} \quad (1.2)$$

is called a basic solution. If, furthermore, $x_B \geq 0$, then x is called a basic feasible solution. Here B is called the basic matrix (or simply the basis). The components

of x_B are called basic variables and those of x_N are called nonbasic variables. If $x_B > 0$, then x is called a nondegenerate basic feasible solution and if at least one component of x_B is 0, then x is called a degenerate basic feasible solution.

It is well-known that if (P) has an optimal solution, then it has a basic optimal solution. The simplex method, in fact, terminates with such an optimal solution (if any). At termination, the (primal) simplex method also produces an optimal solution for (D) as a byproduct. More precisely, if x is the resulting basic optimal solution given by (1.2), then $y = B^{-T}c_B$, $s_B = 0$, $s_N = c_N - N^T B^{-T}c_B \geq 0$, where the subscripts denote the appropriate partitions with respect to the columns of B and N , is a basic optimal solution for (D). The matrix B satisfying these properties simultaneously (primal and dual optimality) is called the optimal basic matrix or simply the optimal basis. Then one way to perform sensitivity analysis on perturbations of the right-hand side vector is to determine the range of perturbations for which the optimal basis remains unchanged for the resulting set of problems. More precisely, if b is replaced by $b + t\Delta b$, where $t \in \mathbb{R}$ and $\Delta b \in \mathbb{R}^m$, then the range of t is given by the set of values for which

$$B^{-1}(b + t\Delta b) = x_B + tB^{-1}\Delta b \geq 0. \quad (1.3)$$

We will call the bounds arising from this approach the optimal basis bounds (see e.g. [9]).

1.2.2 Optimal Partition Approach

Feasible points x and (y, s) are respectively optimal for (P) and (D) if they satisfy the complementary slackness property, i.e.,

$$Ax = b, \quad A^T y + s = c, \quad x_i s_i = 0, \quad i = 1, \dots, n, \quad (1.4)$$

where x_i and s_i denote the i th components of x and s , respectively. Let us denote the set of optimal solutions for (P) and (D) by Ω_P and Ω_D , respectively. We define two index sets as

$$\begin{aligned} \mathcal{B} &= \{j \in \{1, \dots, n\} : x_j > 0 \text{ for some } x \in \Omega_P\}, \\ \mathcal{N} &= \{j \in \{1, \dots, n\} : s_j > 0 \text{ for some } (y, s) \in \Omega_D\}. \end{aligned} \quad (1.5)$$

It is easy to see that $\mathcal{B} \cap \mathcal{N} = \emptyset$ by the complementary slackness property. Moreover, Goldman and Tucker [19] has shown that $\mathcal{B} \cup \mathcal{N} = \{1, \dots, n\}$, i.e., \mathcal{B} and \mathcal{N} actually partition the index set $\{1, \dots, n\}$. Therefore, $(\mathcal{B}, \mathcal{N})$ is called the optimal partition.

The optimal partition approach to sensitivity analysis on perturbations of the right-hand side vector [2, 32] can then be defined as determining the range of perturbations for which the optimal partition remains unchanged for the resulting set of problems. If b is replaced by $b + t\Delta b$, then the range of t is given by the optimal values of the two LPs given by

$$\begin{aligned} \text{(AUX)} \quad & \min_{x_B, t} \quad (\max_{x_B, t}) \quad t \\ & \text{subject to} \\ & Bx_B = b + t\Delta b, \\ & x_B \geq 0, \end{aligned}$$

where B is the submatrix of A consisting only of those columns corresponding to the indices in \mathcal{B} . In contrast to the case for an optimal basis, B need not be a square matrix by the preceding arguments.

We will call the bounds arising from this approach the optimal partition bounds.

1.2.3 Comparison

There are various differences between the concepts of an optimal basis and the optimal partition. First, an LP might have multiple basic optimal solutions and each such solution has a different optimal basis. However, the optimal partition is unique for a given LP instance. Secondly, if an optimal basis is regarded as a partition of indices of the optimal solution x into two sets corresponding to those in x_B and those in x_N , then the cardinalities of the two sets will always be m and $n - m$, respectively. On the other hand, the cardinalities of \mathcal{B} and \mathcal{N} as in (1.5) can be anywhere between 0 and n .

As we mentioned before, the (primal) simplex method also produces an optimal solution (y, s) given by $y = B^{-T}c_B$, $s_B = 0$, $s_N = c_N - N^T B^{-T}c_B \geq 0$, for (D) as a byproduct. Note that perturbations of the right-hand side vector does not affect the constraints of (D). The optimal basis approach can alternatively be defined as finding the range of perturbations for which (y, s) as above remains basic optimal for the resulting set of dual problems. The optimal partition approach, on the other hand, aims to determine the range of perturbations of b for which the whole dual optimal set as opposed to a single optimal solution remains invariant. Consequently, generally speaking, the optimal partition approach takes into account much more

information than the optimal basis approach. Clearly, if there is a unique optimal solution for (P) and (D), then the two approaches coincide.

Despite the fact that the optimal partition approach uses more information about (P) and (D), its computational cost generally outweighs that of the optimal basis approach since two LPs need to be solved. The optimal basis bound is simply given by the single vector inequality (1.3).

The optimal partition has close connections with interior-point methods. First of all, note that any optimal solution (x, y, s) in the relative interior of the primal-dual optimal set will suffice to determine the indices in \mathcal{B} and \mathcal{N} . This is due to the fact that any such solution can be written as a strict convex combination of all the basic optimal solutions. Furthermore, primal-dual interior-point methods converge to an optimal solution in the relative interior of the primal-dual optimal set, thereby readily identifying the two index sets \mathcal{B} and \mathcal{N} . The simplex method, however, has to go through all the alternate basic optimal solutions before it can identify the indices in \mathcal{B} and \mathcal{N} accurately.

In spite of its computational advantage over the optimal partition approach, the optimal basis approach might yield misleading and inconsistent information on sensitivity analysis if there are multiple basic optimal solutions. Jansen, de Jong, Roos and Terlaky [32] present and solve LP instances using different solvers that provide information about post-optimality analysis and report strikingly different results for the same LP instance. The following small example will be useful to illustrate what might go wrong with the optimal basis approach if there are multiple basic optimal solutions.

Figure 1.1: Feasible region of Example 1.2.1

Figure 1.2: Optimal value function of Example 1.2.1

Example 1.2.1 Consider the following LP instance with $m = 1$ and $n = 2$:

$$\min x_1 + x_2 \quad \text{subject to} \quad x_1 + x_2 = 1, \quad x \geq 0.$$

Since this is an LP in two variables, we can draw the graph of the feasible region (see Figure 1.1). Clearly, any feasible solution (indicated by the solid line in Figure 1.1) is optimal since the objective vector $c = [1, 1]^T$ is orthogonal to the affine space defined by the linear equality.

Assume that the cost coefficient of x_1 is replaced by $1+t$. Figure 1.1 includes the affect of such a change on the objective vector c . Let us consider the optimal basis bound. As illustrated in Figure 1.1, there are two basic optimal solutions given by $x^1 := [0, 1]^T$ and $x^2 := [1, 0]^T$. If the simplex method terminates at x^1 , then the optimal basis bound is $[0, +\infty)$ since x^1 will remain optimal for all values of t in that range. A similar argument reveals that the optimal basis bound is given by $(-\infty, 0]$ if x^2 is the final optimal solution. \square

As Example 1.2.1 illustrates, the optimal basis approach might yield totally inconsistent information on sensitivity analysis if there are multiple optimal solutions. Let us take a closer look at Example 1.2.1. Figure 1.2 depicts the optimal value as a function of t , which we denote by $v(t)$. The original problem instance as given by Example 1.2.1 corresponds to $t = 0$. The graph illustrates that $t = 0$ is actually a breakpoint of the optimal value function. $v(t)$ remains constant at value 1 for $t > 0$,

whereas it decreases to $-\infty$ as t decreases from 0 to $-\infty$. In fact, the optimal partition for the original instance ($t = 0$) is given by $\mathcal{B} = \{1, 2\}$ and $\mathcal{N} = \emptyset$ since $x = [1/2, 1/2]^T$, for instance, is an optimal solution. However, if $t > 0$, then x^1 is the unique optimal solution, in which case $\mathcal{B} = \{2\}$ and $\mathcal{N} = \{1\}$, and if $t < 0$, then x^2 is the unique optimal solution with $\mathcal{B} = \{1\}$ and $\mathcal{N} = \{2\}$. Therefore, the optimal partition changes as soon as t moves from 0 in either direction. Consequently, as shown by Adler and Monteiro [2] and Jansen, de Jong, Roos and Terlaky [32], the breakpoints of the optimal value function are precisely those points where the optimal partition changes. In Example 1.2.1, the optimal partition bound is therefore $\{0\}$.

We conclude then that the optimal partition bounds uniquely and unambiguously identify the breakpoints of the optimal value function. Despite its computational disadvantage, the optimal partition approach is a more reliable way of performing sensitivity analysis than the optimal basis approach. As we mentioned before, the two approaches coincide only if both (P) and (D) have a unique optimal solution.

1.2.4 Sensitivity in SDP

The feasible regions defined by the constraints of LP and SDP can geometrically be quite different. The feasible region of an LP is always a polyhedron whereas an SDP feasible region is often nonpolyhedral. In fact, the representation of the 2×2 symmetric positive semidefinite matrices in 3-dimensional space is a right circular cone. There have been some attempts to generalize the concept of basic solution

from LP to a larger class of optimization problems (see Anderson and Nash [6] for infinite-dimensional LPs over general cones and Pataki [52] for SDP). Yet, an optimal solution for SDP does not typically resemble a basic feasible solution for LP. Consequently, the optimal basis approach to sensitivity analysis in LP cannot directly and practically be generalized to perform sensitivity analysis in SDP.

The optimal partition approach, on the other hand, can be extended to SDP as shown by Goldfarb and Scheinberg [18]. In order to do this, one needs an appropriate definition of the optimal partition in the context of SDP. Let (X, y, S) be a primal-dual optimal solution for (SP) and (SD) such that X and (y, S) are in the respective relative interiors of the primal and dual optimal sets. Denote the range spaces of X and S by \mathfrak{R}_P and \mathfrak{R}_D , respectively. It turns out that the range space of any other primal optimal solution \tilde{X} is a subset of \mathfrak{R}_P [7]. A similar relation holds between the range space of the S -component of any other dual optimal solution (\tilde{y}, \tilde{S}) and \mathfrak{R}_D . The optimality conditions for SDP (under some regularity assumptions) imply that $XS = 0$. Consequently, \mathfrak{R}_P and \mathfrak{R}_D are mutually orthogonal subspaces. Let us define \mathfrak{R}_N to be the orthogonal complement of $\mathfrak{R}_P \oplus \mathfrak{R}_D$. Therefore, \mathbb{R}^n can be decomposed into three mutually orthogonal subspaces given by $(\mathfrak{R}_P, \mathfrak{R}_N, \mathfrak{R}_D)$, which is called the optimal partition.

Similarly to the LP case, one way to perform sensitivity analysis on perturbations of the right-hand side vector is to determine the range of perturbations for which the optimal partition remains unchanged for the resulting set of problems.

Let $Q = [Q_P \ Q_N \ Q_D]$ be an orthogonal matrix such that the columns of each component of Q respectively form a basis for the optimal partition $(\mathfrak{R}_P, \mathfrak{R}_N, \mathfrak{R}_D)$.

If the right-hand side vector b is replaced by $b + t\Delta b$, where $t \in \mathbb{R}$ and $\Delta b \in \mathbb{R}^m$, then the range of t is given by the optimal values of the two SDPs

$$\begin{aligned}
 \text{(AUX-SDP)} \quad & \min_{t,U} (\max_{t,U}) && t \\
 & A_i \bullet Q_P U Q_P^T &= & b + t\Delta b_i, \quad i = 1, \dots, m, \\
 & U &\succeq & 0,
 \end{aligned}$$

where the size of U is determined by the dimension of \mathfrak{R}_P . Goldfarb and Scheinberg [18] prove that the optimal values of (AUX-SDP) do indeed give the range of t for which the optimal partition remains invariant for the resulting set of problems.

With some abuse of language, we will also call the bounds arising from this approach the optimal partition bounds. Whether the term refers to LP or SDP will always be clear from the context.

1.3 Motivation and Overview

Interior-point methods possess very attractive theoretical properties such as polynomial-time worst-case complexity. Furthermore, in practice, the algorithms tend to perform much better than their predicted worst-case behavior. Although the competition for a better algorithm between the simplex method and interior-point methods is not yet over, the latter tends to be the algorithm of choice especially for large-scale LPs and problems having certain structures. In addition, the work of Nesterov and Nemirovski [47] extends the interior-point theory to a much larger class of optimization problems and most of the desirable theoretical properties carry over.

Despite their generality and favorable theoretical and practical properties, interior-point methods have been considered unsuitable for the purposes of sensitivity analysis.

The fact that interior-point methods converge to the relative interior of the primal-dual optimal set in the presence of multiple optimal solutions has been viewed as a drawback. This perception is presumably a consequence of the way sensitivity analysis has been traditionally done, i.e., the optimal basis approach outlined in Section 1.2.1. In spite of its shortcomings as shown in [32] or by Example 1.2.1, the optimal basis approach, for which the simplex method is perfectly suited, is still explained as the only way to perform sensitivity analysis in linear programming textbooks. In the context of sensitivity analysis, possible advantages of using optimal solutions in the relative interior of the optimal set over a basic optimal solution have only recently been investigated [24]. The optimal partition approach [2, 32] has been shown to yield consistent and unambiguous information on sensitivity analysis. Moreover, the close relation between the optimal partition and interior-point methods points to possible advantages over the use of the simplex method.

Another possible reason for the aforementioned perception arises from the way interior-point methods work. As their name suggests, such methods generate iterates which lie in the relative interior of the underlying cone (positive orthant for LP and cone of symmetric positive definite matrices for SDP). In most cases, however, an optimal solution lies on the boundary of the cone. Consequently, in contrast to the situation with the simplex method, an optimal solution generally cannot directly

be used in the context of interior-point methods.

Some researchers suggested ways of recovering a basic optimal solution from a nonbasic optimal solution (such as one in the relative interior of the optimal set) so that the optimal basis approach can be used for the purposes of sensitivity analysis [65]. However, such a procedure requires extra computational effort and the resulting basic optimal solution might still suffer from the shortcomings of the optimal basis approach.

On the other hand, the optimal partition approach involves solving two LPs, each of which can be as large as the original LP. Therefore, the computational cost associated with it could be prohibitively high.

The motivation for this dissertation originated from investigating the possibility of performing sensitivity analysis relying entirely on interior-point methods and with very modest computational effort. In other words, assuming that an LP is solved using an interior-point method, we aimed to retrieve information about sensitivity analysis from the iterates resulting from the algorithm without ever switching to a basic optimal solution and using only techniques arising from the interior-point theory. Clearly, an affirmative answer to such a question would enhance the usefulness of interior-point methods.

However, the nature of the interior-point methods as described above necessitates the proper formulation of how to conduct sensitivity analysis using interior-point ideas. In this dissertation, we have formulated a proper approach, henceforth the interior-point approach. Briefly, assuming that an interior-point method is terminated at a near-optimal solution for the original LP, our approach is motivated

by determining the range of perturbations for which a single interior-point iteration at this near-optimal solution would recover a feasible and near-optimal solution for the perturbed problem. This question alone requires further formulation of an appropriate interior-point iteration among many other choices. In Chapter 2, we propose an interior-point iteration for LP and SDP. We first analyze the LP case. For perturbations of the right-hand side and the cost vectors, we obtain tight bounds (interior-point bounds) on the perturbations that allow interior-point methods to recover a feasible and near-optimal solution in a single interior-point iteration. We then study the asymptotic behavior of these interior-point bounds. Under a nondegeneracy assumption, we show that the interior-point bounds asymptotically coincide with the symmetrized version of the optimal basis (or equivalently, the optimal partition) bounds. We then generalize our interior-point perspective to SDP using the Monteiro-Zhang family of search directions and specialize the bounds to the AHO [5], H..K..M [29, 39, 42] and NT [48, 49] directions.

In Chapter 3, we study the asymptotic behavior of the interior-point bounds for degenerate LPs in comparison with the optimal partition bounds for perturbations of the right-hand side and the cost vectors. We prove that the same asymptotic coincidence as in Chapter 2 continues to hold for a specific class of degenerate LPs. For general degenerate LPs, we can still prove a weaker asymptotic relationship between the two bounds.

We consider the interior-point bounds for SDP in Chapter 4. We study the asymptotic behavior of the interior-point bounds for SDP in comparison with the optimal partition bounds for perturbations of the right-hand side vector and the cost

matrix. Under appropriate nondegeneracy assumptions, we prove that our interior-point bounds using the Monteiro-Zhang family of search directions evaluated on the central path asymptotically coincide with the symmetrized optimal partition bounds. We then extend the same asymptotic result to a very narrow central path neighborhood using the Nesterov-Todd direction.

Chapters 2, 3, and 4 are based on the papers [71, 70, 69], each of which is self-contained with its own introduction and conclusion.

Finally, we report some computational results in Chapter 5 to shed some light on the behavior of the interior-point bounds in practice. We apply our interior-point bounds for perturbations of the right-hand side and cost parameters on randomly generated LP and SDP instances. The computational results are in agreement with the theoretical findings of Chapters 2, 3 and 4. Some concluding remarks are also included in this chapter.

Chapter 2

Sensitivity Analysis in Linear Programming and Semidefinite Programming Using Interior-Point Methods*

2.1 Introduction

This paper is concerned with sensitivity analysis for linear programming (LP) and semidefinite programming (SDP) problems using interior-point methods. Sensitivity analysis (also called post-optimality analysis) is the study of the behavior of the optimal solution with respect to changes in the input parameters of the original

*Joint paper with Michael J. Todd. Appeared in *Mathematical Programming* 90 (2), pp. 229–261 (2001). Published by Springer-Verlag.

optimization problem. It is often as important as solving the original problem itself, partly because in real life applications, the parameters are not always precise and are subject to some source of error.

For the LP case, sensitivity analysis based on the optimal basis matrix has been well-studied. Recently, an interior-point method approach using the analytic central optimal solution as opposed to an optimal basic solution has been analyzed by several researchers. Greenberg [24], Jansen, de Jong, Roos and Terlaky [32] and S. Zhang [73] discuss the advantages of the central optimal solution over a basic solution. Adler and Monteiro [2] show that it is possible to perform parametric analysis using the optimal partition (i.e., for each index, knowing whether the corresponding component of an optimal primal solution or of an optimal dual slack vector can be positive). Roos, Terlaky, and Vial [55] develop a parametric analysis of the optimal value from the central optimal solution perspective. Nunez and Freund [50] and Holder, Sturm and S. Zhang [31] study the behavior of the central path under perturbations of the input data.

For the SDP case, Goldfarb and Scheinberg [18] investigate the properties of the optimal value function under perturbations of the input parameters. Sturm and S. Zhang [59] study the properties of the central path with respect to perturbations of the right-hand side vector.

Our study in this paper is different from the above studies in the sense that it is motivated by asking how the interior-point method from a near-optimal pair of strictly feasible solutions for a problem and its dual compares with the results obtained from a nondegenerate optimal basic solution under perturbations of the

right-hand side and the cost parameters for the LP case. We focus on obtaining explicit bounds on the perturbations of the input parameters so that a single iteration of the interior-point method (with very modest cost) regains feasibility for the perturbed problem and its dual. Further, the new iterates have duality gap smaller than that of the original iterates. We show that under the unique, nondegenerate solution assumption, the interior-point approach yields asymptotically exactly the same bounds as those that keep the current basis optimal (after symmetrization with respect to the origin); since these are the bounds natural when using the simplex method, we call these the bounds from the simplex approach. We also extend our analysis to the SDP case and obtain bounds on perturbations of the right-hand side and the cost parameters using the AHO [5], H..K..M [29, 39, 42] and NT [48, 49] directions.

Let us note that the question of using a small number of interior-point iterations to regain feasibility when the problem data change also arises in cutting-plane methods for convex feasibility problems (see, e.g., Goffin, Haurie, and Vial [15] and Goffin and S.-Mokhtarian [16] and the references therein). However, in our case the dimensions of the problems do not change, we apply the iterations from a near-optimal pair of points rather than an analytic center, and we explicitly limit ourselves to a single iteration rather than a small number.

We stress that the bounds we obtain are valid in the presence of degeneracy, which appears in most practical LP models; it is only the comparison with the simplex approach that makes nondegeneracy assumptions. We give an example to show the difficulties when there is degeneracy; however, a follow-up paper (Chap-

ter 3) will show that even in this case our bounds achieve a certain fraction of some natural bounds that depend only on the problem, not on an algorithmic approach.

After this paper was written (and revised), we became aware of a related paper by Kim, Park, and Park [37], henceforth KPP. The authors also consider changes in the right-hand side or the cost parameters and investigate when a single interior-point-like step from a near-optimal pair of strictly feasible solutions for a problem and its dual can regain feasibility and maintain near-optimality. However, KPP only change either the primal or dual solution: if the right-hand side (cost vector) changes, they change only the primal (dual) solution. Their step cannot be motivated by a slight change in the usual Newton step in a primal-dual interior-point iteration, but their change to the primal (dual) solution coincides with ours. KPP show that, in the nondegenerate case, the condition on the change in the data that allows feasibility to be regained is asymptotically exactly that keeping the optimal basis the same. However, to show that the new pair of solutions remains near-optimal requires another condition, which they show holds asymptotically; but it may be the case that the duality gap of the new pair exceeds that of the original pair by a considerable amount. This contrasts with our result, which requires a more stringent condition (the symmetrization of KPP's) to assure feasibility, but which then guarantees a reduction in the duality gap. We also believe that our analysis of the asymptotic behavior of the projection matrices is more complete than theirs.

Our paper is organized as follows. In the next section, we investigate the LP case. We present bounds on perturbations of the right-hand side and the cost vectors using the interior-point approach and the simplex approach and then compare the

bounds resulting from the two approaches. The analysis of perturbations of the right-hand side vector and the cost matrix for the SDP case in the general form as well as using the three specific search directions is given in Section 3. We conclude the paper with a discussion in Section 4.

2.2 Linear Programming

We consider the LP given in the following standard form:

$$\begin{aligned}
 (LPP) \quad & \min_x \quad c^T x \\
 & \text{s.t.} \\
 & Ax = b, \\
 & x \geq 0,
 \end{aligned}$$

where c and $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and $A \in \mathbb{R}^{m \times n}$. Throughout this section, the coefficient matrix A will be fixed; thus we parametrize the above LP by b and c , and we denote it by $LPP(b, c)$. The associated dual LP is given by the following:

$$\begin{aligned}
 (LPD) \quad & \max_{y,s} \quad b^T y \\
 & \text{s.t.} \\
 & A^T y + s = c, \\
 & s \geq 0,
 \end{aligned}$$

where $y \in \mathbb{R}^m$ and $s \in \mathbb{R}^n$. Similarly, the dual LP will be denoted by $LPD(b, c)$. Without loss of generality, we assume that A has full row rank.

We say the triple (x, y, s) is a (strictly) feasible point for $LPP(b, c)$ and $LPD(b, c)$ if x and (y, s) are (strictly) feasible for these two problems respectively. (Here a

feasible solution is called strictly feasible if all inequalities are satisfied strictly.)

2.2.1 Interior-Point Approach

We assume that there exists a strictly feasible point (x, y, s) for $LPP(b, c)$ and $LPD(b, c)$. It is well known that the duality gap corresponding to such a point is given by $c^T x - b^T y = x^T s > 0$. X and S will denote the diagonal matrices corresponding to x and s , respectively, and e will denote the vector of ones in the appropriate dimension.

First, we will briefly review the concept of the central path in LP. The central path is a path of strictly feasible points $(x(\mu), y(\mu), s(\mu))$ parametrized by a positive scalar μ . Each point on the central path satisfies the following system for some $\mu > 0$:

$$\begin{aligned} A^T y + s &= c, \\ Ax &= b, \\ XSe &= \mu e, \end{aligned} \tag{2.1}$$

with $x > 0$ and $s > 0$. Under the assumption above, such a solution exists for each positive μ .

An interior-point iteration is usually a Newton step for this nonlinear system of equations for some μ , possibly with different right-hand sides. Suppose (x, y, s) is the current iterate, and we seek an approximation to the point on the central path corresponding to parameter μ (say equal to $\gamma x^T s/n$). Then the Newton step

$(\Delta x, \Delta y, \Delta s)$ is given by the solution of the following system:

$$\begin{aligned} A^T \Delta y + \Delta s &= r_d, \\ A \Delta x &= r_p, \\ S \Delta x + X \Delta s &= r_{xs}, \end{aligned} \tag{2.2}$$

where $r_p = b - Ax$, $r_d = c - A^T y - s$, and $r_{xs} = \mu e - XSe$. Here, r_p , r_d and r_{xs} are simply the primal, dual and complementary slackness residuals, respectively.

However, we might want to use different right-hand sides. If the right-hand side b or cost vector c is changed to b' or c' , we may wish to use this instead of b or c to compute r_p or r_d . Similarly, we may want to strive for a different product of the primal and dual variables than μe , as in target-following methods. We will say that the Newton step from (x, y, s) *targeting* the feasible point (x', y', s') of $LPP(b', c')$ and $LPD(b', c')$ that satisfies $X'S'e = v$ is the triple $(\Delta x, \Delta y, \Delta s)$ solving (2.2) for $r_p = b' - Ax$, $r_d = c' - A^T y - s$, and $r_{xs} = v - XSe$. (This is a slight abuse of language, since such a point might not exist, but the Newton step is still defined.)

If A has full row rank, then the system (2.2) has a unique solution given by:

$$\begin{aligned} \Delta y &= (AD^2A^T)^{-1}(r_p + AD^2r_d - AS^{-1}r_{xs}), \\ \Delta s &= r_d - A^T \Delta y, \\ \Delta x &= S^{-1}(r_{xs} - X \Delta s), \end{aligned} \tag{2.3}$$

where $D = S^{-\frac{1}{2}}X^{\frac{1}{2}}$. The key observation here is that if A has full row rank, then AD^2A^T will be symmetric positive definite, and hence invertible.

To avoid extra computation, we note that the results below need not be applied to the final iterate generated by a primal-dual interior-point method. If we backtrack

to the previous iterate, a factorization of the matrix AD^2A^T necessary to compute the Newton step will already be available. Then an iteration simply reduces to solving two triangular systems followed by a few matrix-vector products. Hence in practice we may choose to let (x, y, s) be the penultimate iterate of the method used to solve the original problems.

Next, we present our results about perturbations of b and c .

Proposition 2.2.1 *Assume that (x, y, s) is a strictly feasible point for $LPP(b, c)$ and $LPD(b, c)$ and the right-hand side vector b is replaced by $b' := b + \Delta b$, where $\Delta b \in \mathbb{R}^m$. Suppose a Newton step is taken from (x, y, s) targeting the feasible point (x', y', s') of $LPP(b', c)$ and $LPD(b', c)$ that satisfies $X'S'e = XSe$. If, and only if,*

$$\|S^{-1}A^T(AD^2A^T)^{-1}\Delta b\|_\infty \leq 1, \quad (2.4)$$

where $D = X^{\frac{1}{2}}S^{-\frac{1}{2}}$, then a full Newton step can be taken and the resulting iterate will be feasible for the new problems. Moreover, in this case the new iterate will have duality gap at most $x^T s$.

Proof: Using the above notation in (2.2) and by the hypothesis, we find $r_p = \Delta b$, $r_d = 0$, and $r_{xs} = 0$. Let's consider the third equation in (2.2):

$$S\Delta x + X\Delta s = 0. \quad (2.5)$$

Rewriting this equality component-wise, we have:

$$s_i\Delta x_i + x_i\Delta s_i = 0 \quad \text{so} \quad \frac{\Delta x_i}{x_i} + \frac{\Delta s_i}{s_i} = 0, \quad i = 1, \dots, n, \quad (2.6)$$

where x_i denotes the i th component of x . However, the next iterate will be feasible iff $x_i + \Delta x_i \geq 0$ and $s_i + \Delta s_i \geq 0$, $i = 1, \dots, n$, since the equality constraints will automatically be satisfied if a full Newton step is taken. Combining these inequalities with (2.6), we have that the new iterate will satisfy nonnegativity for both x and s if and only if $|\frac{\Delta s_i}{s_i}| \leq 1$, $i = 1, \dots, n$. Thus, it is necessary and sufficient to have $\|S^{-1}\Delta s\|_\infty \leq 1$. But using (2.3), we have:

$$\Delta y = (AD^2A^T)^{-1}\Delta b \quad \text{and} \quad \Delta s = -A^T\Delta y. \quad (2.7)$$

Hence

$$\|S^{-1}\Delta s\|_\infty = \|S^{-1}A^T(AD^2A^T)^{-1}\Delta b\|_\infty,$$

and this proves the first part of the proposition.

The duality gap of the new iterate will be given by:

$$(x + \Delta x)^T(s + \Delta s) = x^T s + x^T \Delta s + s^T \Delta x + \Delta x^T \Delta s. \quad (2.8)$$

Multiplying (2.5) by e^T from the left, we have $x^T \Delta s + s^T \Delta x = 0$. From (2.6), Δx_i and Δs_i have opposite signs, and so $\Delta x^T \Delta s \leq 0$. Thus, we have:

$$(x + \Delta x)^T(s + \Delta s) \leq x^T s \quad (2.9)$$

as claimed. □

We note that the simple bound

$$\|\Delta b\|_\infty \leq \frac{1}{\|S^{-1}A^T(AD^2A^T)^{-1}\|_\infty} \quad (2.10)$$

implies that Δb satisfies the condition (2.4); moreover (2.10) defines the largest L_∞ -box around the origin guaranteeing this condition. The proof is straightforward. A similar statement holds for the next result on perturbations of the cost vector c .

Proposition 2.2.2 *Assume that (x, y, s) is a strictly feasible point for $LPP(b, c)$ and $LPD(b, c)$ and the cost vector c is replaced by $c' := c + \Delta c$, where $\Delta c \in \mathbb{R}^n$. Suppose a Newton step is taken from (x, y, s) targeting the feasible point (x', y', s') of $LPP(b, c')$ and $LPD(b, c')$ that satisfies $X'S'e = XSe$. If, and only if,*

$$\|S^{-1}(I - A^T(AD^2A^T)^{-1}AD^2)\Delta c\|_\infty \leq 1, \quad (2.11)$$

where $D = X^{\frac{1}{2}}S^{-\frac{1}{2}}$, then a full Newton step can be taken and the resulting iterate will be feasible for the new problems. Moreover, in this case the new iterate will have duality gap at most $x^T s$.

Proof: Once again using (2.2), we have $r_p = 0$, $r_d = \Delta c$ and $r_{xs} = 0$. By the argument used in the proof of Proposition 2.2.1, it is necessary and sufficient that $\|S^{-1}\Delta s\|_\infty \leq 1$. Note that (2.3) implies

$$\Delta y = (AD^2A^T)^{-1}AD^2\Delta c \quad \text{and} \quad \Delta s = (I - A^T(AD^2A^T)^{-1}AD^2)\Delta c.$$

Therefore,

$$\|S^{-1}\Delta s\|_\infty = \|S^{-1}(I - A^T(AD^2A^T)^{-1}AD^2)\Delta c\|_\infty,$$

and this proves the first part of the proposition. Essentially the same arguments as in the previous proposition hold to prove the decrease in the duality gap. \square

The proposed Newton system to regain feasibility for the new problems uses $r_{xs} = 0$ in (2.2). This choice can be motivated in the following way. If (x, y, s) is

near-optimal to start with, the pairwise products $x_i s_i$ then are very small. Therefore, by targeting a point for the new problems with the same pairwise products, we hope to be able to maintain near-optimality while regaining feasibility. For reoptimization after a data perturbation, the simplex method always maintains complementarity ($x_i s_i = 0$) while working towards primal or dual feasibility. Consequently, the proposed Newton system seems to be a natural analogue of the simplex method in this respect. Moreover, since the primal and dual steps are not orthogonal when the right-hand side or cost vector are changed, we need to control the second-order term in the change of the duality gap, and our choice does this nicely, guaranteeing that the new duality gap for the perturbed problem will be at least as small as the original one. Finally, our choice of right-hand side implies that the proportional changes $X^{-1}\Delta x$ and $S^{-1}\Delta s$ are negatives of one another, so our conditions become simply L_∞ bounds on a single vector.

Goffin and Sharifi-Mokhtarian [16] also use a similar choice for the Newton step in a different setting: they study the analytic center cutting plane method for solving convex feasibility problems which approximates analytic centers of polyhedra containing the convex set generated via cutting planes. After adding a cut (which possibly makes the current approximate center infeasible) the center is updated based on an infeasible primal-dual Newton's method to restore primal-dual feasibility, where a similar choice to ours is used for the old variables to keep the analysis manageable. However, the motivation and the analysis are very different from ours.

Finally, we give the version of the two propositions above for directional perturbations, i.e., the right-hand side vector b is replaced by $b + \beta d_b$, and the cost vector

c is replaced by $c + \beta d_c$, where $\beta \in \mathbb{R}$, $d_b \in \mathbb{R}^m$ and $d_c \in \mathbb{R}^n$.

Proposition 2.2.3 *Assume that (x, y, s) is a strictly feasible point for $LPP(b, c)$ and $LPD(b, c)$ and the right-hand side vector b and the cost vector c are replaced by $b' := b + \beta d_b$ and $c' := c + \beta d_c$, respectively, where $\beta \in \mathbb{R}$, $d_b \in \mathbb{R}^m$, and $d_c \in \mathbb{R}^n$. Suppose a Newton step is taken from (x, y, s) targeting the feasible point (x', y', s') of $LPP(b', c')$ and $LPD(b', c')$ that satisfies $X'S'e = XSe$. Then a full Newton step will yield a feasible iterate for the new problem with duality gap at most $x^T s$ if and only if*

$$|\beta| \leq \frac{1}{\|S^{-1}(I - A^T(AD^2A^T)^{-1}AD^2)d_c - S^{-1}A^T(AD^2A^T)^{-1}d_b\|_\infty}, \quad (2.12)$$

where $D = X^{\frac{1}{2}}S^{-\frac{1}{2}}$.

Proof: Using (2.3), we have $r_p = \beta d_b$, $r_d = \beta d_c$, and $r_{xs} = 0$ by the hypothesis.

Therefore,

$$\|S^{-1}\Delta s\|_\infty = |\beta| \|S^{-1}(I - A^T(AD^2A^T)^{-1}AD^2)d_c - S^{-1}A^T(AD^2A^T)^{-1}d_b\|_\infty, \quad (2.13)$$

from which the result follows immediately. \square

2.2.2 Simplex Approach

Here we give the bounds derived from the optimal basis. Since the simplex method yields an optimal basic solution, as noted above we call this the simplex approach.

First, we consider changes in the right-hand side vector b . It is clear that as long as Δb satisfies certain conditions, the optimal basis for the original LP will remain optimal for the new LP.

Let x^* be an optimal solution for the original LP, and assume that it is partitioned as x_B^* and x_N^* , corresponding to the basic and nonbasic variables, respectively. Similarly, assume that the columns of the coefficient matrix A are partitioned into B and N accordingly. Let the right-hand side vector b be replaced by $b + \Delta b$, where $\Delta b \in \mathbb{R}^m$. Then the optimal basis will remain optimal for the new problem if and only if primal feasibility is retained:

$$B^{-1}(b + \Delta b) \geq 0 \quad \text{or} \quad B^{-1}\Delta b \geq -B^{-1}b = -x_B^*. \quad (2.14)$$

Clearly, the simplex approach yields “one-sided” bounds as opposed to the “two-sided” bounds we have in the interior-point approach.

Next, we consider changes in the cost vector c . Assume that c is replaced by $c + \Delta c$, where $\Delta c \in \mathbb{R}^n$. Once again, partition c as c_B and c_N , and Δc as Δc_B and Δc_N , corresponding to the basic and nonbasic variables, respectively. The optimal basis will remain optimal if and only if dual feasibility is retained (i.e., the dual slack variable s^* remains nonnegative):

$$c_N^T + \Delta c_N^T - c_B^T B^{-1}N - \Delta c_B^T B^{-1}N \geq 0 \quad \text{or} \quad \Delta c_N - N^T B^{-T} \Delta c_B \geq -s_N^*, \quad (2.15)$$

where s_N^* and s_B^* partition the dual optimal slack s^* . Hence, as long as Δc satisfies the above inequality, the same optimal basis will remain optimal for the new problem.

In the next subsection, we compare the two approaches under the assumption of a unique, nondegenerate optimal solution. Before doing that, we illustrate with a small example what can go wrong with the interior-point bounds (2.4) and (2.11) in the degenerate case. Let (P) be given by $\min\{x_2 - x_1 : x_1 - x_2 = 0, x_2 + x_3 = 1, x \geq$

0}. Then (P) has multiple optimal solutions given by $(x_1, x_2, x_3) = (\beta, \beta, 1 - \beta)$ where $\beta \in [0, 1]$ with an optimal value of 0. The dual problem in this case has a unique but degenerate optimal solution. Let the right hand side be perturbed to $(0, 1)^T + t(2, 1)^T$. It has been shown by Adler and Monteiro [2] and Jansen, de Jong, Roos, and Terlaky [32] that maintaining the optimal partition rather than an optimal basis gives more accurate information about the range of t . The optimal partition-based bounds for t in this example are $(-1/3, +\infty)$. Fixing a near-optimal dual strictly feasible point at $y = (-1 - \epsilon, -2\epsilon)^T$, $s = (\epsilon, \epsilon, 2\epsilon)^T$ for small $\epsilon > 0$, we evaluate the interior-point bound (2.4) at various primal strictly feasible points as a function of β . The interior-point bound yields $\pm\beta/(2\beta + 1)$ as the limits for t ; the upper bound increases from 0 to the desired symmetrized value of $1/3$ as β goes from 0 to 1. One can come up with a similar example for perturbations of c . This shows that the interior-point bounds depend on the near-optimal solution at which they are evaluated in the presence of degeneracy, contrary to the situation under nondegeneracy as we show in the next subsection. However, in a follow-up paper (Chapter 3), we will show that we can still say something about the quality of the interior-point bounds even under degeneracy.

2.2.3 Comparison of the Simplex and Interior-Point Approaches

Recall that Propositions 2.2.1 through 2.2.3 hold for any strictly feasible pair of solutions for $LPP(b, c)$ and $LPD(b, c)$. Clearly, they cannot be applied directly

to the optimal solution pair since strict feasibility is violated. Hence, we need to obtain a “good” strictly feasible point for $LPP(b, c)$ and $LPD(b, c)$ so that we can compare the conditions and bounds from the simplex approach with those arising from the interior-point approach. Throughout this subsection, we will assume that the original LP has a unique, nondegenerate solution, with basic and nonbasic variables indicated by the subscripts B and N as above. Thus the optimal primal solution is $x^* = (x_B^*; x_N^*)$ and the optimal dual solution $(y^*, s^*) = (y^*, (s_B^*; s_N^*))$ with $x_B^* > 0$, $x_N^* = 0$, $s_B^* = 0$, and $s_N^* > 0$.

We will first compare the conditions and bounds where those for the interior-point approach are generated from a strictly feasible point that is close to optimal and also close to the central path. We show that asymptotically the same conditions and bounds are generated by the two approaches, as long as the simplex (or basis) conditions are “symmetrized” to make them two-sided like those from the interior-point approach. Then we will consider *any* strictly feasible point that is close to optimal and show that similar results continue to hold.

The basis condition (2.14) on Δb can be written as $(X_B^*)^{-1}B^{-1}\Delta b \geq -e$. We will call the symmetrized condition the strengthening where this vector must lie between $-e$ and e , or

$$\|(X_B^*)^{-1}B^{-1}\Delta b\|_\infty \leq 1. \quad (2.16)$$

If we follow a similar treatment on Δc , the basis condition (2.15) on Δc can be written as $(S_N^*)^{-1}(\Delta c_N - N^T B^{-T} \Delta c_B) \geq -e$; as above, the symmetrized condition is then

$$\|(S_N^*)^{-1}(\Delta c_N - N^T B^{-T} \Delta c_B)\|_\infty \leq 1. \quad (2.17)$$

Recall that the central path is the set of solutions for positive μ of the system

$$\begin{aligned} A^T y + s &= c, \\ Ax &= b, \\ XSe &= \mu e, \end{aligned} \tag{2.18}$$

with $x > 0$ and $s > 0$. Adler and Monteiro [1] show that the above system indeed defines a continuous and differentiable path of solutions parametrized by μ , and that as μ approaches 0, the points on the central path converge to the analytic center of the optimal face. They also analyze the limiting behavior of the central path and show that the derivative of the path as a function of μ has a limit as μ tends to 0. Here is their result, which holds regardless of degeneracy if x and s are partitioned with respect to the optimal partition.

Theorem 2.2.1 *Let $(x^*, y^*, s^*) = \lim_{\mu \rightarrow 0} (x(\mu), y(\mu), s(\mu))$. Let x and s be partitioned as x_B, x_N, s_B and s_N . Then $\lim_{\mu \rightarrow 0} \dot{x}_N(\mu) = (s_N^*)^{-1}$ and $\lim_{\mu \rightarrow 0} \dot{s}_B(\mu) = (x_B^*)^{-1}$.*

Here, $(s_N^*)^{-1}$ denotes the vector of inverses of the components of s_N^* and similarly for $(x_B^*)^{-1}$. We refer the reader to the proofs of Theorems 5.1 and 5.3 in [1]. Note that our assumption of a unique, nondegenerate solution implies that the optimal partition coincides with the basis partition. Hence, we immediately get a closed form expression for the derivative of $x_B(\mu)$: note that $Bx_B(\mu) + Nx_N(\mu) = b$ implies $B\dot{x}_B(\mu) + N\dot{x}_N(\mu) = 0$ or $\dot{x}_B(\mu) = -B^{-1}N\dot{x}_N(\mu)$. Hence, by Theorem 2.2.1, we have:

$$\lim_{\mu \rightarrow 0} \dot{x}_B(\mu) = -B^{-1}N(s_N^*)^{-1}. \tag{2.19}$$

Similarly, we also get a closed form expression for the derivative of $s_N(\mu)$ as follows: we have $B^T y(\mu) + s_B(\mu) = c_B$ or $y(\mu) = B^{-T}(c_B - s_B(\mu))$ and so $N^T y(\mu) + s_N(\mu) = c_N$ gives $s_N(\mu) = c_N - N^T B^{-T}(c_B - s_B(\mu))$. Differentiating this last equation with respect to μ , taking the limit as μ tends to 0 and using Theorem 2.2.1, we have:

$$\lim_{\mu \rightarrow 0} \dot{s}_N(\mu) = N^T B^{-T}(x_B^*)^{-1}. \quad (2.20)$$

The strictly feasible point we will initially use in our analysis of the interior-point approach is obtained by taking a first-order Taylor approximation from the optimal solution (x^*, y^*, s^*) using the above theorem. Clearly, for small enough μ , the point will be a good approximation to $(x(\mu), y(\mu), s(\mu))$. Consequently, we have the following strictly feasible point:

$$\begin{aligned} x_B &= x_B^* - \mu B^{-1} N (s_N^*)^{-1} && \approx x_B(\mu), \\ x_N &= \mu (s_N^*)^{-1} && \approx x_N(\mu), \\ s_B &= \mu (x_B^*)^{-1} && \approx s_B(\mu), \\ s_N &= s_N^* + \mu N^T B^{-T} (x_B^*)^{-1} && \approx s_N(\mu). \end{aligned} \quad (2.21)$$

With $y = y^* - \mu B^{-T} (x_B^*)^{-1} = B^{-T}(c_B - \mu (x_B^*)^{-1}) \approx y(\mu)$, it is easy to verify that the resulting points (x, y, s) will be strictly feasible for small enough μ ; moreover, it is easy to check that the duality gap of (x, y, s) is μn , the same as that of the corresponding point on the central path. Therefore, in the case of a unique non-degenerate solution to $LPP(b, c)$, we have a strictly feasible point to use in our analysis for the interior-point approach.

From Proposition 2.2.1, we need to compute the following matrix:

$$S^{-1} A^T (A D^2 A^T)^{-1}. \quad (2.22)$$

Instead, it is easier to work with its row permutation:

$$\begin{aligned} \begin{bmatrix} S_B^{-1} \\ S_N^{-1} \end{bmatrix} \begin{bmatrix} B^T \\ N^T \end{bmatrix} \left(\begin{bmatrix} B & N \end{bmatrix} \begin{bmatrix} S_B^{-1} \\ S_N^{-1} \end{bmatrix} \begin{bmatrix} X_B \\ X_N \end{bmatrix} \begin{bmatrix} B^T \\ N^T \end{bmatrix} \right)^{-1} = \\ \begin{bmatrix} S_B^{-1} B^T \\ S_N^{-1} N^T \end{bmatrix} (BS_B^{-1} X_B B^T + NS_N^{-1} X_N N^T)^{-1}. \end{aligned} \quad (2.23)$$

Next, we substitute the values from (2.21). In order to simplify the computations, we will frequently use the following formulae. Suppose M is a square matrix with $\|M\| \leq 1/2$ (we can use any of several norms here, but let us suppose this is the L_2 -operator norm). Then the Neumann lemma [20] implies that $I + M$ is invertible with $\|(I + M)^{-1}\| \leq 2$, and it is then easy to see that

$$(I + M)^{-1} = I - M(I + M)^{-1}.$$

Next suppose that U is invertible and $\|U^{-1}V\| \leq 1/2$. Then applying the result above to $M = U^{-1}V$ we get $U + V = U(I + U^{-1}V)$ invertible, $\|(I + U^{-1}V)^{-1}\| \leq 2$, and

$$(U + V)^{-1} = U^{-1} - U^{-1}V(I + U^{-1}V)^{-1}U^{-1}. \quad (2.24)$$

We will apply this result with $U := BS_B^{-1}X_B B^T$ and $V := NS_N^{-1}X_N N^T$. Note that $U^{-1} = B^{-T}S_B(X_B)^{-1}B^{-1}$ and that U^{-1} and V are $O(\mu)$ (by this we mean each entry is of the stated order).

Now we return to (2.23). We find that

$$S_B^{-1}B^T U^{-1} = (X_B)^{-1}B^{-1} = (X_B^*)^{-1}B^{-1} + O(\mu), \quad (2.25)$$

and from this the top part of the matrix is $(X_B^*)^{-1}B^{-1} + O(\mu)$. Since $(S_N)^{-1} = (S_N^*)^{-1} + O(\mu) = O(1)$, the bottom part of the matrix is $O(\mu)$ since U^{-1} and V are.

Hence (2.23) is

$$\begin{bmatrix} (X_B^*)^{-1}B^{-1} + O(\mu) \\ O(\mu) \end{bmatrix}.$$

This generates the necessary and sufficient condition

$$\left\| \begin{bmatrix} (X_B^*)^{-1}B^{-1} + O(\mu) \\ O(\mu) \end{bmatrix} \Delta b \right\|_{\infty} \leq 1, \quad (2.26)$$

which is asymptotically the same as the symmetrized basis condition (2.16).

Next, we consider a change in the cost vector c . From Proposition 2.2.2, we need to evaluate the following:

$$S^{-1}(I - A^T(AD^2A^T)^{-1}AD^2). \quad (2.27)$$

Permuting both the rows and the columns yields the following:

$$\begin{bmatrix} S_B^{-1} \\ S_N^{-1} \end{bmatrix} \left(I - \begin{bmatrix} B^T \\ N^T \end{bmatrix} (U + V)^{-1} \begin{bmatrix} BX_B S_B^{-1} & NX_N S_N^{-1} \end{bmatrix} \right),$$

where U and V are as defined before.

Let us examine each block of this 2×2 block matrix. The top left block is $(S_B)^{-1} - (S_B)^{-1}B^T(U + V)^{-1}BX_B(S_B)^{-1}$. Using our expressions for $(U + V)^{-1}$ and for $(S_B)^{-1}B^T U^{-1}$ in (2.24) and (2.25), we find that this equals

$$(S_B)^{-1} - (S_B)^{-1} + (X_B)^{-1}B^{-1}V(I + U^{-1}V)^{-1}B^{-T}, \quad (2.28)$$

which is $O(\mu)$ since V is. Similarly, the top right block can be written as $-(S_B)^{-1}B^T(U + V)^{-1}NX_N(S_N)^{-1}$, which simplifies using the same two equations

to

$$(X_B)^{-1}B^{-1}(I - V(I + U^{-1}V)^{-1}U^{-1})NX_N(S_N)^{-1},$$

and this is again $O(\mu)$ because X_N is.

The bottom left block is $-(S_N)^{-1}N^T(U + V)^{-1}BX_B(S_B)^{-1}$. Once again using these equations, we find that this simplifies to $-(S_N)^{-1}N^T(I - U^{-1}V(I + U^{-1}V)^{-1})B^{-T}$, which equals (since U^{-1} and V are $O(\mu)$)

$$-(S_N)^{-1}N^TB^{-T} + O(\mu^2) = -(S_N^*)^{-1}N^TB^{-T} + O(\mu).$$

Finally, the bottom right block is $(S_N)^{-1} - (S_N)^{-1}N^T(U + V)^{-1}NX_N(S_N)^{-1}$. Using (2.24) we can approximate this (since U^{-1} and X_N are $O(\mu)$) as

$$(S_N)^{-1} + O(\mu^2) = (S_N^*)^{-1} + O(\mu).$$

Our necessary and sufficient condition then reduces to

$$\left\| \left[\begin{array}{cc} O(\mu) & O(\mu) \\ -(S_N^*)^{-1}N^TB^{-T} + O(\mu) & (S_N^*)^{-1} + O(\mu) \end{array} \right] \left[\begin{array}{c} \Delta c_B \\ \Delta c_N \end{array} \right] \right\|_{\infty} \leq 1, \quad (2.29)$$

and again this is asymptotically identical to the basis condition (2.17).

We conclude this section by generalizing our results (2.26) and (2.29). In deriving these results, we used an approximation to the point on the central path based on a first-order Taylor approximation from the optimal solution. In the next theorem, we show that the same asymptotic result can be obtained using *any* strictly feasible solution (x, y, s) with a small duality gap $\mu n := x^T s$, which makes our results algorithmically more applicable. In the theorem (i.e., in the bounds (2.26) and (2.29)) and in the rest of this section, we use $O(\mu)$ to denote a scalar, vector, or

matrix whose entries may depend on (x, y, s) but are bounded by a multiple of μ ; this multiple can depend on B and N and on (x^*, y^*, s^*) , but does not depend on the strictly feasible solution (x, y, s) . This is the meaning of the term “uniformly” in the statement.

Theorem 2.2.2 *Under the assumption of a unique, nondegenerate solution, the expressions (2.4) and (2.11) yield the asymptotic results (2.26) and (2.29), respectively for all strictly feasible points (x, y, s) uniformly in μ where $\mu := x^T s/n$. These bounds converge to the symmetrized simplex bounds (2.16) and (2.17) as μ approaches zero.*

To prove Theorem 2.2.2, we use the following lemma. In fact, the lemma holds for any feasible point (x, y, s) and even for a point where feasibility is violated by $O(\mu)$, but the statement below suffices for our needs.

Lemma 2.2.1 *Under the assumption of a unique, nondegenerate solution, let (x, y, s) be any strictly feasible solution with duality gap μn , let (x^*, y^*, s^*) be the optimal solution and let the coefficient matrix A be partitioned as B and N , corresponding to the basic and nonbasic variables, respectively. Then x and s satisfy:*

$$\begin{aligned} x_B &= x_B^* + O(\mu), & x_N &= O(\mu), \quad x_N > 0, \\ s_B &= O(\mu), \quad s_B > 0, & s_N &= s_N^* + O(\mu), \end{aligned} \tag{2.30}$$

where the subscripts indicate the appropriate partitions with respect to B and N .

Proof: Note that x^* is the unique solution to $A\hat{x} = b$, $(s^*)^T \hat{x} \leq 0$, $\hat{x} \geq 0$. Since x satisfies this system with the second right-hand side changed to μn , the result

for x follows from Hoffman's lemma [30]. A similar argument applies to the dual problem. \square

Now we are ready to prove Theorem 2.2.2.

Proof of Theorem 2.2.2: Let (x, y, s) be any strictly feasible solution with duality gap μn . By Lemma 2.2.1, x and s have the form (2.30). Let us re-examine how we obtained (2.26) from (2.22). The only use we made of the form of (x, y, s) was that it satisfied (2.30). The major difference is that now we cannot bound S_B^{-1} by $O(\mu^{-1})$; but in our derivation, all occurrences of S_B^{-1} cancel, and no bound is necessary. In particular, see (2.25) which shows how the S_B^{-1} terms disappear.

Next we reconsider the derivation of (2.29) from (2.27). Once again, all we required is (2.30), and the S_B^{-1} terms vanish; see, e.g., (2.28), where two such terms cancel.

We conclude that our earlier proof goes through unchanged, and this establishes the theorem. \square

We find it remarkable that the same bounds are produced asymptotically by any strictly feasible point, whereas it seems that solutions close to the boundary of the feasible region would generate much worse bounds, since perturbations appear much more likely to lead to infeasibility. However, we have shown that this is not the case. This analysis may shed some light on how well interior-point methods work even when their iterates lie very close to the boundary of the feasible region.

We conclude this section with a brief note on what the $O(\mu)$ terms depend on in the asymptotic results (2.26) and (2.29). The analysis reveals that those terms are

determined by the condition number of B and the minimum components of x_B^* and s_N^* , which act as a condition measure for LPs. As B or the LP gets ill-conditioned, the convergence would require increasingly smaller duality gaps.

2.3 Semidefinite Programming

We consider the SDP given in the following standard form:

$$\begin{aligned}
 (SDP) \quad \min_X \quad & C \bullet X \\
 & A_i \bullet X = b_i, \quad i = 1, \dots, m, \\
 & X \succeq 0,
 \end{aligned}$$

where all $A_i \in \mathcal{S}^n$, $b \in \mathbb{R}^m$, $C \in \mathcal{S}^n$ are given, and $X \in \mathcal{S}^n$. Here \mathcal{S}^n denotes the space of $n \times n$ symmetric matrices, and $X \succeq 0$ indicates that X is symmetric positive semidefinite. Similarly, $X \succ 0$ will indicate that X is symmetric positive definite. The notation $P \bullet Q$ represents the usual inner product $\text{Trace}(P^T Q) = \sum_{ij} P_{ij} Q_{ij}$ on $n \times n$ matrices, and the Frobenius norm $\|P\|_F := (P \bullet P)^{1/2}$ is the associated norm. We assume that the set $\{A_i\}$ is linearly independent. The dual problem associated with (SDP) is:

$$\begin{aligned}
 (SDD) \quad \max_{y,S} \quad & b^T y \\
 & \sum_{i=1}^m y_i A_i + S = C, \\
 & S \succeq 0,
 \end{aligned}$$

where $y \in \mathbb{R}^m$ and $S \in \mathcal{S}^n$. Once again, we will parametrize SDP and SDD by b and C , and the matrices A_i will be fixed. Note that LP is a special case of SDP

where all the matrices A_i and C are diagonal; then S is automatically diagonal, and any X can be replaced by its diagonal restriction without loss of generality.

The concept of the central path can be extended to SDP. If we assume that both $SDP(b, C)$ and $SDD(b, C)$ have strictly feasible solutions (i.e., with X and S positive definite), the central path is defined as the set of solutions $(X(\mu), y(\mu), S(\mu))$ for $\mu > 0$ to the following system together with the requirement that X and S are symmetric positive definite:

$$\begin{aligned} \sum_{i=1}^m y_i A_i + S &= C, \\ A_i \bullet X &= b_i, \quad \text{for } i = 1, \dots, m, \\ XS &= \mu I. \end{aligned} \tag{2.31}$$

A crucial observation is that Newton's method cannot be directly applied to (2.31). The reason is that the residual map takes an iterate $(X, y, S) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$ to a point in $\mathbb{R}^m \times \mathcal{S}^n \times \mathbb{R}^{n \times n}$ (since $XS - \mu I$ is in general not symmetric), which is a space of higher dimension. Many authors have suggested different ways of symmetrizing the third equation in (2.31) so that the residual lies in \mathcal{S}^n . Todd [60] analyzes twenty different search directions for SDP.

Next, we introduce some notation that we will use throughout this section. Script letters will denote linear operators on symmetric matrices. In particular, $\mathcal{A} : \mathcal{S}^n \rightarrow \mathbb{R}^m$ is defined by

$$\mathcal{A}U := (A_i \bullet U)_{i=1}^m, \tag{2.32}$$

with adjoint $\mathcal{A}^* : \mathbb{R}^m \rightarrow \mathcal{S}^n$; then

$$\mathcal{A}^*y = \sum_{i=1}^m y_i A_i. \tag{2.33}$$

We use $\|\cdot\|_2$ to denote the L_2 -operator norm on matrices, and $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ to denote the minimum and maximum eigenvalues of a symmetric matrix.

The directions we will examine will be Newton steps for nonlinear systems of the form

$$\begin{aligned} \mathcal{A}^* \tilde{y} + \tilde{S} &= C, \\ \mathcal{A} \tilde{X} &= b, \\ \Theta(\tilde{X}, \tilde{S}) &= \Theta(X', S'), \end{aligned} \tag{2.34}$$

where $\Theta(X, S)$ is some symmetrization of XS and where X' and S' are the targeted points. (Once again, X' and S' typically form the point on the central path satisfying $X'S' = \mu I$ for some $\mu > 0$, and μ is decreased at each iteration towards 0. We assume that $\Theta(X', S')$ is known for such points even if X' and S' are not.) Therefore, the Newton direction will be given by the solution of the following system:

$$\begin{aligned} \mathcal{A}^* \Delta y + \Delta S &= R_d, \\ \mathcal{A} \Delta X &= r_p, \\ \mathcal{E} \Delta X + \mathcal{F} \Delta S &= R_{EF}, \end{aligned} \tag{2.35}$$

where $r_p = b - \mathcal{A}X$ is the primal residual, $R_d = C - \mathcal{A}^*y - S$ is the dual residual, the operators $\mathcal{E} = \mathcal{E}(X, S)$ and $\mathcal{F} = \mathcal{F}(X, S)$ are the derivatives of Θ with respect to \tilde{X} and \tilde{S} respectively, evaluated at (X, S) , and $R_{EF} = R_{EF}(X, S) = \Theta(X', S') - \Theta(X, S)$. We will also use the following notation introduced by Alizadeh, Haeberly, and Overton [5]:

$$(P \odot Q)K := \frac{1}{2}(PKQ^T + QKP^T), \tag{2.36}$$

where $P, Q \in \mathbb{R}^{n \times n}$ and $K \in \mathcal{S}^n$, and we will regard it as an operator from \mathcal{S}^n to \mathcal{S}^n . The adjoint operator is defined as usual by $\mathcal{E}^*U \bullet V = U \bullet \mathcal{E}V$ for all U, V , and

it is easy to see that

$$Q \odot P = P \odot Q, \quad (P \odot Q)^* = P^T \odot Q^T, \quad (2.37)$$

so that $P \odot Q$ is self-adjoint if P and Q are symmetric. If moreover P and Q are positive definite, then

$$(P \odot Q)U \bullet U = \text{Trace}(PUQU) = \text{Trace}(P^{1/2}UQ^{1/2}Q^{1/2}UP^{1/2}) = \|P^{1/2}UQ^{1/2}\|_F^2,$$

so that $P \odot Q$ is also positive definite. If P is nonsingular,

$$(P \odot P)^{-1} = P^{-1} \odot P^{-1},$$

but there is no simple expression for $(P \odot Q)^{-1}$ in general. Note that $I \odot I$ is the identity operator. Very occasionally, we will extend the domain of the operator $P \odot Q$ to all of $\mathbb{R}^{n \times n}$; for possibly nonsymmetric matrices K , we define it by

$$(P \odot Q)K := \frac{1}{2}(PKQ^T + QK^TP^T). \quad (2.38)$$

Assume that \mathcal{E} is nonsingular. Then the operator $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$ takes \mathbb{R}^m into itself, so is represented by an $m \times m$ matrix, called the Schur complement. (It is unnecessary to represent the operators \mathcal{E} and \mathcal{F} as matrices to define or evaluate the Schur complement.) We find that (2.35) has a unique solution iff the Schur complement matrix is nonsingular, and in this case the solution can be found from

$$\begin{aligned} (\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)\Delta y &= r_p - \mathcal{A}\mathcal{E}^{-1}(R_{EF} - \mathcal{F}R_d), \\ \Delta S &= R_d - \mathcal{A}^*\Delta y, \\ \Delta X &= \mathcal{E}^{-1}(R_{EF} - \mathcal{F}\Delta S). \end{aligned} \quad (2.39)$$

In this paper, we will analyze the AHO, H..K..M, and the NT directions, as well as the general family of Monteiro-Y. Zhang search directions [42, 74, 46]. The AHO direction was suggested by Alizadeh, Haeberly and Overton [5]. The H..K..M direction was independently introduced by Helmberg, Rendl, Vanderbei and Wolkowicz [29] and Kojima, Shindoh and Hara [39], and later rediscovered by Monteiro [42]. Finally, Nesterov and Todd [48, 49] introduced the NT direction.

The reason for considering the above three directions is twofold. Firstly, the H..K..M and NT directions give a unique search direction for every symmetric positive definite X and S and surjective operator \mathcal{A} . The AHO direction also enjoys this property if $XS + SX$ is symmetric positive semidefinite [57, 62] or if (X, y, S) lies in a suitable neighborhood of the central path [45]. Moreover, the first two directions possess the property that $\mathcal{E}^{-1}\mathcal{F}$ is positive definite, which will lead to a reduction in the duality gap for the new problem arising from perturbations of b and C as in the LP case; moreover, $\mathcal{E}^{-1}\mathcal{F}$ is self-adjoint, so that the Schur complement matrix is symmetric in these cases. (Note that an operator \mathcal{G} from \mathcal{S}^n to \mathcal{S}^n is positive definite if $U \bullet \mathcal{G}U > 0$ for every nonzero $U \in \mathcal{S}^n$; it is self-adjoint if $U \bullet \mathcal{G}V = V \bullet \mathcal{G}U$ for every $U, V \in \mathcal{S}^n$.) Again, the AHO direction enjoys the positive-definiteness property (but not self-adjointness in general) if $XS + SX$ is symmetric positive semidefinite. Our second reason for analyzing these three directions is that they are among the search directions used most frequently in practice. In the next subsection, we will present our general results for the Monteiro-Zhang family of search directions for the SDP. Then we will turn our attention to the three specific search directions stated above. Finally, we will show that for these three cases, if the SDP

under consideration is derived from an LP problem, then the bounds reduce to those we obtained above for the LP case.

Since the derivation is somewhat technical, and the results cannot be stated precisely without some initial analysis, the reader may wish to skip Subsections 2.3.2–2.3.4 on a first reading.

2.3.1 General Results

We assume that there is a strictly feasible point (X, y, S) for $SDP(b, C)$ and $SDD(b, C)$ defined in the obvious way. We also assume that \mathcal{A} is a surjective operator, which follows if the A_i s are linearly independent. Clearly, the duality gap corresponding to this point will be given by $C \bullet X - b^T y = X \bullet S > 0$, since both X and S are symmetric positive definite. We further assume that the operators \mathcal{E} and \mathcal{F} are in the following form:

$$\mathcal{E} = S \odot M, \quad \mathcal{F} = MX \odot I, \quad (2.40)$$

where M is a symmetric positive definite matrix; this defines precisely the Monteiro-Zhang family of search directions. As is known and will also be seen shortly, this assumption holds for the AHO, H..K..M and NT directions. From (2.37), the adjoint operators are given by

$$\mathcal{E}^* = S \odot M, \quad \mathcal{F}^* = XM \odot I. \quad (2.41)$$

Under the assumption (2.40), \mathcal{E} is nonsingular. Moreover, $\mathcal{E}^{-1}\mathcal{F}$ is positive definite for the H..K..M and NT directions, and this also holds for the AHO direction if

$XS + SX$ is symmetric positive semidefinite. Note that since \mathcal{A} is surjective, the Schur complement matrix $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$ will then be nonsingular. Finally, $V^{\frac{1}{2}}$ will denote the unique symmetric positive definite square root of the symmetric positive definite matrix V .

First, we consider a change in the right-hand side vector b .

Proposition 2.3.1 *Assume that (X, y, S) is a strictly feasible point for $SDP(b, C)$ and $SDD(b, C)$ and let \mathcal{E} and \mathcal{F} as in (2.35) be given by (2.40). Assume that $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$ is nonsingular. Let the right-hand side vector b be replaced by $b' := b + \Delta b$, where $\Delta b \in \mathbb{R}^m$, and suppose a Newton step for the system (2.34) is taken from (X, y, S) targeting the feasible point (X', y', S') of $SDP(b', C)$ and $SDD(b', C)$ that satisfies $\Theta(X', S') = \Theta(X, S)$. Then a full Newton step can be taken and the resulting iterate will be feasible for the new problems if, and only if, Δb satisfies the following inequalities:*

$$\lambda_{\min} \left(X^{-\frac{1}{2}} (\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^* [(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\Delta b]) X^{-\frac{1}{2}} \right) \geq -1, \quad (2.42)$$

$$\lambda_{\max} \left(S^{-\frac{1}{2}} (\mathcal{A}^* [(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\Delta b]) S^{-\frac{1}{2}} \right) \leq 1. \quad (2.43)$$

Moreover, the duality gap of the new iterate will be at most $X \bullet S$ if $\mathcal{E}^{-1}\mathcal{F}$ is positive definite.

Proof: Note that by the hypothesis, we have $r_p = \Delta b$, $R_d = 0$ and $R_{EF} = 0$. Then, from (2.39), the Newton step $(\Delta X, \Delta y, \Delta S)$ is given by:

$$\begin{aligned} \Delta y &= (\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\Delta b, \\ \Delta S &= -\mathcal{A}^* [(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\Delta b], \\ \Delta X &= \mathcal{E}^{-1}\mathcal{F}\mathcal{A}^* [(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\Delta b]. \end{aligned} \quad (2.44)$$

Then, clearly, the next iterate will be feasible for the new problem if and only if $X + \Delta X \succeq 0$ and $S + \Delta S \succeq 0$. But,

$$X + \Delta X \succeq 0 \quad \text{holds iff} \quad I + X^{-\frac{1}{2}} \Delta X X^{-\frac{1}{2}} \succeq 0. \quad (2.45)$$

(2.45) implies that all the eigenvalues of the symmetric matrix $X^{-\frac{1}{2}} \Delta X X^{-\frac{1}{2}}$ should be greater than or equal to -1. With this observation and combining (2.44) with (2.45), we have exactly (2.42) as a necessary and sufficient condition for the new X iterate to be feasible. Similarly,

$$S + \Delta S \succeq 0 \quad \text{holds iff} \quad I + S^{-\frac{1}{2}} \Delta S S^{-\frac{1}{2}} \succeq 0; \quad (2.46)$$

combining (2.44) with (2.46), and using the same argument, we have exactly (2.43) as a necessary and sufficient condition for the new S iterate to be feasible.

Next, we will show that the duality gap of the new iterate is at most the original duality gap given by $X \bullet S$, assuming $\mathcal{E}^{-1}\mathcal{F}$ is positive definite. Note that

$$(X + \Delta X) \bullet (S + \Delta S) = X \bullet S + X \bullet \Delta S + S \bullet \Delta X + \Delta X \bullet \Delta S. \quad (2.47)$$

By our hypothesis, \mathcal{E} and \mathcal{F} are given by (2.40). Therefore, if we use (2.41), it is easy to verify that

$$\mathcal{E}^* M^{-1} = S, \quad \mathcal{F}^* M^{-1} = X. \quad (2.48)$$

Then, using (2.48), we have:

$$\Delta X \bullet S = \Delta X \bullet \mathcal{E}^* M^{-1} = \mathcal{E} \Delta X \bullet M^{-1}. \quad (2.49)$$

Similarly,

$$\Delta S \bullet X = \Delta S \bullet \mathcal{F}^* M^{-1} = \mathcal{F} \Delta S \bullet M^{-1}. \quad (2.50)$$

However, (2.35) and our hypothesis imply $\mathcal{E}\Delta X + \mathcal{F}\Delta S = 0$. Combining this with (2.49) and (2.50), we obtain

$$\Delta X \bullet S + X \bullet \Delta S = 0. \quad (2.51)$$

Finally,

$$\Delta X \bullet \Delta S = -\Delta S \bullet \mathcal{E}^{-1}\mathcal{F}\Delta S \leq 0, \quad (2.52)$$

since $\mathcal{E}^{-1}\mathcal{F}$ is positive definite. Hence, (2.47), (2.51), and (2.52) imply:

$$(X + \Delta X) \bullet (S + \Delta S) \leq X \bullet S. \quad (2.53)$$

This completes the proof. \square

Next, we consider perturbations of the cost matrix C .

Proposition 2.3.2 *Assume that (X, y, S) is a strictly feasible point for $SDP(b, C)$ and $SDD(b, C)$ and let \mathcal{E} and \mathcal{F} as in (2.35) be given by (2.40). Assume that $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$ is nonsingular. Let the cost matrix C be replaced by $C' := C + \Delta C$, where $\Delta C \in \mathcal{S}^n$, and suppose a Newton step for the system (2.34) is taken from (X, y, S) targeting the feasible point (X', y', S') of $SDP(b, C')$ and $SDD(b, C')$ that satisfies $\Theta(X', S') = \Theta(X, S)$. Then a full Newton step can be taken and the resulting iterate will be feasible for the new problems if, and only if, ΔC satisfies the following inequalities:*

$$\lambda_{\max} \left(X^{-\frac{1}{2}} [\mathcal{E}^{-1}\mathcal{F}\Delta C - \mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\Delta C] X^{-\frac{1}{2}} \right) \leq 1, \quad (2.54)$$

$$\lambda_{\min} \left(S^{-\frac{1}{2}} [\Delta C - \mathcal{A}^*(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\Delta C] S^{-\frac{1}{2}} \right) \geq -1. \quad (2.55)$$

Moreover, the duality gap of the new iterate will be at most $X \bullet S$ if $\mathcal{E}^{-1}\mathcal{F}$ is positive definite.

Proof: Once again, using the hypothesis and the notation in (2.39), we have $r_p = 0$, $R_d = \Delta C$, and $R_{EF} = 0$. Then the Newton step $(\Delta X, \Delta y, \Delta S)$ is given by:

$$\begin{aligned}\Delta y &= (\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\Delta C, \\ \Delta S &= \Delta C - \mathcal{A}^*(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\Delta C, \\ \Delta X &= -\mathcal{E}^{-1}\mathcal{F}\Delta C + \mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\Delta C.\end{aligned}\tag{2.56}$$

Then, proceeding as in the proof of Proposition 2.3.1, we see that the next iterate will be feasible for the new problem if and only if $X + \Delta X \succeq 0$ and $S + \Delta S \succeq 0$. Using a similar argument as in the previous proof, these conditions become that the minimum eigenvalues of $X^{-\frac{1}{2}}\Delta X X^{-\frac{1}{2}}$ and $S^{-\frac{1}{2}}\Delta S S^{-\frac{1}{2}}$ be at least -1 . Then, using (2.56), we obtain exactly the bounds (2.54) and (2.55) we seek. Essentially the same argument as in Proposition 2.3.1 shows that, if $\mathcal{E}^{-1}\mathcal{F}$ is positive definite, the duality gap is bounded above by $X \bullet S$. This completes the proof. \square

Next, as in the LP case, we present our result for directional perturbations.

Proposition 2.3.3 *Assume that (X, y, S) is a strictly feasible point for $SDP(b, C)$ and $SDD(b, C)$ and let \mathcal{E} and \mathcal{F} as in (2.35) be given by (2.40). Assume that $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$ is nonsingular. Let the right-hand side vector b and the cost matrix C be replaced by $b' := b + \beta d_b$, $C' := C + \beta D_C$, respectively, where $\beta \in \mathbb{R}$, $d_b \in \mathbb{R}^m$, and $D_C \in \mathcal{S}^n$. Suppose a Newton step for the system (2.34) is taken from (X, y, S) targeting the feasible point (X', y', S') of $SDP(b', C')$ and $SDD(b', C')$ that satisfies $\Theta(X', S') = \Theta(X, S)$. Then a full Newton step can be taken and the resulting iterate will be feasible for the new problems if, and only if, β satisfies the following:*

$$|\beta| \leq \min\{a, b\},\tag{2.57}$$

where a is the reciprocal of

$$\lambda_{\max} \left(X^{-\frac{1}{2}} (\mathcal{E}^{-1} \mathcal{F} [D_C - \mathcal{A}^* (\mathcal{A} \mathcal{E}^{-1} \mathcal{F} \mathcal{A}^*)^{-1} \mathcal{A} \mathcal{E}^{-1} \mathcal{F} D_C - \mathcal{A}^* (\mathcal{A} \mathcal{E}^{-1} \mathcal{F} \mathcal{A}^*)^{-1} d_b]) X^{-\frac{1}{2}} \right)$$

(or $+\infty$ if this quantity is negative) and b that of

$$-\lambda_{\min} \left(S^{-\frac{1}{2}} [D_C - \mathcal{A}^* (\mathcal{A} \mathcal{E}^{-1} \mathcal{F} \mathcal{A}^*)^{-1} \mathcal{A} \mathcal{E}^{-1} \mathcal{F} D_C - \mathcal{A}^* (\mathcal{A} \mathcal{E}^{-1} \mathcal{F} \mathcal{A}^*)^{-1} d_b] S^{-\frac{1}{2}} \right)$$

(again, $+\infty$ if this quantity is negative). Moreover, the duality gap of the new iterate will be at most $X \bullet S$ if $\mathcal{E}^{-1} \mathcal{F}$ is positive definite.

Proof: As in Propositions 2.3.1 and 2.3.2, the result follows simply by observing that $r_p = \beta d_b$, $R_D = \beta D_C$ and $R_{XS} = 0$ in (2.39), and imposing the conditions $\lambda_{\min}(S^{-\frac{1}{2}} \Delta S S^{-\frac{1}{2}}) \geq -1$ and $\lambda_{\min}(X^{-\frac{1}{2}} \Delta X X^{-\frac{1}{2}}) \geq -1$. \square

Before analyzing the three search directions, we would like to discuss the concept of *scale-invariance*. Given $SDP(b, C)$ and $SDD(b, C)$, if we apply a change of variable in $SDP(b, C)$ such that X is replaced by $\hat{X} = PXP^T$, where P is a nonsingular matrix in $\mathbb{R}^{n \times n}$, $SDP(b, C)$ transforms to

$$\begin{aligned} (\widehat{SDP}) \quad & \min_{\hat{X}} \hat{C} \bullet \hat{X} \\ & \hat{\mathcal{A}} \hat{X} = b, \\ & \hat{X} \succeq 0, \end{aligned}$$

where $\hat{C} := P^{-T} C P^{-1}$, and $\hat{\mathcal{A}}$ and $\hat{\mathcal{A}}^*$ are defined from $\{\hat{A}_i := P^{-T} A_i P^{-1}\}$ as in (2.32) and (2.33). The dual of this problem is

$$\begin{aligned} (\widehat{SDD}) \quad & \max_{\hat{y}, \hat{S}} b^T \hat{y} \\ & \hat{\mathcal{A}}^* \hat{y} + \hat{S} = \hat{C}, \\ & \hat{S} \succeq 0, \end{aligned}$$

which is exactly the transformation of $SDD(b, C)$ with (y, S) replaced by $(\hat{y} := y, \hat{S} := P^{-T}SP^{-1})$. If (X, y, S) is a strictly feasible point for $SDP(b, C)$ and $SDD(b, C)$, then $(\hat{X}, \hat{y}, \hat{S}) = (PXP^T, y, P^{-T}SP^{-1})$ is a strictly feasible point for \widehat{SDP} and \widehat{SDD} .

Now, we are in a position to discuss *P-scale-invariance* and *Q-scale-invariance* introduced by Todd [60]. A method for defining a search direction for semidefinite programming is called *P-scale-invariant* if the direction at any iterate is the same as would result from scaling the problem and iterate by an arbitrary nonsingular matrix P , using the method to determine the direction for the scaled problem, and then scaling back. It is called *Q-scale-invariant* if this is true when P is restricted to the set of orthogonal matrices. Todd shows that the H..K..M and NT directions are *P-scale invariant*, whereas all three directions we will analyze are *Q-scale invariant* (see Propositions 6.6 and 6.7 in [60]).

Furthermore, the Schur complement matrix given by $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$ is invariant under scaling, i.e., $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^* = \hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{\mathcal{A}}^*$. To see this, consider the i th column of the unscaled Schur complement matrix:

$$u = (\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)e_i = \mathcal{A}\mathcal{E}^{-1}\mathcal{F}A_i, \quad (2.58)$$

where e_i is the i th unit vector. Let $K = \mathcal{E}^{-1}\mathcal{F}A_i$. Then we have $\mathcal{E}K = \mathcal{F}A_i$. Using the fact that $\mathcal{E} = S \odot M$ and $\mathcal{F} = MX \odot I$ for our directions where M is a symmetric positive definite matrix (2.40), we have:

$$\frac{1}{2}(SKM + MKS) = \frac{1}{2}(MXA_i + A_iXM). \quad (2.59)$$

Then u in (2.58) is given by

$$u = \begin{bmatrix} A_1 \bullet K \\ \vdots \\ A_m \bullet K \end{bmatrix}. \quad (2.60)$$

As will be seen in the following analysis, the matrix M scales like S , i.e., $\hat{M} = P^{-T}MP^{-1}$, for the H..K..M and NT directions. Then $\hat{\mathcal{E}} = \hat{S} \odot \hat{M}$ and $\hat{\mathcal{F}} = \hat{M}\hat{X} \odot I$. Therefore, the i th column of the scaled Schur complement matrix is given by:

$$u' = (\hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{\mathcal{A}}^*)e_i = \hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{A}_i. \quad (2.61)$$

Let $\hat{K} = \hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{A}_i$. Then $\hat{\mathcal{E}}\hat{K} = \hat{\mathcal{F}}\hat{A}_i$. Using the definitions of $\hat{\mathcal{E}}$ and $\hat{\mathcal{F}}$, and substituting the values for the scaled matrices, we have:

$$\begin{aligned} \frac{1}{2}(\hat{S}\hat{K}\hat{M} + \hat{M}\hat{K}\hat{S}) &= \frac{1}{2}(\hat{M}\hat{X}\hat{A}_i + \hat{A}_i\hat{X}\hat{M}) \quad \text{or} \\ \frac{1}{2}(P^{-T}SP^{-1}\hat{K}P^{-T}MP^{-1} + P^{-T}MP^{-1}\hat{K}P^{-T}SP^{-1}) &= \\ \frac{1}{2}(P^{-T}MP^{-1}PXP^TP^{-T}A_iP^{-1} + P^{-T}A_iP^{-1}PXP^TP^{-T}MP^{-1}). & \end{aligned} \quad (2.62)$$

Multiplying (2.62) by P^T from the left and P from the right, we get:

$$\frac{1}{2}(SP^{-1}\hat{K}P^{-T}M + MP^{-1}\hat{K}P^{-T}S) = \frac{1}{2}(MXA_i + A_iXM). \quad (2.63)$$

Comparing (2.59) with (2.63), we have the same symmetric matrix on the right-hand side. Since S and M are symmetric positive definite, both systems have the same unique solution, so that $\hat{K} = PKP^T$. Hence,

$$u' = \begin{bmatrix} \hat{A}_1 \bullet \hat{K} \\ \vdots \\ \hat{A}_m \bullet \hat{K} \end{bmatrix} = \begin{bmatrix} \text{Trace}(P^{-T}A_1P^{-1}PKP^T) \\ \vdots \\ \text{Trace}(P^{-T}A_mP^{-1}PKP^T) \end{bmatrix} = \begin{bmatrix} A_1 \bullet K \\ \vdots \\ A_m \bullet K \end{bmatrix}. \quad (2.64)$$

From (2.60) and (2.64), we conclude that the Schur complement matrix is invariant under scaling. With this observation, either the original iterate or the scaled one can be used to compute this matrix. We will make use of this observation in our analysis.

For the AHO direction, $M = I$, thus $\hat{M} = M = I = P^{-T}P^{-1}$ iff $P = P^{-T}$. This is the reason why, unlike the other directions, the AHO direction only enjoys Q -scale invariance.

2.3.2 The AHO Direction

The AHO direction [5] is the Newton step for the following symmetrization of the third equation in (2.34):

$$\Theta(\tilde{X}, \tilde{S}) := \frac{1}{2}(\tilde{X}\tilde{S} + \tilde{S}\tilde{X}) = \frac{1}{2}(X'S' + S'X'). \quad (2.65)$$

It corresponds to taking

$$\mathcal{E} = S \odot I, \quad \mathcal{F} = X \odot I, \quad R_{EF} = \frac{1}{2}(X'S' + S'X') - \frac{1}{2}(XS + SX). \quad (2.66)$$

Therefore, $M = I$ for the AHO direction. Recall from Section 2.3.1 that we need the operator \mathcal{E}^{-1} for our analysis. For the AHO direction, \mathcal{E} is given by (2.66), and \mathcal{E}^{-1} does not have a nice closed form expression. However, using the Q -scale invariance property, assuming that (X, y, S) is our current strictly feasible point for $SDP(b, C)$ and $SDD(b, C)$, we let $S = QDQ^T$ be the eigenvalue decomposition of S , where D is a diagonal matrix with strictly positive eigenvalues of S , and Q is an

orthogonal matrix. Then, using $P = Q^{-1}$ as a scaling matrix, we have:

$$\begin{aligned}\hat{X} &= Q^{-1}XQ^{-T} = Q^T X Q, \\ \hat{S} &= Q^T S Q = D, \\ \hat{A}_i &= Q^T A_i Q.\end{aligned}\tag{2.67}$$

With this transformation, $\hat{\mathcal{E}} = \hat{S} \odot I = D \odot I$. Therefore, $\hat{\mathcal{E}}^{-1}$ has a closed form expression: $U := \hat{\mathcal{E}}^{-1}R$ is given by

$$\hat{\mathcal{E}}U = R \quad \text{iff} \quad DU + UD = 2R \quad \text{iff} \quad U_{ij} = \frac{2R_{ij}}{d_i + d_j},\tag{2.68}$$

where U_{ij} denotes the (i, j) entry of the matrix U , and d_i the i th diagonal element of D .

Now we compute the Schur complement matrix for the current iterate (X, y, S) using the scaled iterate. Let $N = \hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{\mathcal{A}}^*$. Then the i th column of N is given by:

$$Ne_i = (\hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{\mathcal{A}}^*)e_i = \hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{A}_i = \hat{\mathcal{A}}\hat{\mathcal{E}}^{-1} \left[\frac{1}{2}(\hat{A}_i\hat{X} + \hat{X}\hat{A}_i) \right] = \hat{\mathcal{A}}\hat{K}_i,\tag{2.69}$$

where

$$(\hat{K}_i)_{kl} = \frac{(\hat{A}_i\hat{X} + \hat{X}\hat{A}_i)_{kl}}{d_k + d_l}\tag{2.70}$$

from (2.68). Let us also write K_i for $Q\hat{K}_iQ^T$. Therefore,

$$Ne_i = \begin{bmatrix} \hat{A}_1 \bullet \hat{K}_i \\ \vdots \\ \hat{A}_m \bullet \hat{K}_i \end{bmatrix} = \begin{bmatrix} A_1 \bullet K_i \\ \vdots \\ A_m \bullet K_i \end{bmatrix}.\tag{2.71}$$

(2.71) implies that the Schur complement matrix N is not symmetric in general (see also Proposition 6.4 in [60]).

As mentioned previously, the AHO direction does not satisfy the well-defined direction property, that is, the AHO direction may fail to exist at a strictly feasible iterate. Therefore, we assume that (X, S) is such that the Schur complement $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$ is nonsingular in this subsection.

First, we consider a change in the right-hand side vector b . If we use the scaled iterate $(\hat{X}, \hat{y}, \hat{S})$ and the fact that $\widehat{\Delta b} = \Delta b$, Proposition 2.3.1 implies that the following bounds on Δb are necessary and sufficient:

$$\begin{aligned}\lambda_{\min} \left(\hat{X}^{-\frac{1}{2}} \left[\hat{\mathcal{E}}^{-1} \hat{\mathcal{F}} \hat{\mathcal{A}}^* (N^{-1} \Delta b) \right] \hat{X}^{-\frac{1}{2}} \right) &\geq -1, \\ \lambda_{\max} \left(\hat{S}^{-\frac{1}{2}} \left[\hat{\mathcal{A}}^* (N^{-1} \Delta b) \right] \hat{S}^{-\frac{1}{2}} \right) &\leq 1.\end{aligned}\tag{2.72}$$

The first inequality in (2.72) yields the following:

$$\begin{aligned}\lambda_{\min} \left(\hat{X}^{-\frac{1}{2}} \left(\hat{\mathcal{E}}^{-1} \left[\frac{1}{2} \sum_{i=1}^m (N^{-1} \Delta b)_i (\hat{A}_i \hat{X} + \hat{X} \hat{A}_i) \right] \right) \hat{X}^{-\frac{1}{2}} \right) = \\ \lambda_{\min} \left(\sum_{i=1}^m (N^{-1} \Delta b)_i \left(\hat{X}^{-\frac{1}{2}} \hat{K}_i \hat{X}^{-\frac{1}{2}} \right) \right) \geq -1,\end{aligned}\tag{2.73}$$

where \hat{K}_i is as in (2.70) and $(N^{-1} \Delta b)_i$ denotes the i th component of the vector $N^{-1} \Delta b$. In this bound, we can also replace $\hat{X}^{-\frac{1}{2}} \hat{K}_i \hat{X}^{-\frac{1}{2}}$ by $X^{-\frac{1}{2}} K_i X^{-\frac{1}{2}}$, since the two are related by an orthogonal similarity. Similarly, the second inequality in (2.72) yields the following:

$$\begin{aligned}\lambda_{\max} \left(D^{-\frac{1}{2}} \left[\sum_{i=1}^m (N^{-1} \Delta b)_i \hat{A}_i \right] D^{-\frac{1}{2}} \right) = \\ \lambda_{\max} \left(\sum_{i=1}^m (N^{-1} \Delta b)_i \left(D^{-\frac{1}{2}} \hat{A}_i D^{-\frac{1}{2}} \right) \right) \leq 1.\end{aligned}\tag{2.74}$$

Again, $D^{-\frac{1}{2}} \hat{A}_i D^{-\frac{1}{2}}$ can be replaced by $S^{-\frac{1}{2}} A_i S^{-\frac{1}{2}}$ if desired. Summarizing, we have

Proposition 2.3.4 *Let (X, y, S) be a strictly feasible point for $SDP(b, C)$ and $SDD(b, C)$ such that $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$ is nonsingular. Assume that the right-hand side vector b is replaced by $b' := b + \Delta b$, where $\Delta b \in \mathbb{R}^m$. Suppose a Newton step is taken from (X, y, S) targeting the feasible point (X', y', S') of $SDP(b', C)$ and $SDD(b', C)$ that satisfies $(X'S' + S'X')/2 = (XS + SX)/2$. Then, if we use the AHO direction, a full Newton step can be taken and the resulting iterate will be feasible for the new problems iff Δb satisfies (2.73) and (2.74). Moreover, the duality gap of the new iterate will be at most $X \bullet S$ if $\mathcal{E}^{-1}\mathcal{F}$ is positive definite.*

□

Next, we consider a change in the cost matrix C . If we use the scaled iterate $(\hat{X}, \hat{y}, \hat{S})$ again and the fact that $\widehat{\Delta C} = Q^T \Delta C Q$, Proposition 2.3.2 implies that the following bounds on $\widehat{\Delta C}$ are necessary and sufficient:

$$\begin{aligned} \lambda_{\max} \left(\hat{X}^{-\frac{1}{2}} \left[\hat{\mathcal{E}}^{-1} \hat{\mathcal{F}} \widehat{\Delta C} - \hat{\mathcal{E}}^{-1} \hat{\mathcal{F}} \hat{\mathcal{A}}^* N^{-1} \hat{\mathcal{A}} \hat{\mathcal{E}}^{-1} \hat{\mathcal{F}} \widehat{\Delta C} \right] \hat{X}^{-\frac{1}{2}} \right) &\leq 1, \\ \lambda_{\min} \left(\hat{S}^{-\frac{1}{2}} \left[\widehat{\Delta C} - \hat{\mathcal{A}}^* N^{-1} \hat{\mathcal{A}} \hat{\mathcal{E}}^{-1} \hat{\mathcal{F}} \widehat{\Delta C} \right] \hat{S}^{-\frac{1}{2}} \right) &\geq -1. \end{aligned} \quad (2.75)$$

Let $\hat{L} = \hat{\mathcal{E}}^{-1} \hat{\mathcal{F}} \widehat{\Delta C} = \hat{\mathcal{E}}^{-1} \left(\frac{1}{2} (\widehat{\Delta C} \hat{X} + \hat{X} \widehat{\Delta C}) \right)$. Then, using (2.68), we get $\hat{L}_{ij} = \frac{(\widehat{\Delta C} \hat{X} + \hat{X} \widehat{\Delta C})_{ij}}{d_i + d_j}$. Hence, the first inequality in (2.75) simplifies to

$$\begin{aligned} \lambda_{\max} \left(\hat{X}^{-\frac{1}{2}} \left(\hat{L} - \hat{\mathcal{E}}^{-1} \hat{\mathcal{F}} \hat{\mathcal{A}}^* (N^{-1} v) \right) \hat{X}^{-\frac{1}{2}} \right) &= \\ \lambda_{\max} \left(\hat{X}^{-\frac{1}{2}} \left(\hat{L} - \hat{\mathcal{E}}^{-1} \left[\frac{1}{2} \sum_{i=1}^m (N^{-1} v)_i (\hat{A}_i \hat{X} + \hat{X} \hat{A}_i) \right] \right) \hat{X}^{-\frac{1}{2}} \right) &= \\ \lambda_{\max} \left(\hat{X}^{-\frac{1}{2}} \left(\hat{L} - \sum_{i=1}^m (N^{-1} v)_i \hat{K}_i \right) \hat{X}^{-\frac{1}{2}} \right) &\leq 1, \end{aligned} \quad (2.76)$$

where \hat{K}_i is again given by (2.70) and

$$v = \begin{bmatrix} \hat{L} \bullet \hat{A}_1 \\ \vdots \\ \hat{L} \bullet \hat{A}_m \end{bmatrix}.$$

Similarly, the second inequality in (2.75) yields:

$$\begin{aligned} \lambda_{\min} \left(D^{-\frac{1}{2}} \left[\widehat{\Delta C} - \hat{\mathcal{A}}^* \left((N^{-1}) \hat{\mathcal{A}} \hat{\mathcal{E}}^{-1} \hat{\mathcal{F}} \widehat{\Delta C} \right) \right] D^{-\frac{1}{2}} \right) = \\ \lambda_{\min} \left(D^{-\frac{1}{2}} \left[\widehat{\Delta C} - \sum_{i=1}^m (N^{-1}v)_i \hat{A}_i \right] D^{-\frac{1}{2}} \right) \geq -1. \end{aligned} \quad (2.77)$$

(As before, we can express these bounds in terms of unscaled quantities using K_i as above and $L := Q \hat{L} Q^T$.) Hence, we obtain

Proposition 2.3.5 *Let (X, y, S) be a strictly feasible point for $SDP(b, C)$ and $SDD(b, C)$ such that $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$ is nonsingular. Assume that the cost matrix C is replaced by $C' := C + \Delta C$, where $\Delta C \in \mathcal{S}^n$. Suppose a Newton step is taken from (X, y, S) targeting the feasible point (X', y', S') of $SDP(b, C')$ and $SDD(b, C')$ that satisfies $(X'S' + S'X')/2 = (XS + SX)/2$. Then, if we use the AHO direction, a full Newton step can be taken and the resulting iterate will be feasible for the new problems iff $\widehat{\Delta C} = Q^T \Delta C Q$ satisfies (2.76) and (2.77). Moreover, the duality gap of the new iterate will be at most $X \bullet S$ if $\mathcal{E}^{-1}\mathcal{F}$ is positive definite.*

□

2.3.3 The H..K..M Direction

The H..K..M direction [29, 39, 42] is the Newton step for the following symmetrization of the third equation in (2.34):

$$\Theta(\tilde{X}, \tilde{S}) := \frac{1}{2}(S\tilde{X}\tilde{S} + \tilde{S}\tilde{X}S) = \frac{1}{2}(SX'S' + S'X'S). \quad (2.78)$$

Here, the operators \mathcal{E} and \mathcal{F} are given by

$$\mathcal{E} = S \odot S, \quad \mathcal{F} = SX \odot I, \quad R_{EF} = \frac{1}{2}(SX'S' + S'X'S) - SXS. \quad (2.79)$$

Therefore, $M = S$ for the H..K..M direction. Alternatively, so that \mathcal{E} does not need to be inverted, we have:

$$\mathcal{E} = I \odot I, \quad \mathcal{F} = X \odot S^{-1}, \quad R_{EF} = \frac{1}{2}(X'S'S^{-1} + S^{-1}S'X') - X. \quad (2.80)$$

Note that the H..K..M direction is P -scale invariant. Therefore, we apply the scaling transformation using $P = S^{\frac{1}{2}}$. Then we have the following scaled matrices:

$$\begin{aligned} \hat{X} &= S^{\frac{1}{2}}XS^{\frac{1}{2}}, \\ \hat{S} &= I, \\ \hat{A}_i &= S^{-\frac{1}{2}}A_iS^{-\frac{1}{2}}, \\ \hat{C} &= S^{-\frac{1}{2}}CS^{-\frac{1}{2}}. \end{aligned} \quad (2.81)$$

We also have $\hat{\mathcal{E}} = \mathcal{E} = I \odot I$ and $\hat{\mathcal{F}} = \hat{X} \odot I$.

Now we compute the Schur complement matrix for the current iterate (X, y, S) using the scaled iterate. Let $N = \hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{\mathcal{A}}^*$. Then the i th column of N is given by:

$$Ne_i = (\hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{\mathcal{A}}^*)e_i = \hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{A}_i = \hat{\mathcal{A}} \left[\frac{1}{2}(\hat{A}_i\hat{X} + \hat{X}\hat{A}_i) \right]. \quad (2.82)$$

Therefore,

$$Ne_i = \frac{1}{2} \begin{bmatrix} \hat{A}_1 \bullet (\hat{A}_i \hat{X} + \hat{X} \hat{A}_i) \\ \vdots \\ \hat{A}_m \bullet (\hat{A}_i \hat{X} + \hat{X} \hat{A}_i) \end{bmatrix} = \begin{bmatrix} \text{Trace}(\hat{A}_1 \hat{X} \hat{A}_i) \\ \vdots \\ \text{Trace}(\hat{A}_m \hat{X} \hat{A}_i) \end{bmatrix}. \quad (2.83)$$

In the above derivation, we used the obvious facts that $\text{Trace}(A) = \text{Trace}(A^T)$ and $\text{Trace}(PK) = \text{Trace}(KP)$, for any square matrices A , P , and K . (2.83) implies that the Schur complement matrix N is always symmetric (see also Proposition 6.4 in [60]).

The H..K..M direction is a well-defined direction, i.e., it exists and is unique for every symmetric positive definite X and S and every surjective \mathcal{A} . Moreover, the operator $\mathcal{E}^{-1}\mathcal{F}$ is self-adjoint and positive definite (see Proposition 6.3 in [60]).

First, we consider a change in the right-hand side vector b . If we use the scaled iterate $(\hat{X}, \hat{y}, \hat{S})$ and the fact that $\widehat{\Delta b} = \Delta b$, Proposition 2.3.1 implies that the following bounds on Δb are necessary and sufficient:

$$\begin{aligned} \lambda_{\min} \left(\hat{X}^{-\frac{1}{2}} \left[\hat{\mathcal{E}}^{-1} \hat{\mathcal{F}} \hat{\mathcal{A}}^*(N^{-1} \Delta b) \right] \hat{X}^{-\frac{1}{2}} \right) &\geq -1, \\ \lambda_{\max} \left(\hat{S}^{-\frac{1}{2}} \left[\hat{\mathcal{A}}^*(N^{-1} \Delta b) \right] \hat{S}^{-\frac{1}{2}} \right) &\leq 1. \end{aligned} \quad (2.84)$$

The first inequality in (2.84) yields the following:

$$\begin{aligned} \lambda_{\min} \left(\hat{X}^{-\frac{1}{2}} \left[\frac{1}{2} \sum_{i=1}^m (N^{-1} \Delta b)_i (\hat{A}_i \hat{X} + \hat{X} \hat{A}_i) \right] \hat{X}^{-\frac{1}{2}} \right) = \\ \lambda_{\min} \left(\frac{1}{2} \sum_{i=1}^m (N^{-1} \Delta b)_i \left[\hat{X}^{-\frac{1}{2}} \hat{A}_i \hat{X}^{\frac{1}{2}} + \hat{X}^{\frac{1}{2}} \hat{A}_i \hat{X}^{-\frac{1}{2}} \right] \right) \geq -1. \end{aligned} \quad (2.85)$$

Similarly, the second inequality in (2.84) yields the following:

$$\lambda_{\max} \left(\sum_{i=1}^m (N^{-1} \Delta b)_i \hat{A}_i \right) \leq 1. \quad (2.86)$$

Note that (2.86) bounds the maximum eigenvalue of $E := \sum_{i=1}^m (N^{-1}\Delta b)_i \hat{A}_i$, while (2.85) bounds the minimum eigenvalue of $(U^{-1}EU + UEU^{-1})/2$ for $U := \hat{X}^{\frac{1}{2}}$. By Lemma 3.3 of Monteiro [42], the former eigenvalue is always bounded by the maximum eigenvalue of $(U^{-1}EU + UEU^{-1})/2$. Thus, we have

Proposition 2.3.6 *Let (X, y, S) be a strictly feasible point for $SDP(b, C)$ and $SDD(b, C)$. Assume that the right-hand side vector b is replaced by $b' := b + \Delta b$, where $\Delta b \in \mathbb{R}^m$ and a Newton step is taken from (X, y, S) targeting the feasible point (X', y', S') of $SDP(b', C)$ and $SDD(b', C)$ that satisfies $\frac{1}{2}(SX'S' + S'X'S) = XS$. Then, if we use the H..K..M direction, a full Newton step can be taken and the resulting iterate will be feasible for the new problems iff Δb satisfies (2.85) and (2.86).*

A sufficient condition is that

$$\left\| \frac{1}{2} \sum_{i=1}^m (N^{-1}\Delta b)_i \left[\hat{X}^{-\frac{1}{2}} \hat{A}_i \hat{X}^{\frac{1}{2}} + \hat{X}^{\frac{1}{2}} \hat{A}_i \hat{X}^{-\frac{1}{2}} \right] \right\|_2 \leq 1.$$

Moreover, the duality gap of the new iterate will be at most $X \bullet S$.

□

(It is easy to see that the sufficient condition in the theorem is in fact necessary and sufficient for both perturbations Δb and $-\Delta b$ to yield feasible full Newton steps.)

Next, we consider a change in the cost matrix C . If we use the scaled iterate $(\hat{X}, \hat{y}, \hat{S})$ again and the fact that $\widehat{\Delta C} = S^{-\frac{1}{2}} \Delta C S^{-\frac{1}{2}}$, Proposition 2.3.2 implies that the following bounds on $\widehat{\Delta C}$ are necessary and sufficient:

$$\begin{aligned} \lambda_{\max} \left(\hat{X}^{-\frac{1}{2}} \left[\hat{\mathcal{E}}^{-1} \hat{\mathcal{F}} \widehat{\Delta C} - \hat{\mathcal{E}}^{-1} \hat{\mathcal{F}} \hat{A}^* N^{-1} \hat{A} \hat{\mathcal{E}}^{-1} \hat{\mathcal{F}} \widehat{\Delta C} \right] \hat{X}^{-\frac{1}{2}} \right) &\leq 1, \\ \lambda_{\min} \left(\hat{S}^{-\frac{1}{2}} \left[\widehat{\Delta C} - \hat{A}^* N^{-1} \hat{A} \hat{\mathcal{E}}^{-1} \hat{\mathcal{F}} \widehat{\Delta C} \right] \hat{S}^{-\frac{1}{2}} \right) &\geq -1. \end{aligned} \tag{2.87}$$

The first inequality in (2.87) yields the following:

$$\begin{aligned}
& \lambda_{\max} \left(\hat{X}^{-\frac{1}{2}} \left[\frac{1}{2}(\widehat{\Delta C} \hat{X} + \hat{X} \widehat{\Delta C}) - \hat{\mathcal{F}} \hat{\mathcal{A}}^* \left(N^{-1} \hat{\mathcal{A}} \left[\frac{1}{2}(\widehat{\Delta C} \hat{X} + \hat{X} \widehat{\Delta C}) \right] \right) \right] \hat{X}^{-\frac{1}{2}} \right) = \\
& \lambda_{\max} \left(\hat{X}^{-\frac{1}{2}} \left[\frac{1}{2}(\widehat{\Delta C} \hat{X} + \hat{X} \widehat{\Delta C}) - \frac{1}{2} \sum_{i=1}^m (N^{-1} v)_i (\hat{A}_i \hat{X} + \hat{X} \hat{A}_i) \right] \hat{X}^{-\frac{1}{2}} \right) = \\
& \frac{1}{2} \lambda_{\max} \left(\hat{X}^{-\frac{1}{2}} \widehat{\Delta C} \hat{X}^{\frac{1}{2}} + \hat{X}^{\frac{1}{2}} \widehat{\Delta C} \hat{X}^{-\frac{1}{2}} - \sum_{i=1}^m (N^{-1} v)_i \left(\hat{X}^{-\frac{1}{2}} \hat{A}_i \hat{X}^{\frac{1}{2}} + \hat{X}^{\frac{1}{2}} \hat{A}_i \hat{X}^{-\frac{1}{2}} \right) \right) \leq 1,
\end{aligned} \tag{2.88}$$

where $v_i = \hat{A}_i \bullet \widehat{\Delta C} \hat{X}$.

Similarly, the second inequality in (2.87) yields:

$$\lambda_{\min} \left(\widehat{\Delta C} - \hat{\mathcal{A}}^* \left(N^{-1} \hat{\mathcal{A}} \widehat{\Delta C} \right) \right) = \lambda_{\min} \left(\widehat{\Delta C} - \sum_{i=1}^m (N^{-1} v)_i \hat{A}_i \right) \geq -1, \tag{2.89}$$

where v is the same as above. Note that (2.89) bounds the minimum eigenvalue of $E := \widehat{\Delta C} - \sum_{i=1}^m (N^{-1} v)_i \hat{A}_i$, while (2.88) bounds the maximum eigenvalue of $(U^{-1} E U + U E U^{-1})/2$ for $U := \hat{X}^{\frac{1}{2}}$. Again, the results of Monteiro [42] show that the former is at most the minimum eigenvalue of $(U^{-1} E U + U E U^{-1})/2$. Thus, we have shown

Proposition 2.3.7 *Let (X, y, S) be a strictly feasible point for $SDP(b, C)$ and $SDD(b, C)$. Assume that the cost matrix C is replaced by $C' := C + \Delta C$, where $\Delta C \in \mathcal{S}^n$ and a Newton step is taken from (X, y, S) targeting the feasible point (X', y', S') of $SDP(b, C')$ and $SDD(b, C')$ that satisfies $X' S' = X S$. Then, if we use the H..K..M direction, a full Newton step can be taken and the resulting iterate will be feasible for the new problems iff $\widehat{\Delta C} = S^{-\frac{1}{2}} \Delta C S^{-\frac{1}{2}}$ satisfies (2.88) and*

(2.89). A sufficient condition is that

$$\left\| \frac{1}{2} \left(\hat{X}^{-\frac{1}{2}} \widehat{\Delta C} \hat{X}^{\frac{1}{2}} + \hat{X}^{\frac{1}{2}} \widehat{\Delta C} \hat{X}^{-\frac{1}{2}} \right) - \frac{1}{2} \sum_{i=1}^m (N^{-1}v)_i \left(\hat{X}^{-\frac{1}{2}} \hat{A}_i \hat{X}^{\frac{1}{2}} + \hat{X}^{\frac{1}{2}} \hat{A}_i \hat{X}^{-\frac{1}{2}} \right) \right\|_2 \leq 1.$$

Moreover, the duality gap of the new iterate will be at most $X \bullet S$.

□

2.3.4 The NT Direction

The NT direction [48, 49] is the Newton step for the following symmetrization of the third equation in (2.34):

$$\Theta(\tilde{X}, \tilde{S}) := \frac{1}{2} (W^{-1} \tilde{X} \tilde{S} + \tilde{S} \tilde{X} W^{-1}) = \frac{1}{2} (W^{-1} X' S' + S' X' W^{-1}), \quad (2.90)$$

where W is the scaling matrix defined by $W = X^{\frac{1}{2}} (X^{\frac{1}{2}} S X^{\frac{1}{2}})^{-\frac{1}{2}} X^{\frac{1}{2}}$ so that $W S W = X$. Here, the operators \mathcal{E} and \mathcal{F} are given by

$$\mathcal{E} = S \odot W^{-1}, \quad \mathcal{F} = W^{-1} X \odot I,$$

$$R_{EF} = \frac{1}{2} \left((W^{-1} X' S' + S' X' W^{-1}) - (W^{-1} X S + S X W^{-1}) \right). \quad (2.91)$$

Therefore, $M = W^{-1}$ for the NT direction. It has been shown [62] that if the targeted point satisfies $X' S' = \nu I$ for some $\nu > 0$, then the NT direction can alternatively be defined in the following convenient way, in which case \mathcal{E} does not need to be inverted:

$$\mathcal{E} = I \odot I, \quad \mathcal{F} = W \odot W, \quad R_{EF} = \nu S^{-1} - X. \quad (2.92)$$

We claim that the representation (2.92) can be generalized to the case when X' and S' are arbitrary matrices if R_{EF} is appropriately chosen. First of all, note that (2.91) implies that the third equation of (2.35) is given by

$$(S\Delta X + \Delta SX)W^{-1} + W^{-1}(\Delta XS + X\Delta S) = W^{-1}(X'S' - XS) + (S'X' - SX)W^{-1}. \quad (2.93)$$

Postmultiplying (2.93) by W , we obtain

$$S\Delta X + \Delta SX + W^{-1}(\Delta XS + X\Delta S)W = W^{-1}(X'S' - XS)W + S'X' - SX. \quad (2.94)$$

Now we will show that there exists a unique symmetric matrix R_{EF} such that the following system

$$\Delta X + W\Delta SW = R_{EF}, \quad (2.95)$$

which is related to (2.92) is equivalent to (2.94) for arbitrary X' and S' . Proceeding as in [62], we see that (2.95) is equivalent to each of the following two equations:

$$W^{-1}\Delta XSW + \Delta SX = W^{-1}R_{EF}SW, \quad (2.96)$$

$$S\Delta X + W^{-1}X\Delta SW = SR_{EF}. \quad (2.97)$$

The first equality follows from premultiplying (2.95) by W^{-1} and postmultiplying by $W^{-1}X$, together with $SW = W^{-1}X$, and the second equality is a consequence of premultiplying (2.95) by S . Adding up (2.96) and (2.97), we obtain the same expression on the left hand side of (2.94). Hence, it follows that R_{EF} should satisfy

$$W^{-1}R_{EF}SW + SR_{EF} = W^{-1}(X'S' - XS)W + S'X' - SX. \quad (2.98)$$

Postmultiplying (2.98) by W^{-1} , we obtain

$$W^{-1}R_{EF}S + SR_{EF}W^{-1} = W^{-1}(X'S' - XS) + (S'X' - SX)W^{-1}. \quad (2.99)$$

Using the notations in (2.36) and (2.38), (2.99) is equivalent to

$$(W^{-1} \odot S)R_{EF} = (W^{-1} \odot I)[X'S' - XS]. \quad (2.100)$$

On the left, $W^{-1} \odot S$ is viewed as an operator from \mathcal{S}^n to itself, while on the right $W^{-1} \odot I$ takes $\mathbb{R}^{n \times n}$ to \mathcal{S}^n . However, note that $W^{-1} \odot S$ is positive definite, and therefore (2.100) has a unique solution R_{EF} . This shows that the solution to (2.95) will also satisfy (2.93) if R_{EF} is given by (2.100). However, the solutions to the two Newton systems (2.35) where the third equations are given by (2.93) and (2.95) respectively exist and are unique, and hence must agree. Therefore, we conclude that an alternative representation of the NT direction is given by

$$\mathcal{E} = I \odot I, \quad \mathcal{F} = W \odot W, \quad R_{EF}, \quad (2.101)$$

where R_{EF} is defined as the solution to (2.100). Observe that computation of R_{EF} involves solving a Lyapunov system. However, for our purposes, the right hand side of (2.100) is exactly equal to $\Theta(X', S') - \Theta(X, S)$, and in our case this is 0; therefore $R_{EF} = 0$.

Note that the NT direction is also P -scale invariant. Therefore, we apply the scaling transformation using $P = W^{-\frac{1}{2}}$. Then we have the following scaled matrices:

$$\begin{aligned} \hat{X} &= W^{-\frac{1}{2}} X W^{-\frac{1}{2}}, \\ \hat{S} &= W^{\frac{1}{2}} S W^{\frac{1}{2}}, \\ \hat{A}_i &= W^{\frac{1}{2}} A_i W^{\frac{1}{2}}, \\ \hat{C} &= W^{\frac{1}{2}} C W^{\frac{1}{2}}. \end{aligned} \quad (2.102)$$

Hence $\hat{X} = \hat{S}$ and so $\hat{W} = I$. Thus we have $\hat{\mathcal{E}} = \hat{\mathcal{F}} = I \odot I$.

Now we compute the Schur complement matrix for the current iterate (X, y, S) using the scaled iterate. Let $N = \hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{\mathcal{A}}^*$. Then the i th column of N is given by:

$$\begin{aligned} Ne_i &= (\hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{\mathcal{A}}^*)e_i = \hat{\mathcal{A}}\hat{A}_i \\ &= \begin{bmatrix} \hat{A}_1 \bullet \hat{A}_i \\ \vdots \\ \hat{A}_m \bullet \hat{A}_i \end{bmatrix}. \end{aligned} \quad (2.103)$$

Thus the Schur complement matrix N is always symmetric (see also Proposition 6.4 in [60]).

The NT direction is a well-defined direction, i.e., it exists and is unique for every symmetric positive definite X and S and every surjective \mathcal{A} . Moreover, the operator $\mathcal{E}^{-1}\mathcal{F}$ is self-adjoint and positive definite (see Proposition 6.3 in [60]).

First we consider a change in the right-hand side vector b . If we use the scaled iterate $(\hat{X}, \hat{y}, \hat{S})$ and the fact that $\widehat{\Delta b} = \Delta b$, Proposition 2.3.1 implies that the following bounds on Δb are necessary and sufficient:

$$\begin{aligned} \lambda_{\min} \left(\hat{X}^{-\frac{1}{2}} \left[\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{\mathcal{A}}^*(N^{-1}\Delta b) \right] \hat{X}^{-\frac{1}{2}} \right) &\geq -1, \\ \lambda_{\max} \left(\hat{S}^{-\frac{1}{2}} \left[\hat{\mathcal{A}}^*(N^{-1}\Delta b) \right] \hat{S}^{-\frac{1}{2}} \right) &\leq 1. \end{aligned} \quad (2.104)$$

The first inequality in (2.104) yields the following:

$$\lambda_{\min} \left(\sum_{i=1}^m (N^{-1}\Delta b)_i \hat{X}^{-\frac{1}{2}} \hat{A}_i \hat{X}^{-\frac{1}{2}} \right) \geq -1. \quad (2.105)$$

Similarly, since $\hat{X} = \hat{S}$, the second inequality in (2.104) requires that the maximum eigenvalue of the same matrix be at most 1. Therefore, unlike the situation for the other directions, both bounds relate to the same matrix, and our necessary and

sufficient condition simplifies, becoming similar to the sufficient condition involving the 2-norm for the H..K..M direction or the necessary and sufficient condition involving the L_∞ -norm of a vector for LP. We state this nice result about the NT direction as:

Proposition 2.3.8 *Let (X, y, S) be a strictly feasible point for $SDP(b, C)$ and $SDD(b, C)$. Assume that the right-hand side vector b is replaced by $b' := b + \Delta b$, where $\Delta b \in \mathbb{R}^m$. Suppose a Newton step is taken from (X, y, S) targeting the feasible point (X', y', S') of $SDP(b', C)$ and $SDD(b', C)$ that satisfies $\Theta(X', S') = \Theta(X, S)$. Then, if we use the NT direction, a full Newton step can be taken and the resulting iterate will be feasible for the new problems iff Δb satisfies*

$$\left\| \sum_{i=1}^m (N^{-1} \Delta b)_i \hat{X}^{-\frac{1}{2}} \hat{A}_i \hat{X}^{-\frac{1}{2}} \right\|_2 \leq 1.$$

Moreover, the duality gap of the new iterate will be at most $X \bullet S$.

□

Next, we consider a change in the cost matrix C . If we use the scaled iterate $(\hat{X}, \hat{y}, \hat{S})$ again and the fact that $\widehat{\Delta C} = W^{\frac{1}{2}} \Delta C W^{\frac{1}{2}}$, Proposition 2.3.2 implies that the following bounds on $\widehat{\Delta C}$ are necessary and sufficient:

$$\begin{aligned} \lambda_{\max} \left(\hat{X}^{-\frac{1}{2}} \left[\hat{\mathcal{E}}^{-1} \hat{\mathcal{F}} \widehat{\Delta C} - \hat{\mathcal{E}}^{-1} \hat{\mathcal{F}} \hat{A}^* N^{-1} \hat{A} \hat{\mathcal{E}}^{-1} \hat{\mathcal{F}} \widehat{\Delta C} \right] \hat{X}^{-\frac{1}{2}} \right) &\leq 1, \\ \lambda_{\min} \left(\hat{S}^{-\frac{1}{2}} \left[\widehat{\Delta C} - \hat{A}^* N^{-1} \hat{A} \hat{\mathcal{E}}^{-1} \hat{\mathcal{F}} \widehat{\Delta C} \right] \hat{S}^{-\frac{1}{2}} \right) &\geq -1. \end{aligned} \tag{2.106}$$

The first inequality in (2.106) yields the following:

$$\lambda_{\max} \left(\hat{X}^{-\frac{1}{2}} \left[\widehat{\Delta C} - \sum_{i=1}^m (N^{-1} v)_i \hat{A}_i \right] \hat{X}^{-\frac{1}{2}} \right) =$$

$$\lambda_{\max} \left(\hat{X}^{-\frac{1}{2}} \widehat{\Delta C} \hat{X}^{-\frac{1}{2}} - \sum_{i=1}^m (N^{-1}v)_i \hat{X}^{-\frac{1}{2}} \hat{A}_i \hat{X}^{-\frac{1}{2}} \right) \leq 1, \quad (2.107)$$

where $v_i = \hat{A}_i \bullet \widehat{\Delta C}$.

Similarly, the second inequality in (2.106) requires that the minimum eigenvalue of the same matrix be at most -1 . Once again, the two bounds involve the same matrix, and so can be written concisely as a bound on its 2-norm. We state this as:

Proposition 2.3.9 *Let (X, y, S) be a strictly feasible point for $SDP(b, C)$ and $SDD(b, C)$. Assume that the cost matrix C is replaced by $C' := C + \Delta C$, where $\Delta C \in \mathcal{S}^n$. Suppose that a Newton step is taken from (X, y, S) targeting the feasible point (X', y', S') of $SDP(b, C')$ and $SDD(b, C')$ that satisfies $\Theta(X', S') = \Theta(X, S)$. Then, if we use the NT direction, a full Newton step can be taken and the resulting iterate will be feasible for the new problems iff $\widehat{\Delta C} = W^{\frac{1}{2}} \Delta C W^{\frac{1}{2}}$ satisfies*

$$\left\| \hat{X}^{-\frac{1}{2}} \left[\widehat{\Delta C} - \sum_{i=1}^m (N^{-1}v)_i \hat{A}_i \right] \hat{X}^{-\frac{1}{2}} \right\|_2 \leq 1.$$

Moreover, the duality gap of the new iterate will be at most $X \bullet S$.

□

2.3.5 Comparison with LP

As mentioned before, LP is a special case of SDP where all the matrices A_i , C , and hence X and S , are restricted to be diagonal. If the LP is given in the standard form, then A_i is the diagonal matrix corresponding to the i th row of the coefficient matrix A in LP, C is the diagonal matrix whose components are given by the cost

vector c , and X and S are the diagonal matrices similarly obtained from x and s , respectively.

In this section, given an LP, we analyze the relationship between the LP bounds (2.4) and (2.11) and their counterparts resulting from the three directions for the corresponding SDP. The analysis will not refer to any of the three directions specifically, but we will only assume that the operators \mathcal{E} and \mathcal{F} are given by (2.40) and that the matrix M is also diagonal whenever X and S are diagonal. This property holds for all three directions as well as the so-called dual H..K..M direction [39, 42], which uses $M := X^{-1}$.

Since the matrices defining the operators \mathcal{E} and \mathcal{F} are diagonal and since diagonal matrices commute, some of the computations in the previous sections can be significantly simplified. In particular, if Σ is diagonal, then $\mathcal{E}\Sigma = SM\Sigma$, $\mathcal{F}\Sigma = MX\Sigma$. Therefore, $\mathcal{E}^{-1}\mathcal{F}\Sigma = S^{-1}X\Sigma$. By this observation, the i th column of the Schur complement matrix $N = \mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$ is then given by

$$\begin{aligned}
 Ne_i &= \mathcal{A}\mathcal{E}^{-1}\mathcal{F}A_i = \mathcal{A}S^{-1}XA_i \implies \\
 Ne_i &= \begin{bmatrix} \text{Trace}(A_1S^{-1}XA_i) \\ \vdots \\ \text{Trace}(A_mS^{-1}XA_i) \end{bmatrix}. \tag{2.108}
 \end{aligned}$$

However, (2.108) implies that $N = AD^2A^T$, where $D^2 = XS^{-1}$, which is exactly the Schur complement matrix in LP.

Let us first focus on perturbations of b . The bounds (2.42) and (2.43) arising from Proposition 2.3.1 can be simplified using the fact that all operations yield diagonal matrices and that diagonal matrices commute. In particular, (2.42) is

equivalent to

$$\begin{aligned} \lambda_{\min} \left(X^{-\frac{1}{2}} (\mathcal{E}^{-1} \mathcal{F} \mathcal{A}^* [(\mathcal{A} \mathcal{E}^{-1} \mathcal{F} \mathcal{A}^*)^{-1} \Delta b]) X^{-\frac{1}{2}} \right) &= \\ \lambda_{\min} (S^{-1} \mathcal{A}^* [(\mathcal{A} \mathcal{E}^{-1} \mathcal{F} \mathcal{A}^*)^{-1} \Delta b]) &\geq -1, \end{aligned} \quad (2.109)$$

which can be combined with (2.43) to yield the following norm bound:

$$\begin{aligned} \|S^{-1} \mathcal{A}^* [(\mathcal{A} \mathcal{E}^{-1} \mathcal{F} \mathcal{A}^*)^{-1} \Delta b]\|_2 &= \\ \left\| \sum_{i=1}^m ((AD^2 A^T)^{-1} \Delta b)_i A_i S^{-1} \right\|_2 &= \\ \|S^{-1} A^T (AD^2 A^T)^{-1} \Delta b\|_\infty &\leq 1, \end{aligned} \quad (2.110)$$

where the last equality follows from the fact that the L_2 operator norm of a diagonal matrix is the same as the L_∞ norm of the vector of its diagonal entries. Therefore, the SDP bounds for Δb reduce exactly to the bound given in Proposition 2.2.1.

We now consider perturbations of C . In a similar manner, the bounds (2.54) and (2.55) arising from Proposition 2.3.2 can be combined into a single norm bound given by

$$\|S^{-1} (\Delta C - \mathcal{A}^* (\mathcal{A} \mathcal{E}^{-1} \mathcal{F} \mathcal{A}^*)^{-1} \mathcal{A} \mathcal{E}^{-1} \mathcal{F} \Delta C)\|_2 \leq 1. \quad (2.111)$$

Using our previous observations, $\mathcal{A} \mathcal{E}^{-1} \mathcal{F} \Delta C = AD^2 \Delta c$, where Δc is the vector obtained from the diagonal entries of ΔC . Therefore, (2.111) can be rewritten as

$$\begin{aligned} \|S^{-1} \Delta C - \sum_{i=1}^m ((AD^2 A^T)^{-1} AD^2 \Delta c)_i A_i S^{-1}\|_2 &= \\ \|S^{-1} (I - A^T (AD^2 A^T)^{-1} AD^2) \Delta c\|_\infty &\leq 1, \end{aligned} \quad (2.112)$$

which again is the same as the bound given in Proposition 2.2.2.

Therefore, we have proved the following proposition:

Proposition 2.3.10 *Given an LP, the interior-point bounds for all three directions for the corresponding SDP are exactly the same as those for the original LP for perturbations of b and c .*

2.4 Discussion

In this paper, we have analyzed perturbations of the right-hand side and the cost parameters in LP and SDP and presented tight bounds on the perturbations so that the result of a single interior-point iteration would yield feasible solutions to the perturbed problem and its dual. For the LP case where the solution is unique and nondegenerate, we showed that the bounds arising from the interior-point method asymptotically coincide with those from the optimal basis after symmetrizing with respect to the origin. Moreover, as long as the perturbations are within the bounds, one interior-point iteration at a strictly feasible point for the original problem and its dual results in a feasible point for the perturbed problem and its dual with a duality gap no greater than that of the original iterates.

Under the assumption of a unique and nondegenerate solution in LP, the optimal partition coincides with the basis partition. This no longer holds under degeneracy, however, since the optimal basis is not unique. Therefore, the bounds obtained from the simplex approach depend on the basis being used and a direct analysis as in the nondegenerate case is not very meaningful. In order to overcome this shortcoming of the basis approach, an optimal partition perspective has been developed [2, 32] and shown to yield more accurate information on sensitivity. Consequently, in the presence of degeneracy, the optimal partition-based bounds seem to be a natural basis for comparison with the interior-point bounds. However, the analysis is considerably more complex and uses different tools: it will be the subject of a subsequent paper (Chapter 3). However, it is worthwhile to note that the bounds

resulting from our interior-point approach still apply regardless of degeneracy as long as both primal and dual LPs have strictly feasible solutions.

For the SDP case, the analysis gets harder and the conditions on the perturbations for feasibility to be regained in one step more complicated to state, involving eigenvalue bounds on two different matrices, except for the simpler NT case. However, all three of our search directions yield the same bounds (and the same as given by the LP interior-point approach) in the case that an LP is cast as an SDP. Since an optimal solution for an SDP does not typically resemble a basic feasible solution, a comparison as in the LP case is not possible. However, in [18], Goldfarb and Scheinberg extend the optimal partition approach to sensitivity analysis in SDP. The resulting bounds can therefore be used for comparison with the interior-point bounds. Furthermore, from the theoretical results, it is not clear as to how the bounds resulting from the three search directions compare with one another in practice. We intend to study such questions in future research (see Chapters 4 and 5).

Chapter 3

An Interior-Point Approach to Sensitivity Analysis in Degenerate Linear Programs

3.1 Introduction

Sensitivity analysis (or post-optimality analysis) is the study of how the optimal solution of an optimization problem changes with respect to the changes in the problem data. The possible presence of errors in the problem data often makes sensitivity analysis as important as solving the original problem itself.

In the context of linear programming (LP), sensitivity analysis can be performed using an optimal basis approach (as in the simplex method) or an optimal partition approach, where the optimal partition refers to knowing, for each index, whether the

corresponding component of an optimal primal solution or of an optimal dual slack vector can be positive. The latter approach has close connections with interior-point methods since such methods, when properly terminated, provide an optimal solution in the relative interior of the optimal face, from which the optimal partition is readily available. In fact, as will shortly be discussed in detail, the optimal partition approach has been developed by Adler and Monteiro [2] and Jansen, de Jong, Roos and Terlaky [32] as a promising alternative in order to circumvent the drawbacks of the classical optimal basis approach in the presence of degeneracy. Later, Monteiro and Mehrotra [44] extended this approach by relaxing the requirement that the optimal partition be known. They also provided two methods to estimate the range of perturbations, each of which can be performed at any optimal solution, regardless of where it lies on the optimal face. More recently, Greenberg, Holder, Roos and Terlaky [26] related the dimension of the optimal set to the dimension of the set of objective perturbations for which the optimal partition is invariant. Greenberg [25] considered the simultaneous perturbations of the right-hand side and the cost vectors from an optimal partition perspective.

Recently, we studied perturbations of the right-hand side and the cost parameters in linear programming [71] (see Chapter 2), motivated by how interior-point methods from a near-optimal pair of strictly feasible solutions for a problem and its dual would compare with the optimal basis approach obtained from a nondegenerate optimal basic solution for such perturbations. The proposed interior-point perspective stems from the objectives of regaining feasibility and maintaining near-optimality in a *single* iteration of the interior-point method. This requires the setup

of the “right” Newton system among many possible choices in order to achieve both objectives simultaneously. Such a perspective provides a basis for the comparison of the interior-point and the simplex approaches to sensitivity analysis.

Under the assumption of a unique, nondegenerate optimal solution, the authors showed that the Newton system proposed in [71] is the “right” one in the sense that it yields asymptotically the same bounds on perturbations as those that keep the current basis optimal (after symmetrization with respect to the origin). Similar results, but changing only one of the primal or dual near-optimal solutions, were obtained by Kim, Park and Park [37].

However, most LPs arising from real-life problems are degenerate. Our goal in this paper is to investigate the quality of the bounds from the interior-point perspective in the absence of the strong assumption of nondegeneracy. This will lead to a complete analysis of the interior-point perspective proposed in [71] (Chapter 2). In doing so, we need something to compare our interior-point bounds with. In contrast to the nondegenerate case, the presence of multiple optimal bases makes a simplex-based approach unsuitable, as will be explained shortly. We therefore compare our bounds to those obtained from considering how much the right-hand side or the cost vector can change while maintaining the same optimal partition. Consequently, we use completely different tools for our analysis in this paper.

The next section is devoted to the preliminaries including the introduction of the tools relevant for the analysis as well as the restatement of our interior-point approach. Section 3.3 discusses the equivalence between the primal and dual formulations and shows that it suffices to consider perturbations of the right-hand side

only. We analyze the interior-point bounds under a special case of degeneracy in Section 3.4 and extend the analysis to the general degenerate case in Section 3.5. Section 3.6 concludes the paper with some remarks.

3.2 Preliminaries

We consider the LP in the following standard form:

$$\min_x c^T x, \text{ subject to } Ax = b, \quad x \geq 0. \quad (\text{P})$$

The associated dual LP is given by

$$\max_{y,s} b^T y, \text{ subject to } A^T y + s = c, \quad s \geq 0. \quad (\text{D})$$

Here, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$ constitute the data, and $(x, y, s) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ are the decision variables. Throughout this paper, the coefficient matrix A will be fixed and we will consider one-dimensional perturbations of the right-hand side vector b and the cost vector c , i.e., b will be replaced by $b + t\Delta b$ and c by $c + t\Delta c$, where Δb and Δc will be fixed in \mathbb{R}^m and \mathbb{R}^n , respectively, and $t \in \mathbb{R}$ will be the parameter. This is also called *parametric analysis* in the literature.

We will make the following assumptions:

1. The coefficient matrix A has full row rank.
2. Both (P) and (D) have strictly feasible solutions, i.e., there exist $x > 0$, $s > 0$ and y such that $Ax = b$ and $A^T y + s = c$.

The classical approach to sensitivity analysis has been based on the simplex method. Assuming that an optimal solution exists, the simplex method terminates with a basic optimal solution along with a corresponding basis. A natural criterion for the allowable perturbations in the data is then given by the following: how much perturbation in the data can one allow so that the current basis remains optimal for the perturbed LP?

Let us consider the parametric right-hand side (RHS) problem, i.e., let b be replaced by $b + t\Delta b$. Define $v(t) = \min\{c^T x : Ax = b + t\Delta b, x \geq 0\}$. It is well-known that v is a convex, piecewise linear, continuous function of t . The parametric RHS problem includes finding out all the “breakpoints” of $v(t)$.

Fixing a value of t , say at 0 for the purposes of this paper, the classical approach to sensitivity analysis then provides the set of values of t for which an optimal basis for $t = 0$ remains optimal for the resulting LPs parametrized by t . This is called the *optimality interval* associated with an optimal basis. Note that the optimal basis approach indeed yields the breakpoints of $v(t)$ around 0 under primal and dual nondegeneracy (which holds only if 0 itself is not a breakpoint of $v(t)$). However, the presence of primal and/or dual degeneracies is a shortcoming for this approach since, for example, multiple optimal bases might yield different optimality intervals. This shortcoming has been observed by several researchers. Adler and Monteiro [2], and Jansen, de Jong, Roos and Terlaky [32] developed an optimal partition approach to sensitivity analysis and showed that the optimality intervals associated with the optimal partitions uniquely and unambiguously identify the breakpoints of $v(t)$ and the intervals between the consecutive breakpoints. By the symmetry

between (P) and (D), which will be treated in more detail in Section 3.3, the same conclusions also hold for the parametric analysis of the cost vector c .

The idea of the optimal partition is based on a well-known result of Goldman and Tucker [19]. The optimality conditions for (P) and (D) are given by primal and dual feasibility and complementary slackness, that is, a triple (x, y, s) is optimal for (P) and (D) if and only if it satisfies

$$Ax = b, \quad A^T y + s = c, \quad x_i s_i = 0, \quad i = 1, \dots, n, \quad x \geq 0, \quad s \geq 0, \quad (3.1)$$

where x_i and s_i denote the i th components of x and s , respectively. Let Ω_P and Ω_D denote the set of optimal solutions for (P) and (D), respectively. Then, we can define two index sets as

$$\begin{aligned} \mathcal{B} &= \{j \in \{1, \dots, n\} : x_j > 0 \text{ for some } x \in \Omega_P\}, \\ \mathcal{N} &= \{j \in \{1, \dots, n\} : s_j > 0 \text{ for some } (y, s) \in \Omega_D\}. \end{aligned} \quad (3.2)$$

The optimality conditions (3.1) imply that $\mathcal{B} \cap \mathcal{N} = \emptyset$. The Goldman-Tucker result indicates that \mathcal{B} and \mathcal{N} actually partition the index set $\{1, \dots, n\}$, i.e., $\mathcal{B} \cup \mathcal{N} = \{1, \dots, n\}$. Therefore, there exist at least one primal solution $x \in \Omega_P$ and one dual solution $(y, s) \in \Omega_D$ such that $x + s > 0$. Such a solution will be called *strictly complementary* and $(\mathcal{B}, \mathcal{N})$ will be called the *optimal partition*. In contrast to the possibility of multiple optimal bases, the optimal partition is unique for a given LP instance.

We will denote by B and N the columns of A corresponding to the indices in \mathcal{B} and \mathcal{N} , respectively, and we will also partition the cost vector c as c_B and c_N ,

and the variables x and s as x_B and x_N , and s_B and s_N accordingly. Note that if (x, y, s) is a strictly complementary solution, then we have $x_B > 0$, $x_N = 0$, $s_B = 0$ and $s_N > 0$.

Let us again restrict our attention to one-dimensional perturbations of the right-hand side vector b . The optimal partition approach is based on maintaining the whole dual optimal set invariant rather than an optimal basis as in the classical simplex approach. Note that perturbations of b do not affect the dual feasible region. Consequently, the range of t is given by solving two auxiliary LPs. More precisely, if b is replaced by $b + t\Delta b$, and if Ω_D denotes the dual optimal set for (D) (i.e., $t = 0$), then the lower and upper bounds on t are given by the optimal values of

$$\begin{aligned}
 \text{(AUX1)} \quad & \min_{x,\lambda} \quad (\max_{x,\lambda}) \quad \lambda \\
 & \text{subject to} \\
 & Ax = b + \lambda\Delta b, \\
 & x \geq 0, \\
 & (s^*)^T x = 0, \quad \forall (y^*, s^*) \in \Omega_D.
 \end{aligned}$$

We will call the resulting bounds *the optimal partition bounds*. Note that both problems are always feasible since $\lambda = 0$ together with any $x \in \Omega_P$ satisfy all the constraints. Fixing the dual optimal set Ω_D is equivalent to fixing the optimal partition $(\mathcal{B}, \mathcal{N})$ by the Goldman-Tucker result. Therefore, the (possibly infinite) last constraint set in (AUX1) can be replaced by the equivalent single constraint $x^T s^* = 0$, where s^* is any point in the relative interior of Ω_D (hence $s_N^* > 0$). This condition, in turn, is the same as setting $x_N = 0$. Consequently, (AUX1) can be

written in the following simplified form:

$$\begin{aligned} \text{(AUX1)} \quad & \min_{x_B, \lambda} \quad (\max_{x_B, \lambda}) \quad \lambda \\ & \text{subject to} \end{aligned}$$

$$Bx_B = b + \lambda\Delta b,$$

$$x_B \geq 0.$$

The analogous derivation for the one-dimensional perturbations of the cost vector c leads to the following auxiliary problems, whose optimal values give the optimal partition bounds for t when c is replaced by $c + t\Delta c$:

$$\begin{aligned} \text{(AUX2)} \quad & \min_{y, s_N, \lambda} \quad (\max_{y, s_N, \lambda}) \quad \lambda \\ & \text{subject to} \end{aligned}$$

$$B^T y = c_B + \lambda\Delta c_B,$$

$$N^T y + s_N = c_N + \lambda\Delta c_N,$$

$$s_N \geq 0.$$

Here, Δc_B and Δc_N constitute the corresponding partition of Δc .

Before getting into the symmetrized bounds we would like to recall an important result about the dimensions of the optimal solution sets Ω_P and Ω_D . In what follows, $\dim(\cdot)$ denotes the dimension and $|\cdot|$ denotes the cardinality of a set. The reader is referred to Lemma IV.44 in [55] for a proof.

Proposition 3.2.1 $\dim(\Omega_P) = |\mathcal{B}| - \text{rank}(B)$; $\dim(\Omega_D) = m - \text{rank}(B)$.

□

3.2.1 Symmetrized Bounds

The auxiliary problems (AUX1) and (AUX2) can be reformulated in the following way. Let us consider (AUX1) and let $x^* \in \Omega_P$. Then, the equality constraint can be rewritten as

$$Bx_B = Bx_B^* + \lambda\Delta b \quad \text{or} \quad B(x_B - x_B^*) = \lambda\Delta b.$$

Therefore, by a change of variable, if we let $u = x_B - x_B^*$, then (AUX1) is equivalent to

$$\begin{aligned} \text{(AUX1)} \quad & \min_{u,\lambda} (\max_{u,\lambda}) \quad \lambda \\ & \text{subject to} \end{aligned}$$

$$\begin{aligned} Bu &= \lambda\Delta b, \\ u &\geq -x_B^*. \end{aligned}$$

Next, we will tighten the constraints in the above formulation by putting upper bounds on u as well, and our choice for the upper bound will be x_B^* , which will give the largest L_∞ -box around the origin which is contained in the feasible region:

$$\begin{aligned} \text{(SA1)} \quad & \min_{u,\lambda} (\max_{u,\lambda}) \quad \lambda \\ & \text{subject to} \end{aligned}$$

$$\begin{aligned} Bu &= \lambda\Delta b, \\ -x_B^* &\leq u \leq x_B^*. \end{aligned}$$

We will call (SA1) the *symmetrized LP* and the resulting optimal solutions the *symmetrized bounds*. The formulation of (SA1) reveals that if (u^*, λ^*) solves the maximization problem, then $(-u^*, -\lambda^*)$ solves the minimization problem. Therefore, it suffices to solve one LP as opposed to solving two LPs to obtain the optimal

partition bounds from (AUX1). A similar treatment of (AUX2) gives rise to the following symmetrized LP:

$$\begin{aligned}
 \text{(SA2)} \quad & \min_{v,w,\lambda} \quad (\max_{v,w,\lambda}) \quad \lambda \\
 & \text{subject to} \\
 & B^T v \quad = \quad \lambda \Delta c_B, \\
 & N^T v + w \quad = \quad \lambda \Delta c_N, \\
 & -s_N^* \leq w \leq s_N^*,
 \end{aligned}$$

which is obtained by replacing $y - y^*$ by v and $s_N - s_N^*$ by w , where $(y^*, s^*) \in \Omega_D$. Finally, a similar symmetrization has been applied to w .

Next, we would like to discuss the relationship between the auxiliary and the symmetrized LPs. First of all, let us assume that both (P) and (D) have unique and nondegenerate solutions. Then, Proposition 3.2.1 implies that B is actually a square and nonsingular matrix, hence invertible. In fact, B is the optimal basis. Consequently, (AUX1) and (AUX2) are trivial to solve and their optimal solutions coincide with the optimal basis bounds arising from the simplex method. With this observation, the constraints of (AUX1) reduce to

$$\lambda B^{-1} \Delta b \geq -x_B^* \quad \text{or} \quad \lambda (X_B^*)^{-1} B^{-1} \Delta b \geq -e, \quad (3.3)$$

where X_B^* is the diagonal matrix whose components are given by x_B^* and e denotes the vector of ones in the appropriate dimension. Similarly, the constraints of (SA1) can be rewritten as

$$-e \leq \lambda (X_B^*)^{-1} B^{-1} \Delta b \leq e \quad \text{or} \quad |\lambda| \|(X_B^*)^{-1} B^{-1} \Delta b\|_\infty \leq 1, \quad (3.4)$$

where $\|\cdot\|_\infty$ is the L_∞ -norm. A similar treatment reveals that the constraints of (AUX2) are equivalent to

$$\lambda(S_N^*)^{-1}(\Delta c_N - N^T B^{-T} \Delta c_B) \geq -e, \quad (3.5)$$

where S_N^* is defined similarly, and that those of (SA2) to

$$|\lambda| \|(S_N^*)^{-1}(\Delta c_N - N^T B^{-T} \Delta c_B)\|_\infty \leq 1. \quad (3.6)$$

The derivations (3.3)–(3.6) imply the following relationship between the auxiliary and the symmetrized LPs: let (λ^-, λ^+) denote the optimal partition bounds given by the optimal solutions of the auxiliary LPs (including possibly $\pm\infty$). Then, the symmetrized bounds for t are $(-\lambda^s, \lambda^s)$, where

$$\lambda^s = \min(|\lambda^-|, \lambda^+). \quad (3.7)$$

Therefore, the symmetrized bounds are indeed equal to the “symmetrization” of the optimal partition bounds.

Next, let us assume that (P) has a unique but degenerate solution. Then, by Proposition 3.2.1, B is nonsquare but it has full column rank. Therefore, (AUX1) is still easy to solve. If Δb does not lie in the range space of B , then the optimal solutions of (AUX1) and (SA1) are all zero (which implies that $t = 0$ is a breakpoint of $v(t)$). Otherwise, there exists a unique vector v such that $Bv = \Delta b$, and hence, the constraints of (AUX1) are equivalent to

$$\lambda(X_B^*)^{-1}v \geq -e. \quad (3.8)$$

Similarly, the constraints of (SA1) can be stated as

$$|\lambda| \|(X_B^*)^{-1}v\|_\infty \leq 1. \quad (3.9)$$

Once again, we conclude that a similar symmetry as in (3.7) continues to hold between (SA1) and (AUX1). In a similar manner, one can show that such a relationship holds between (SA2) and (AUX2) if (D) has a unique but degenerate solution.

The preceding discussion shows that the optimal solutions of the auxiliary and the symmetrized LPs have the nice relationship (3.7) as long as there is a unique optimal solution that one can use to symmetrize the constraints of the auxiliary LPs to obtain the symmetrized LPs. An interesting question then is whether the same nice relationship continues to hold between the auxiliary and the symmetrized LPs if there are multiple optimal solutions, that is whether the symmetrized bounds are independent of the choice of the optimal solution used to symmetrize the constraints. Unfortunately, the answer is no as shown by the following example. Let (P) be given by $\min\{x_2 - x_1 : x_1 - x_2 = 0, x_2 + x_3 = 1, x \geq 0\}$. Then (P) has multiple optimal solutions given by $(x_1, x_2, x_3) = (\beta, \beta, 1 - \beta)$ where $\beta \in [0, 1]$, with an optimal value of 0. If the right-hand side is perturbed to $(0, 1)^T + t(2, 1)^T$, then the reader can easily verify that (AUX1) yields $(-1/3, +\infty)$ as the optimal partition bounds, whereas the symmetrized bounds are $(-\beta, +\beta)$ if one uses the optimal solutions with $\beta < 1/3$ to symmetrize the constraints, and $(-1/3, 1/3)$ if those with $\beta \geq 1/3$ are used. This example illustrates that in case of multiple optimal solutions, the symmetrized bounds are dependent on the optimal solution used in the formulation of the symmetrized LPs. Therefore, the relationship (3.7) no longer holds between the symmetrized and the auxiliary LPs.

However, we will keep using the symmetrized LPs for two reasons. First of all,

at least in the unique solution case, they bear a nice relationship to the auxiliary LPs. For our analysis, we will always choose an optimal solution in the relative interior of the optimal set; therefore the symmetrization will hopefully allow more room for the decision variables of the symmetrized LPs. Secondly, the symmetrized LPs are easier to deal with than the auxiliary LPs and the symmetrized bounds will provide a good comparison basis for our interior-point approach proposed in [71] (Chapter 2), as will be analyzed in the subsequent sections.

3.2.2 Interior-Point Approach and Central Path Neighborhoods

We will start with a brief review of the primal-dual path-following interior-point methods. The reader is referred to [67] for an extensive treatment. The central path is a path of strictly feasible points $(x(\nu), y(\nu), s(\nu))$ parametrized by a positive scalar ν . Each point on the central path satisfies the following system for some $\nu > 0$:

$$\begin{aligned} Ax &= b, \\ A^T y + s &= c, \\ XSe &= \nu e, \end{aligned} \tag{3.10}$$

with $x > 0$ and $s > 0$. Under the two assumptions in Section 3.2, such a solution exists and is unique for each positive ν . Interior-point methods are iterative algorithms that generate iterates which “follow” the central path in the direction of decreasing ν towards the primal-dual optimal set $\Omega_P \times \Omega_D$. The iterates generated typically lie in some neighborhood of the central path. For any given feasible iterate

(x, y, s) , the duality gap is given by $c^T x - b^T y = x^T s \geq 0$ and we define the duality measure μ as $\mu := \mu(x, s) := x^T s/n$. Let \mathcal{S} and \mathcal{S}^0 denote the set of feasible and strictly feasible primal-dual points respectively, that is,

$$\mathcal{S} = \{(x, y, s) : Ax = b, A^T y + s = c, (x, s) \geq 0\}, \quad (3.11)$$

$$\mathcal{S}^0 = \{(x, y, s) \in \mathcal{S} : (x, s) > 0\}. \quad (3.12)$$

One of the commonly used neighborhoods in interior-point methods is the so-called *wide neighborhood*, denoted by $\mathcal{N}_{-\infty}(\gamma)$:

$$\mathcal{N}_{-\infty}(\gamma) = \{(x, y, s) \in \mathcal{S}^0 : x_i s_i \geq \gamma \mu, \forall i = 1, 2, \dots, n\}, \quad (3.13)$$

where $\gamma \in (0, 1]$.

At each iteration, given $(x, y, s) \in \mathcal{N}_{-\infty}(\gamma)$, the algorithm determines a search direction $(\Delta x, \Delta y, \Delta s)$. This direction is usually obtained by seeking an approximation to the point on the central path corresponding to some parameter $\nu \leq \mu$, and then applying Newton's method to the nonlinear system of equations (3.10). Finally, a (damped) step is taken in this direction in such a way that the resulting iterate still lies in $\mathcal{N}_{-\infty}(\gamma)$.

However, as in the context of target-following methods, one might seek an approximation to a point other than the one on the central path. We will say that a Newton step from (x, y, s) *targeting* the feasible pair of points (x', y', s') that satisfies $X'S'e = v$ is the direction $(\Delta x, \Delta y, \Delta s)$ obtained from the Newton's method

applied to (3.10) with νe replaced by $X'S'e$:

$$\begin{aligned} A\Delta x &= b - Ax, \\ A^T\Delta y + \Delta s &= c - A^T y, \\ S\Delta x + X\Delta s &= v - XSe. \end{aligned} \tag{3.14}$$

Next, we describe the interior-point approach proposed by the authors in [71]. Given a primal-dual pair of LPs (P) and (D), let us assume that b or c is perturbed in some fixed direction. Assuming (x, y, s) is strictly primal-dual feasible for (P) and (D), a full Newton step is taken from (x, y, s) targeting “a feasible point” (x', y', s') of the perturbed LPs which satisfies $X'S'e = XSe$. (It is possible that there is no such feasible point for the perturbed LPs, however, the Newton step as given above is still well-defined.) We state the results formally, referring the reader to [71] for the proofs. Note, in particular, that the duality gap of the resulting feasible iterate for the perturbed LPs is no greater than that of the original iterate.

Proposition 3.2.2 *Assume that (x, y, s) is a strictly feasible point for (P) and (D) and the right-hand side vector b is replaced by $b + t\Delta b$, where $t \in \mathbb{R}$ and $\Delta b \in \mathbb{R}^m$. Suppose a Newton step is taken from (x, y, s) targeting the feasible pair of points (x', y', s') of the perturbed pair of LPs that satisfies $X'S'e = XSe$. Then a full Newton step will yield a feasible iterate for the new problem if and only if*

$$|t| \leq \frac{1}{\|S^{-1}A^T(AD^2A^T)^{-1}\Delta b\|_\infty}, \tag{3.15}$$

where $D = X^{\frac{1}{2}}S^{-\frac{1}{2}}$. Moreover, in this case the new iterate will have duality gap at most $x^T s$.

Proposition 3.2.3 *Assume that (x, y, s) is a strictly feasible point for (P) and (D) and the cost vector c is replaced by $c + t\Delta c$, where $t \in \mathbb{R}$ and $\Delta c \in \mathbb{R}^n$. Suppose a Newton step is taken from (x, y, s) targeting the feasible pair of points (x', y', s') of the perturbed pair of LPs that satisfies $X'S'e = XSe$. Then a full Newton step will yield a feasible iterate for the new problem if and only if*

$$|t| \leq \frac{1}{\|S^{-1}(I - A^T(AD^2A^T)^{-1}AD^2)\Delta c\|_\infty}, \quad (3.16)$$

where $D = X^{\frac{1}{2}}S^{-\frac{1}{2}}$. Moreover, in this case the new iterate will have duality gap at most $x^T s$.

Under primal-dual nondegeneracy, the bounds arising from Propositions 3.2.2 and 3.2.3 computed at near-optimal solutions for (P) and (D) asymptotically equal the symmetrized bounds arising from (SA1) and (SA2) [71]. The goal of this paper is to investigate the quality of these bounds in the absence of the nondegeneracy assumption.

We first present a nice characterization of the distance of the strictly feasible primal-dual points (x, y, s) from strictly complementary optimal solutions in terms of the duality gap μn . Using this characterization, we derive some bounds on the components of such points. In what follows, x_B, x_N, s_B and s_N denote the partitions of x and s according to the optimal partition $(\mathcal{B}, \mathcal{N})$ as before. Furthermore, we will use the bounds $O(\mu)$, $\Omega(\mu)$ and $\Theta(\mu)$ interchangeably for scalars as well as vectors and matrices by which we mean each entry satisfies the stated bounds. $O(\mu)$ will indicate that the quantity (in absolute value) is bounded above by some positive multiple of μ , where the multiple depends on the primal-dual instance (P) and (D)

but does not depend on the particular strictly feasible point or on μ . Similarly, $\Omega(\mu)$ will indicate a lower bound by some positive multiple of μ and $\Theta(\mu)$ will mean a lower and upper bound by some positive multiples of μ .

The following proposition will be useful for the analysis that follows. Actually, the proposition continues to hold for any feasible solutions and even for a point where feasibility is violated by $O(\mu)$. The statement below suffices for the purposes of this paper.

Proposition 3.2.4 *Let (x, y, s) be a strictly feasible point for (P) and (D) with duality gap μn . Then, there exists a strictly complementary optimal solution (x^*, y^*, s^*) of (P) and (D) such that*

$$(x, y, s) = (x^*, y^*, s^*) + O(\mu). \quad (3.17)$$

Proof: Optimal solutions of (P) and (D) satisfy the linear system $Ax = b$, $A^T y + s = c$, $c^T x - b^T y = 0$, $x \geq 0$, $s \geq 0$. Any strictly feasible point (x, y, s) satisfies the same linear system with the third equality replaced by $c^T x - b^T y = \mu n$. Hoffman's lemma [30] indicates that there exists a solution $(\hat{x}, \hat{y}, \hat{s})$ of the first system such that $(\hat{x}, \hat{y}, \hat{s}) = (x, y, s) + O(\mu)$. The result follows immediately if $(\hat{x}, \hat{y}, \hat{s})$ is strictly complementary. If not, there exists an arbitrarily small perturbation of $(\hat{x}, \hat{y}, \hat{s})$ which leads to a strictly complementary solution and (3.17) follows since $\mu > 0$. \square

The following corollary immediately follows from Proposition 3.2.4 since $x_N^* = 0$ and $s_B^* = 0$ for any optimal solution of (P) and (D).

Corollary 3.2.1 *Let (x, y, s) be a strictly feasible point for (P) and (D) with duality gap μn . Then,*

$$x_N = O(\mu), \quad s_B = O(\mu). \quad (3.18)$$

□

Note that both Proposition 3.2.4 and Corollary 3.2.1 hold for *any* primal-dual strictly feasible (x, y, s) . Next, we derive some more bounds by restricting the iterates to lie in a wide neighborhood given by (3.13).

Proposition 3.2.5 *Let $(x, y, s) \in \mathcal{N}_{-\infty}(\gamma)$ with duality gap μn for (P) and (D). Then,*

$$XSe = \Theta(\mu), \quad s_N = \Omega(1), \quad x_B = \Omega(1), \quad X_N S_N^{-1} e = O(\mu), \quad S_B X_B^{-1} e = O(\mu). \quad (3.19)$$

Proof: Since $(x, y, s) \in \mathcal{N}_{-\infty}(\gamma)$, we have $x_i s_i \geq \gamma \mu$. Moreover, $x^T s = \mu n$. Therefore, $x_i s_i \leq \mu n$ since $x > 0$ and $s > 0$. This proves $XSe = \Theta(\mu)$. By Corollary 3.2.1, $x_N = O(\mu)$. Then, $XSe = \Theta(\mu)$ implies $s_N = \Omega(1)$. A similar argument shows $x_B = \Omega(1)$. Finally, $x_N = O(\mu)$ together with $s_N = \Omega(1)$ imply $X_N S_N^{-1} e = O(\mu)$. The proof of $S_B X_B^{-1} e = O(\mu)$ is similar. □

3.3 Equivalence

In this section, we show that the interior-point bounds are independent of the problem formulation. It is well-known that although (P) and (D) do not look symmetric, they can easily be reformulated so that (D) is in the form of (P) and

vice versa. We briefly review this reformulation. Let $(\hat{x}, \hat{y}, \hat{s})$ be such that $A\hat{x} = b$ and $A^T\hat{y} + \hat{s} = c$. Let us consider (D) first. The objective function can be rewritten as

$$b^T y = \hat{x}^T A^T y = \hat{x}^T c - \hat{x}^T s, \quad (3.20)$$

where we used $A\hat{x} = b$ and the fact that every feasible pair (y, s) for (D) satisfies $A^T y + s = c$. Note that the first term is a constant: therefore maximizing $b^T y$ is the same as minimizing $\hat{x}^T s$. Let $K \in \mathbb{R}^{(n-m) \times n}$ be such that its rows form a basis for the null space of A . Then, premultiplying the equality constraints in (D) by K yields

$$Ks = Kc =: \hat{c}. \quad (3.21)$$

Moreover, if s satisfies (3.21), then $c - s$ lies in the null space of K , for which the columns of A^T form a basis by definition of K . Therefore, there exists y such that $A^T y = c - s$. Consequently, (D) is equivalent to

$$\min_s \hat{x}^T s, \text{ subject to } Ks = \hat{c}, \quad s \geq 0. \quad (\text{D}')$$

Note in particular that K has full row rank by its definition. If we take the dual of (D'), we obtain

$$\max_{u,x} \hat{c}^T u, \text{ subject to } K^T u + x = \hat{x}, \quad x \geq 0. \quad (\text{P}')$$

It is not hard to see that (P) and (P') are also equivalent by a similar argument. Therefore, the roles of (P) and (D) can be interchanged via this reformulation.

Let us now focus on perturbations of c , i.e., let c be replaced by $c + t\Delta c$. By the above reformulation, this is the same as replacing the right-hand side of (D') by

$\hat{c} + tK\Delta c$. Therefore, Proposition 3.2.2 can be used to evaluate the interior-point bound at a strictly feasible primal-dual pair (s, x) (note that the roles of x and s are interchanged). We need to compute

$$X^{-1}K^T(KSX^{-1}K^T)^{-1}K\Delta c. \quad (3.22)$$

On the other hand, one can also use Proposition 3.2.3 to compute the interior-point bound directly at (x, s) , which requires the evaluation of

$$S^{-1}(I - A^T(AXS^{-1}A^T)^{-1}AXS^{-1})\Delta c. \quad (3.23)$$

A simple manipulation of (3.22) gives rise to another equivalent formula:

$$X^{-1/2}S^{-1/2}\Psi X^{1/2}S^{-1/2}\Delta c, \quad (3.24)$$

where Ψ is the orthogonal projection matrix onto the range space of $X^{-1/2}S^{1/2}K^T$.

Similarly, (3.23) is equivalent to

$$X^{-1/2}S^{-1/2}\Xi X^{1/2}S^{-1/2}\Delta c, \quad (3.25)$$

where Ξ is the orthogonal projection matrix onto the null space of $AX^{1/2}S^{-1/2}$. Therefore, in order to prove that (3.22) and (3.23) are equivalent, it suffices to show that Ψ and Ξ project onto the same subspace, or that the null space of $AX^{1/2}S^{-1/2}$ equals the range space of $X^{-1/2}S^{1/2}K^T$. This is easily proven by an inclusion argument: if w satisfies $AX^{1/2}S^{-1/2}w = 0$, then $X^{1/2}S^{-1/2}w = K^T u$ for some unique u . Thus, w is in the range space of $X^{-1/2}S^{1/2}K^T$. Conversely, if $w = X^{-1/2}S^{1/2}K^T u$ for some u , then $AX^{1/2}S^{-1/2}w = AK^T u = 0$. This proves the equivalence of the interior-point bounds.

We next argue that the range of t resulting from the optimal partition bounds is also independent of the formulation. If the two LPs are formulated in the form of (P) and (D), then (AUX2) yields the range of t for perturbations of c . Premultiplying the equality constraints of (AUX2) by $K = [K_B, K_N]$ leads to (AUX1') given by

$$\min_{w, \lambda} (\max_{w, \lambda}) \lambda \text{ s.t. } K_N w = \lambda K \Delta c, w \geq -s_N^*, \quad (\text{AUX1}') \quad (3.26)$$

which exactly yields the range of t for perturbations of the right-hand side of (D') if one uses the form (D') and (P'). Similarly, if (w, λ) is feasible for (AUX1'), then

$$\begin{bmatrix} \lambda \Delta c_B \\ \lambda \Delta c_N - w \end{bmatrix}$$

lies in the null space of K . Then, by our previous observation, there exists v such that $B^T v = \lambda \Delta c_B$, $N^T v + w = \lambda \Delta c_N$, which is exactly the constraints of (AUX2), completing the proof of the claim.

Using this observation, we will carry out our analysis for perturbations of b only in the subsequent sections, and state the corresponding results for changes in c as corollaries. We begin with a special case of degeneracy first and then consider the most general case.

3.4 Unique Primal Solution

Throughout this section, we assume that (P) has a unique but degenerate optimal solution x^* . Note that by Proposition 3.2.1, we have $|\mathcal{B}| = \text{rank}(B)$, i.e., B has linearly independent columns. In this particular case, Proposition 3.2.4 provides another useful bound on x_B for a strictly feasible primal-dual point (x, y, s) .

Corollary 3.4.1 *Assume that (P) has a unique optimal solution x^* . Let (x, y, s) be primal-dual strictly feasible for (P) and (D) with duality gap μn . Then,*

$$x_B = x_B^* + O(\mu). \quad (3.27)$$

□

An analogous corollary follows if (D) has a unique solution.

Corollary 3.4.2 *Assume that (D) has a unique optimal solution (y^*, s^*) . Let (x, y, s) be primal-dual strictly feasible for (P) and (D) with duality gap μn . Then,*

$$s_N = s_N^* + O(\mu). \quad (3.28)$$

□

Next, we will analyze one-dimensional perturbations of b .

3.4.1 Perturbations of b

In this subsection, we assume that the right-hand side vector b is replaced by $b + t\Delta b$, where $\Delta b \in \mathbb{R}^m$ and $t \in \mathbb{R}$. We also assume that $(x, y, s) \in \mathcal{N}_{-\infty}(\gamma)$ is a primal-dual strictly feasible point for (P) and (D) for some $\gamma \in (0, 1]$. We will compare the interior-point bounds arising from Proposition 3.2.2 at (x, y, s) with the optimal values of (SA1), i.e., the symmetrized bounds. The interior-point bounds are given by the L_∞ -norm of

$$S^{-1}A^T(AD^2A^T)^{-1}\Delta b, \quad (3.29)$$

where $D^2 = XS^{-1}$.

Let us now consider (SA1). Since B has full column rank, Δb either does not lie in the range space of B , in which case the optimal values of (SA1) as well as (AUX1) are all 0, or there exists a unique $v \in \mathbb{R}^{|\mathcal{B}|}$ such that $Bv = \Delta b$, in which case the constraints of (SA1) reduce to (3.9), from which the symmetrized bounds can be obtained easily. We will consider both situations in turn.

Let us start with the second case. Without loss of generality, we can assume that Δb has unit L_2 -norm, which implies a bound on v . Then, we need to compute

$$u = (AD^2A^T)^{-1}\Delta b = (AD^2A^T)^{-1}Bv \quad (3.30)$$

in order to obtain (3.29). However, (3.30) is equivalent to

$$AD^2A^T u = Bv \quad \text{or} \quad BX_B S_B^{-1} B^T u + NX_N S_N^{-1} N^T u = Bv, \quad (3.31)$$

where B and N are the partitions of the coefficient matrix A with respect to \mathcal{B} and \mathcal{N} as before. Since B has linearly independent columns, there exists a matrix $C \in \mathbb{R}^{m \times (m-|\mathcal{B}|)}$ such that the augmented matrix $[B \ C]$ is square and nonsingular: let W be its inverse. Therefore, premultiplying the second equality in (3.31) by W , we obtain

$$\begin{bmatrix} I \\ 0 \end{bmatrix} X_B S_B^{-1} [I \ 0] \tilde{u} + \tilde{N} X_N S_N^{-1} \tilde{N}^T \tilde{u} = \begin{bmatrix} I \\ 0 \end{bmatrix} v, \quad (3.32)$$

where $\tilde{u} = W^{-T}u$, $\tilde{N} = WN$ and I is a $|\mathcal{B}| \times |\mathcal{B}|$ identity matrix. Therefore, if we partition \tilde{N} and \tilde{u} accordingly as

$$\tilde{N} = \begin{bmatrix} \tilde{N}_1 \\ \tilde{N}_2 \end{bmatrix}, \quad \tilde{u} = \begin{bmatrix} \tilde{u}_1 \\ \tilde{u}_2 \end{bmatrix},$$

(3.32) can then be decomposed in the following way:

$$\begin{bmatrix} D_B^2 + \tilde{N}_1 D_N^2 \tilde{N}_1^T & \tilde{N}_1 D_N^2 \tilde{N}_2^T \\ \tilde{N}_2 D_N^2 \tilde{N}_1^T & \tilde{N}_2 D_N^2 \tilde{N}_2^T \end{bmatrix} \begin{bmatrix} \tilde{u}_1 \\ \tilde{u}_2 \end{bmatrix} = \begin{bmatrix} v \\ 0 \end{bmatrix}, \quad (3.33)$$

where D_B and D_N are the corresponding partitions of D . By (3.29), we need to compute

$$S^{-1} A^T u = S^{-1} (W A)^T \tilde{u} = \begin{bmatrix} S_B^{-1} \tilde{u}_1 \\ S_N^{-1} (\tilde{N}_1^T \tilde{u}_1 + \tilde{N}_2^T \tilde{u}_2) \end{bmatrix}. \quad (3.34)$$

For notational convenience, let us define

$$F := \tilde{N}_1 D_N, \quad G := \tilde{N}_2 D_N.$$

Note that G has full row rank since A does. The bottom equality in (3.33) can be rewritten as

$$G F^T \tilde{u}_1 + G G^T \tilde{u}_2 = 0, \quad \text{so} \quad \tilde{u}_2 = -(G G^T)^{-1} G F^T \tilde{u}_1. \quad (3.35)$$

Substituting (3.35) in the top equality in (3.33) gives

$$(D_B^2 + F F^T - F G^T (G G^T)^{-1} G F^T) \tilde{u}_1 = v, \quad \text{or} \quad (D_B^2 + F(I - P_G) F^T) \tilde{u}_1 = v, \quad (3.36)$$

where P_G is the orthogonal projection matrix onto the range space of G^T . Therefore, $I - P_G$ is the orthogonal projection matrix onto the null space of G .

We briefly review the Neumann lemma now [20]. Let U be an invertible matrix and let V satisfy $\|U^{-1}V\| \leq 1/2$. (The particular norm being used does not really matter: we will always use $\|\cdot\|$ for the Euclidean norm or the operator norm arising

from it.) Then, $I + U^{-1}V$ is invertible with $\|I + U^{-1}V\| \leq 2$. Moreover $U + V$ is invertible and given by

$$(U + V)^{-1} = U^{-1} - U^{-1}V(I + U^{-1}V)^{-1}U^{-1}. \quad (3.37)$$

Now, we apply this result to (3.36) with $U := D_B^2$ and $V := F(I - P_G)F^T$. Proposition 3.2.5 implies that both U^{-1} and V are $O(\mu)$ since $I - P_G$ is a projection matrix and has unit Euclidean norm. Therefore, assuming the duality gap μn is small,

$$\begin{aligned} \tilde{u}_1 &= (D_B^2 + F(I - P_G)F^T)^{-1} v, \\ &= D_B^{-2}v - D_B^{-2}F(I - P_G)F^T (I + D_B^{-2}F(I - P_G)F^T)^{-1} D_B^{-2}v. \end{aligned} \quad (3.38)$$

It then follows that

$$S_B^{-1}\tilde{u}_1 = X_B^{-1}v - X_B^{-1}F(I - P_G)F^T (I + D_B^{-2}F(I - P_G)F^T)^{-1} D_B^{-2}v. \quad (3.39)$$

However, by Proposition 3.2.5, F is $O(\mu^{1/2})$, D_B^{-2} is $O(\mu)$ and X_B^{-1} is $O(1)$. Consequently, the second term on the right hand side of (3.39) is $O(\mu^2)$ since $\|I - P_G\| \leq 1$. Finally, Corollary 3.4.1 implies $X_B^{-1} = (X_B^*)^{-1} + O(\mu)$. Therefore,

$$S_B^{-1}\tilde{u}_1 = (X_B^*)^{-1}v + O(\mu). \quad (3.40)$$

We have thus obtained the top part of (3.34). For the lower part, we get

$$\begin{aligned} S_N^{-1}(\tilde{N}_1^T \tilde{u}_1 + \tilde{N}_2^T \tilde{u}_2) &= S_N^{-1}D_N^{-1}(F^T \tilde{u}_1 + G^T \tilde{u}_2), \\ &= (X_N S_N)^{-1/2}(I - P_G)F^T \tilde{u}_1, \end{aligned} \quad (3.41)$$

where we substituted (3.35) for \tilde{u}_2 . Proposition 3.2.5 implies $(X_N S_N)^{-1/2} = O(\mu^{-1/2})$ and $F = O(\mu^{1/2})$. By (3.40), $\tilde{u}_1 = O(\mu)$ since s_B is. Combining these

bounds with $\|I - P_G\| \leq 1$ leads to

$$S_N^{-1}(\tilde{N}_1^T \tilde{u}_1 + \tilde{N}_2^T \tilde{u}_2) = O(\mu). \quad (3.42)$$

Using (3.34), we conclude that the L_∞ -norm of the quantity (3.29) we need to evaluate is given by

$$\|S^{-1}A^T(AD^2A^T)^{-1}\Delta b\|_\infty = \left\| \begin{bmatrix} (X_B^*)^{-1}v + O(\mu) \\ O(\mu) \end{bmatrix} \right\|_\infty. \quad (3.43)$$

The reciprocal of (3.43) gives the desired interior-point bound. Consequently, if the duality gap μn is small, we conclude by comparing (3.43) with (3.9) that the interior-point approach yields exactly the same bound as the optimal solution to (SA1) asymptotically in μ .

Next, we address the situation where Δb does not lie in the range space of B . In this case, the optimal values of both (AUX1) and (SA1) are clearly 0. Δb can be uniquely written as

$$\Delta b = Bv_B + Cv_C, \quad (3.44)$$

where $[B \ C]$ is nonsingular as before and v_C is a nonzero vector. Once again, we need to compute (3.29). We follow a similar treatment as before, and corresponding to (3.33) we obtain:

$$\begin{bmatrix} D_B^2 + \tilde{N}_1 D_N^2 \tilde{N}_1^T & \tilde{N}_1 D_N^2 \tilde{N}_2^T \\ \tilde{N}_2 D_N^2 \tilde{N}_1^T & \tilde{N}_2 D_N^2 \tilde{N}_2^T \end{bmatrix} \begin{bmatrix} \tilde{u}_1 \\ \tilde{u}_2 \end{bmatrix} = \begin{bmatrix} v_B \\ v_C \end{bmatrix}. \quad (3.45)$$

The bottom part can be expanded as

$$\tilde{N}_2 X_N \left[S_N^{-1} \tilde{N}_1^T \tilde{u}_1 + S_N^{-1} \tilde{N}_2^T \tilde{u}_2 \right] = v_C. \quad (3.46)$$

However, (3.34) implies that the term in the brackets is exactly the bottom part of the quantity (3.29) we seek. Let us denote that term by p and let $X_N p = q$. Then, (3.46) is equivalent to $\tilde{N}_2 q = v_C$. Since v_C is nonzero, the norm of q is bounded below, that is, $\|q\| \geq \alpha > 0$ where α is the norm of the least squares solution. Therefore, $\|q\|_\infty \geq \beta$ with $\beta := (\alpha/\sqrt{n - |\mathcal{B}|})$ (see e.g. [20]). (Note that $|\mathcal{B}| < n$ since $|\mathcal{B}| = n$ can happen only if $m = n$, in which case Δb is always in the range of B .) However, $\|q\|_\infty \leq \|X_N\|_\infty \|p\|_\infty$ since $X_N p = q$. This implies

$$\|p\|_\infty \geq \frac{\|q\|_\infty}{\|X_N\|_\infty} \geq \frac{\beta}{\|X_N\|_\infty} = \Omega(1/\mu), \quad (3.47)$$

where the last equality follows from Corollary 3.2.1. Therefore, as μ tends to 0, $\|p\|_\infty$ tends to ∞ , which implies that the interior-point bound given by its reciprocal tends to 0 as desired.

We remark that if $\mathcal{B} = \emptyset$, then $x^* = 0$ is the only optimal solution of (P), which can happen only if $b = 0$. In this case, the top part of (3.34) disappears. The interior-point bound is then given by the reciprocal of $\|p\|_\infty$, where p is as defined after (3.46). By the preceding argument, the interior-point bound tends to 0 as μ approaches 0. This is still in agreement with the optimal partition bounds since any nonzero perturbation of b leads to a change in the optimal partition and hence, the optimal partition bounds in this case are also equal to 0. Therefore, we have proved the following theorem:

Theorem 3.4.1 *Let $(x, y, s) \in \mathcal{N}_{-\infty}(\gamma)$ be a primal-dual strictly feasible point for (P) and (D). Assume that (P) has a unique but degenerate optimal solution and that b is replaced by $b + t\Delta b$ where $t \in \mathbb{R}$ and $\Delta b \in \mathbb{R}^m$. Then the interior-point*

bound evaluated at (x, y, s) yields exactly the same value as the optimal solution of (SA1) asymptotically in μ , where $\mu = x^T s/n$.

□

The following corollary of Theorem 3.4.1 is an immediate consequence of the equivalence between (P) and (D) as discussed in Section 3.3. One uses Corollary 3.4.2 in place of Corollary 3.4.1 in the preceding analysis.

Corollary 3.4.3 *Let $(x, y, s) \in \mathcal{N}_{-\infty}(\gamma)$ be a primal-dual strictly feasible point for (P) and (D). Assume that (D) has a unique but degenerate optimal solution and that c is replaced by $c + t\Delta c$ where $t \in \mathbb{R}$ and $\Delta c \in \mathbb{R}^n$. Then the interior-point bound evaluated at (x, y, s) yields exactly the same value as the optimal solution of (SA2) asymptotically in μ , where $\mu = x^T s/n$.*

□

It does not appear that we can obtain better results for perturbations of c in the case of a unique primal optimal solution (but not dual optimal solution) than those arising from the analysis of the general case in the next section. A similar remark holds for perturbations of b in the case of a unique dual optimal solution (but not primal optimal solution).

3.5 General Case

In this section, we turn our attention to the most general case where both (P) and (D) may have multiple optimal solutions. As the small example given at the end of

Section 3.2.1 reveals, some complications arise in the presence of multiple optimal solutions. For instance, unlike the previous case, the symmetrized bounds become dependent on the optimal solution of (P) used in the formulation of (SA1) if (P) has multiple optimal solutions. Furthermore, they do not necessarily coincide with the “symmetrizations” of the optimal partition bounds arising from (AUX1). Similar remarks hold for the relationship between (SA2) and (AUX2) if (D) has multiple optimal solutions.

Despite this complication arising from the presence of multiple optimal solutions, we aim to be able to say something about the quality of the interior-point bounds at least in comparison with the symmetrized bounds. In the next subsection, we analyze perturbations of b in this general setting.

3.5.1 Perturbations of b

Let (P) have multiple optimal solutions and let b be replaced by $b+t\Delta b$, where $t \in \mathbb{R}$ and $\Delta b \in \mathbb{R}^m$. Suppose that $(x, y, s) \in \mathcal{N}_{-\infty}(\gamma)$ is primal-dual strictly feasible where $\gamma \in (0, 1]$. For such a point, Proposition 3.2.4 guarantees the existence of a strictly complementary solution (x^*, y^*, s^*) whose distance from (x, y, s) is bounded above by the duality gap $n\mu$. We will compare the interior-point bounds evaluated at (x, y, s) with the optimal values of (SA1). Among other optimal solutions of (P), the x^* above will be the particular choice of the primal optimal solution to be used in the formulation of (SA1). The use of such an optimal solution in the relative interior of the primal optimal set is likely to leave more room for the decision variables of (SA1) since $x_B^* > 0$.

Let us first consider (SA1). The constraints of (SA1) are

$$\begin{aligned} Bu &= \lambda \Delta b, \\ -x_B^* &\leq u \leq x_B^*. \end{aligned} \tag{3.48}$$

Let $\text{rank}(B) = r$ and $|\mathcal{B}| = k$. Clearly we have $r \leq m$ and $r < k$ since Proposition 3.2.1 implies $\dim(\Omega_P) = k - r$, which is positive by our assumption. This, in turn, implies that $r > 0$ since $r = 0$ if and only if $\mathcal{B} = \emptyset$ (assuming no columns of A are identically zero). A QR factorization of B yields $B = QR$, where $Q \in \mathbb{R}^{m \times m}$ is orthogonal and $R \in \mathbb{R}^{m \times k}$ is upper triangular with

$$R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix}, \tag{3.49}$$

where R_1 has r rows. Note that R_1 has full row rank.

Premultiplying the equality constraints in (3.48) by Q^T yields

$$\begin{aligned} \begin{bmatrix} R_1 \\ 0 \end{bmatrix} u &= \lambda \begin{bmatrix} \widetilde{\Delta b}_1 \\ \widetilde{\Delta b}_2 \end{bmatrix}, \\ -x_B^* &\leq u \leq x_B^*, \end{aligned} \tag{3.50}$$

with $\widetilde{\Delta b} = Q^T \Delta b$. Clearly, (3.50) reveals that (SA1) has a nontrivial optimal solution λ^* if and only if $\widetilde{\Delta b}_2 = 0$.

First, we consider the nontrivial case. (Since $\widetilde{\Delta b}$ is nonzero, this implies that $k > 0$.) Let (λ^*, u^*) be an optimal solution to the maximization problem with $\lambda^* \neq 0$. Note that λ^* is finite since u^* is bounded (this follows since $\mathcal{B} \neq \emptyset$). Then, we have

$$\widetilde{\Delta b} = Q^T \Delta b = (1/\lambda^*) R u^*. \tag{3.51}$$

The interior-point approach, on the other hand, requires the evaluation of (3.29) at (x, y, s) . By (3.51), we then need to evaluate the L_∞ -norm of

$$(1/\lambda^*)S^{-1}A^T(AD^2A^T)^{-1}QRu^*. \quad (3.52)$$

Let

$$w = (AD^2A^T)^{-1}QRu^* \quad \text{or} \quad AD^2A^T w = QRu^*. \quad (3.53)$$

Premultiplying the second equality in (3.53) by Q^T gives

$$\begin{bmatrix} R_1 & \tilde{N}_1 \\ 0 & \tilde{N}_2 \end{bmatrix} \begin{bmatrix} D_B^2 & 0 \\ 0 & D_N^2 \end{bmatrix} \begin{bmatrix} R_1^T & 0 \\ \tilde{N}_1^T & \tilde{N}_2^T \end{bmatrix} \begin{bmatrix} \tilde{w}_1 \\ \tilde{w}_2 \end{bmatrix} = \begin{bmatrix} R_1 \\ 0 \end{bmatrix} u^*, \quad (3.54)$$

where \tilde{w}_1 and \tilde{w}_2 are the appropriate partitions of $\tilde{w} = Q^T w$ and \tilde{N}_1 and \tilde{N}_2 are those of $\tilde{N} = Q^T N$. Let us define

$$F := \tilde{N}_1 D_N, \quad G := R_1 D_B, \quad H := \tilde{N}_2 D_N. \quad (3.55)$$

(3.54) can then be decomposed into two equations as

$$\begin{aligned} (GG^T + FF^T)\tilde{w}_1 + FH^T\tilde{w}_2 &= R_1 u^*, \\ HF^T\tilde{w}_1 + HH^T\tilde{w}_2 &= 0. \end{aligned} \quad (3.56)$$

Note, in particular, that both G and H have full row rank since R_1 and A do. From the second equation in (3.56), we obtain

$$\tilde{w}_2 = -(HH^T)^{-1}HF^T\tilde{w}_1. \quad (3.57)$$

Substituting (3.57) in the first equation of (3.56) leads to

$$\begin{aligned} (GG^T + FF^T - FH^T(HH^T)^{-1}HF^T)\tilde{w}_1 &= R_1 u^*, \quad \text{or} \\ (GG^T + F(I - P_H)F^T)\tilde{w}_1 &= R_1 u^*, \end{aligned} \quad (3.58)$$

where $I - P_H$ is the orthogonal projection matrix onto the null space of H . Proposition 3.2.5 implies that the second term in parentheses in the second equation above is $O(\mu)$ since $\|I - P_H\| \leq 1$. In order to apply Neumann's lemma, we need to show that $(GG^T)^{-1}$ is bounded.

Lemma 3.5.1 $(GG^T)^{-1} = O(\mu)$.

Proof: We use the “thin” QR factorization of $G^T = D_B R_1^T = YZ$, where Y has orthonormal columns and Z is upper triangular and nonsingular. Then, $(GG^T)^{-1} = Z^{-1}Z^{-T}$. Therefore, it suffices to find an upper bound on Z^{-1} . We have

$$D_B R_1^T = YZ, \quad \text{or} \quad R_1 R_1^T = R_1 D_B^{-1} Y Z. \quad (3.59)$$

Therefore, $I = (R_1 R_1^T)^{-1} R_1 D_B^{-1} Y Z$, or $Z^{-1} = (R_1 R_1^T)^{-1} R_1 D_B^{-1} Y$. However, by Proposition 3.2.5, $D_B^{-1} = O(\mu^{1/2})$, which implies that $Z^{-1} = O(\mu^{1/2})$ completing the proof. \square

We can now apply Neumann's lemma to (3.58). Using the same notation as in (3.37) we have $U := GG^T$ and $V := F(I - P_H)F^T$. Note that both U^{-1} and V are $O(\mu)$. We obtain

$$\tilde{w}_1 = (GG^T)^{-1} G D_B^{-1} u^* - (GG^T)^{-1} V (I + (GG^T)^{-1} V)^{-1} (GG^T)^{-1} G D_B^{-1} u^*, \quad (3.60)$$

where we used $R_1 = G D_B^{-1}$.

By (3.52) and (3.53), we need

$$(1/\lambda^*) S^{-1} A^T w = (1/\lambda^*) \begin{bmatrix} S_B^{-1} R_1^T \tilde{w}_1 \\ S_N^{-1} (\tilde{N}_1^T \tilde{w}_1 + \tilde{N}_2^T \tilde{w}_2) \end{bmatrix}. \quad (3.61)$$

Let us define

$$W = G^T(GG^T)^{-1}. \quad (3.62)$$

For the top part of (3.61) we need to evaluate

$$\begin{aligned} S_B^{-1}R_1^T\tilde{w}_1 &= S_B^{-1}D_B^{-1}G^T\tilde{w}_1, \\ &= (S_B X_B)^{-1/2}P_G D_B^{-1}u^* \\ &\quad - (S_B X_B)^{-1/2}WV(I + (GG^T)^{-1}V)^{-1}W^T D_B^{-1}u^*, \end{aligned} \quad (3.63)$$

where we used (3.60), (3.62) and where P_G is the orthogonal projection matrix onto the range space of G^T . Consider the second term in the right hand side of the second equality. By Proposition 3.2.5, $(S_B X_B)^{-1/2}$ is $O(\mu^{-1/2})$, V is $O(\mu)$ and D_B^{-1} is $O(\mu^{1/2})$. Lemma 3.5.1 implies that $W = G^T(GG^T)^{-1} = YZ^{-T} = O(\mu^{1/2})$ since $\|Y\| \leq 1$. Therefore, the whole expression is $O(\mu^2)$. We conclude that the top part of (3.61) is

$$(1/\lambda^*)(S_B X_B)^{-1/2}P_G(S_B X_B)^{1/2}X_B^{-1}u^* + (1/\lambda^*)O(\mu^2). \quad (3.64)$$

Let us next consider the lower part of (3.61). We need to compute

$$S_N^{-1}\tilde{N}_1^T\tilde{w}_1 + S_N^{-1}\tilde{N}_2^T\tilde{w}_2. \quad (3.65)$$

By (3.60) the first term in (3.65) is given by

$$(S_N X_N)^{-1/2}F^T [W^T - (GG^T)^{-1}V(I + (GG^T)^{-1}V)^{-1}W^T] D_B^{-1}u^*. \quad (3.66)$$

Note that $W = O(\mu^{1/2})$ by the preceding discussion. As for the second term in brackets, we have both $(GG^T)^{-1}$ and V are $O(\mu)$, which implies the whole second

term is $O(\mu^{5/2})$. Thus, the expression in brackets is $O(\mu^{1/2})$. By Proposition 3.2.5, $(S_N X_N)^{-1/2}$ is $O(\mu^{-1/2})$, whereas both F^T and D_B^{-1} are $O(\mu^{1/2})$. We therefore conclude that (3.66) is $O(\mu)$.

For the second term in (3.65), we use (3.57) together with (3.60):

$$-(S_N X_N)^{-1/2} H^T (H H^T)^{-1} H F^T \tilde{w}_1 = -(S_N X_N)^{-1/2} P_H F^T \tilde{w}_1. \quad (3.67)$$

Note that $\tilde{w}_1 = O(\mu)$ by the preceding arguments. The fact that $\|P_H\| \leq 1$ together with $(S_N X_N)^{-1/2}$ being $O(\mu^{-1/2})$ and F^T being $O(\mu^{1/2})$ implies (3.67) is $O(\mu)$.

Therefore, we conclude that the lower part of (3.61) is $O(\mu)$. Combining this result with (3.64) yields the following:

$$r := (1/\lambda^*) \begin{bmatrix} (S_B X_B)^{-1/2} P_G (S_B X_B)^{1/2} X_B^{-1} u^* + O(\mu^2) \\ O(\mu) \end{bmatrix}. \quad (3.68)$$

Consequently, we need to evaluate the L_∞ -norm of (3.68) and take its reciprocal. Observe that $X_B^{-1} = (X_B^*)^{-1} + O(\mu)$ by Proposition 3.2.4. Using this, we derive an upper bound on the L_∞ -norm of (3.68).

$$\|r\|_\infty \leq |1/\lambda^*| \left(\|(S_B X_B)^{-1/2}\|_\infty \|P_G\|_\infty \|(S_B X_B)^{1/2}\|_\infty \|(X_B^*)^{-1} u^*\|_\infty + O(\mu) \right). \quad (3.69)$$

Since $(x, y, s) \in \mathcal{N}_{-\infty}(\gamma)$,

$$x_i s_i = \mu n - \sum_{j \neq i} x_j s_j \leq \mu n - \mu(n-1)\gamma = \mu(n - (n-1)\gamma). \quad (3.70)$$

Thus, $(x_i s_i)^{1/2} \leq (\mu(n - (n-1)\gamma))^{1/2}$ and $(x_i s_i)^{-1/2} \leq 1/(\gamma\mu)^{1/2}$. Furthermore, since $\|P_G\| \leq 1$, we have $\|P_G\|_\infty \leq k^{1/2}$ (see e.g. [20]), where $|\mathcal{B}| = k$. Finally, since u^* is

optimal for (SA1), $\|(X_B^*)^{-1}u^*\|_\infty \leq 1$. Therefore,

$$\begin{aligned} \|r\|_\infty &\leq \frac{1}{|\lambda^*|} \left(\frac{1}{(\gamma\mu)^{1/2}} k^{1/2} (\mu(n - (n-1)\gamma))^{1/2} + O(\mu) \right), \\ &= \frac{1}{|\lambda^*|} \left(\sqrt{\frac{(n - (n-1)\gamma)k}{\gamma}} + O(\mu) \right). \end{aligned} \quad (3.71)$$

We conclude that the interior-point bound, which is the reciprocal of (3.71), is then bounded below by

$$\frac{1}{\|r\|_\infty} \geq \frac{\sqrt{\gamma}}{\sqrt{(n - (n-1)\gamma)k} + O(\mu)} |\lambda^*|. \quad (3.72)$$

Note, in particular, that the lower bound tends to $1/\sqrt{k}$, independent of n , as $\mu \rightarrow 0$ if (x, y, s) is on the central path.

We next consider the case where Δb is not in the range space of B . Again, in this case, the symmetrized bounds as well as the optimal partition bounds are all 0. The QR factorization of B can be rewritten as $B = QR = [Q_1 \ Q_2]R = Q_1R_1$, where we use (3.49) and $[Q_1 \ Q_2]$ is the appropriate partition of Q . Since Q is orthogonal, Δb can uniquely be expressed as

$$\Delta b = Q_1v_1 + Q_2v_2, \quad (3.73)$$

where v_2 is nonzero. Arguing similarly to Section 3.4, we need to evaluate (3.29), which in turn requires the computation of

$$w = (AD^2A^T)^{-1}(Q_1v_1 + Q_2v_2) \quad \text{or} \quad AD^2A^T w = Q_1v_1 + Q_2v_2. \quad (3.74)$$

Premultiplying (3.74) by Q^T leads to

$$\begin{bmatrix} R_1 & \tilde{N}_1 \\ 0 & \tilde{N}_2 \end{bmatrix} \begin{bmatrix} D_B^2 & 0 \\ 0 & D_N^2 \end{bmatrix} \begin{bmatrix} R_1^T & 0 \\ \tilde{N}_1^T & \tilde{N}_2^T \end{bmatrix} \begin{bmatrix} \tilde{w}_1 \\ \tilde{w}_2 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \quad (3.75)$$

which looks like (3.45). Essentially the same arguments as in Section 3.4 reveal that the interior-point bound tends to 0 as μ approaches 0.

Therefore, we have proved the following theorem.

Theorem 3.5.1 *Let $(x, y, s) \in \mathcal{N}_{-\infty}(\gamma)$ be a primal-dual strictly feasible point for (P) and (D) with duality gap μn . Assume that (P) has multiple optimal solutions and that b is replaced by $b + t\Delta b$ where $t \in \mathbb{R}$ and $\Delta b \in \mathbb{R}^m$. If the strictly feasible solution given by Proposition 3.2.4 is used for symmetrization in (SA1), then the ratio of the interior-point bound evaluated at (x, y, s) to the value of the optimal solution of (SA1) is at least*

$$\frac{\sqrt{\gamma}}{\sqrt{(n - (n - 1)\gamma)k} + O(\mu)}, \quad (3.76)$$

where $k = |\mathcal{B}|$.

□

Note that the presence of multiple primal optimal solutions implies $k > 0$, therefore, the expression (3.76) is well-defined. As in Section 3.4, Theorem 3.5.1 leads to the following corollary by the discussion in Section 3.3. Due to the interchange of the roles of the basic and nonbasic variables in the reformulation given in Section 3.3, k in the denominator of (3.76) is replaced by $(n - k)$. Under the assumption of multiple dual optimal solutions, Proposition 3.2.1 indicates that $m > r$, which implies $k < n$ since A has full row rank.

Corollary 3.5.1 *Let $(x, y, s) \in \mathcal{N}_{-\infty}(\gamma)$ be a primal-dual strictly feasible point for (P) and (D) with duality gap μn . Assume that (D) has multiple optimal solutions*

and that c is replaced by $c + t\Delta c$ where $t \in \mathbb{R}$ and $\Delta c \in \mathbb{R}^n$. If the strictly feasible solution given by Proposition 3.2.4 is used for symmetrization in (SA2), then the ratio of the interior-point bound evaluated at (x, y, s) to the value of the optimal solution of (SA2) is at least

$$\frac{\sqrt{\gamma}}{\sqrt{(n - (n - 1)\gamma)(n - k) + O(\mu)}}, \quad (3.77)$$

where $k = |\mathcal{B}|$.

□

3.6 Conclusion

In this paper, we have studied the quality of the bounds arising from the interior-point perspective on sensitivity analysis developed by the authors in [71] (Chapter 2). By relaxing the strong assumption of nondegeneracy, we have been able to consider all possible degeneracy scenarios and to investigate how our bounds compare with those arising from the optimal partition approach to sensitivity analysis.

If the primal problem has a degenerate but unique optimal solution, then our approach yields the same bounds as the “symmetrized” optimal partition bounds for perturbations of b . By the equivalence discussed in Section 3.3, the same relationship holds for perturbations of c if the dual problem has a degenerate but unique optimal solution. This result directly extends the previous result proved in [71] under the assumption of a unique and nondegenerate solution.

We then considered general degenerate LPs. In this case, we were able to show

that our interior-point approach would yield bounds that are at least a certain fraction of the symmetrized bounds, where the fraction depends on certain characteristics of the problem instance and of the iterate at which the bounds are evaluated. Our extensive computational tests suggest that the ratio in practice is much better than the predicted worst-case ratio. Although this result is not as strong as the aforementioned results, our interior-point bounds still yield some useful information on the range of allowable perturbations. The fact that the cost of the evaluation of our bounds is simply the same as that of an interior-point iteration makes it more appealing given the cost of solving two LPs to obtain the range from the optimal partition approach.

Chapter 4

An Interior-Point Perspective on Sensitivity Analysis in Semidefinite Programming

4.1 Introduction

Having solved an optimization problem, one is often concerned with how the optimal solution would be affected by changes in the input parameters of the original problem. This is called sensitivity analysis or post-optimality analysis, which is precisely the subject of this paper.

For linear programming (LP), there is a vast amount of literature on sensitivity analysis. Traditionally, the question has been studied from the perspective of an optimal basis matrix, for which the simplex method is well-suited. This perspective

and an alternative optimal partition approach to sensitivity analysis are described in more detail in Section 4.2.1.

Motivated by the theoretical and practical efficiency of interior-point methods, Yildirim and Todd [71] (Chapter 2) investigated the possibility of performing sensitivity analysis relying entirely on interior-point methods. They developed and analyzed an interior-point perspective on sensitivity analysis in LP and extended it to semidefinite programming (SDP). More precisely, they propose bounds (henceforth interior-point bounds) which are functions of the original problem instance, the perturbation and the feasible point at which they are evaluated, using only the matrix factorizations necessary to carry out an interior-point iteration from this solution. For LP, these bounds are shown to have a nice asymptotic relationship with the optimal partition bounds under nondegeneracy and a specific kind of degeneracy; they still provide useful information in the case of general degeneracy [71, 70] (Chapters 2 and 3).

In contrast to the case for LP, sensitivity analysis in SDP is a relatively new topic. Fundamental differences between the problem structures of LP and SDP necessitate the development of an appropriate way to perform sensitivity analysis in SDP. Goldfarb and Scheinberg [18] propose an optimal partition approach, which has close connections with a related approach in LP. Sturm and S. Zhang [59] study the sensitivity of the central path under perturbations of the right-hand side. Nunez and Freund [51] study the properties of the central path and propose bounds on its behavior under general data perturbations.

Our goal in this paper is to study the asymptotic behavior of the interior-point

bounds of [71] for SDP in comparison with the optimal partition bounds of [18]. For perturbations of the right-hand side vector and the cost matrix, we show that the interior-point bounds evaluated on the central path using the Monteiro-Zhang family of search directions converge to the symmetrized version of the optimal partition bounds (symmetrized partition bounds) under appropriate nondegeneracy assumptions. Furthermore, under certain assumptions, this asymptotic coincidence continues to hold even when strict complementarity fails – a phenomenon that does not happen in LP. We also extend the same convergence results to iterates lying in an appropriate (but very narrow) neighborhood of the central path if the Nesterov-Todd direction is used to evaluate the interior-point bounds.

This paper is organized as follows. In Section 4.1.1, we define the notation that will be used throughout the paper. Section 4.2 introduces SDP and some preliminaries that will be relevant for the analysis. We also outline the optimal partition approach of [18] to sensitivity analysis in SDP and review the interior-point bounds of [71] in this section. The symmetry between the primal and dual problems is discussed in Section 4.3, justifying our consideration of perturbations of the right-hand side vector only. We analyze the asymptotic behavior of the interior-point bounds on the central path using the Nesterov-Todd direction in comparison with the symmetrized partition bounds and then extend the same results to the Monteiro-Zhang family of search directions in Section 4.4. Section 4.5 discusses the extension of the asymptotic results of Section 4.4 to an appropriate central path neighborhood using the Nesterov-Todd direction. We conclude the paper with some remarks in Section 4.6.

4.1.1 Notation

We use the following notation throughout this paper. \mathbb{R}^m will denote m -dimensional Euclidean space and \mathcal{S}^n will denote the set of $n \times n$ symmetric matrices. Upper-case Roman letters will be reserved for matrices. Lower-case Greek letters will always denote scalars. I will denote the identity matrix in the appropriate dimension and Q will be reserved for orthogonal matrices. We will use the trace inner product on $\mathbb{R}^{k \times n}$ defined as

$$U \bullet V := \text{Trace}(U^T V) = \sum_{i,j} U_{ij} V_{ij}$$

for $U, V \in \mathbb{R}^{k \times n}$, where U_{ij} denotes the (i, j) entry of U . Associated with the trace inner product is the Frobenius norm defined as $\|P\|_F := (P \bullet P)^{1/2}$. For $X \in \mathcal{S}^n$, $\lambda(X)$ denotes the spectrum, or the set of eigenvalues. $X \succ 0$ ($X \succeq 0$) will indicate that X is positive (semi)definite, i.e., all entries of $\lambda(X)$ are positive (nonnegative). For $X \succeq 0$, $X^{1/2} \in \mathcal{S}^n$ will be the unique positive semidefinite square root of X . We use $Y \succeq Z$ if $Y - Z \succeq 0$. We will denote by $\|\cdot\|$ the Euclidean norm and the operator norm induced by it. We use script letters to denote linear operators on symmetric matrices. We also define

$$n^{\bar{2}} := \frac{1}{2}n(n+1).$$

Finally, we use the standard big O and related notation. If $u(\mu)$ and $w(\mu)$ are two functions defined for $\mu > 0$ with $w(\mu) > 0$, then $u(\mu) = O(w(\mu))$ indicates that $u(\mu)/w(\mu)$ is bounded independent of μ as $\mu \downarrow 0$. If $u(\mu)/w(\mu)$ is bounded away from 0 as $\mu \downarrow 0$, then we write $u(\mu) = \Omega(w(\mu))$. Finally, $u(\mu) = \Theta(w(\mu))$ iff (if and only if) $u(\mu) = O(w(\mu))$ and $u(\mu) = \Omega(w(\mu))$. We will use $O(w(\mu))$ for a vector or

matrix to indicate that each entry is $O(w(\mu))$.

4.2 Semidefinite Programming

We consider the SDP given in the following standard form:

$$\begin{aligned} \text{(P)} \quad \min_X \quad & C \bullet X \\ & A_i \bullet X = b_i, \quad i = 1, \dots, m, \\ & X \succeq 0, \end{aligned}$$

where all $A_i \in \mathcal{S}^n$, $b \in \mathbb{R}^m$, $C \in \mathcal{S}^n$ are given, and $X \in \mathcal{S}^n$. The dual problem associated with (P) is:

$$\begin{aligned} \text{(D)} \quad \max_{y,S} \quad & b^T y \\ & \sum_{i=1}^m y_i A_i + S = C, \\ & S \succeq 0, \end{aligned}$$

where $y \in \mathbb{R}^m$ and $S \in \mathcal{S}^n$.

For feasible (X, y, S) , the duality gap is given by

$$C \bullet X - b^T y = X \bullet S = \text{Trace}(XS) = \text{Trace}(X^{1/2} S^{1/2} S^{1/2} X^{1/2}) = \|X^{1/2} S^{1/2}\|_F^2 \geq 0,$$

where we used the fact that $\text{Trace}(AB) = \text{Trace}(BA)$ for $A, B \in \mathbb{R}^{n \times n}$. We make the following assumptions throughout this paper; these assumptions guarantee that both (P) and (D) have bounded optimal solutions and that the duality gap is 0 at any primal-dual optimal solution.

Assumption 4.2.1 *The primal-dual Slater condition is satisfied, i.e., (P) and (D) have feasible solutions X and (y, S) with $X \succ 0$ and $S \succ 0$. Such solutions will be called strictly feasible.*

Assumption 4.2.2 *The set $\{A_i, i = 1, \dots, m\}$ is linearly independent.*

Under Assumptions 4.2.1 and 4.2.2, the central path is defined as the set of unique solutions $(X(\mu), y(\mu), S(\mu))$ for each $\mu > 0$ to the following system together with the requirement that X and S be symmetric positive definite:

$$\begin{aligned} A_i \bullet X &= b_i, \quad \text{for } i = 1, \dots, m, \\ \sum_{i=1}^m y_i A_i &+ S = C, \\ XS &= \mu I. \end{aligned} \tag{4.1}$$

As $\mu \downarrow 0$, Goldfarb and Scheinberg [17] claimed that the central path converges to the so-called “analytic center” of the primal-dual optimal set. Recently, Halicka, de Klerk and Roos [27] provided a counterexample in the absence of strict complementarity. However, the central path does converge to a point in the relative interior of the primal-dual optimal set [17, 27], which we denote by (X^*, y^*, S^*) . Let \mathbf{O}_P (\mathbf{O}_P^0) and \mathbf{O}_D (\mathbf{O}_D^0) denote (the relative interiors of) the primal and dual optimal sets, respectively. The following lemma, due to Barker and Carlson [7], shows that any optimal solution in the relative interior of the optimal set will have maximal range space.

Lemma 4.2.1 *For any $\tilde{X} \in \mathbf{O}_P$ ($(\tilde{y}, \tilde{S}) \in \mathbf{O}_D$) and any $\hat{X} \in \mathbf{O}_P^0$ ($(\hat{y}, \hat{S}) \in \mathbf{O}_D^0$), the range space of \tilde{X} (\tilde{S}) is a subset of the range space of \hat{X} (\hat{S}).*

The analytic center lies in the relative interior of the primal and dual optimal sets [17, Lemma 4.2]. By Lemma 4.2.1, both X^* and S^* then have respectively maximal ranks r and s among all the primal-dual optimal solutions. Since (X^*, y^*, S^*) satisfies $X^*S^* = S^*X^* = 0$, X^* and S^* then share a common set of eigenvectors

and hence can be diagonalized simultaneously. By applying an orthogonal transformation to the problem data if necessary, we can assume that both X^* and S^* are diagonal and given by

$$X^* = \begin{bmatrix} \Lambda^* & 0 \\ 0 & 0 \end{bmatrix}, \quad S^* = \begin{bmatrix} 0 & 0 \\ 0 & \Omega^* \end{bmatrix}, \quad (4.2)$$

where $\Lambda^* \in \mathcal{S}^r$ and $\Omega^* \in \mathcal{S}^s$ with Λ^* and Ω^* positive definite. Clearly, $r + s \leq n$. We will say that *strict complementarity* holds if $r + s = n$. Consequently, the two partitions in (4.2) are identical iff strict complementarity holds. In contrast to the case for linear programming, a strictly complementary optimal solution might not exist for SDP.

Primal-dual path-following interior-point methods for SDP are iterative algorithms which use a Newton-like method to generate search directions to find an approximate solution to the nonlinear system (4.1). Starting from an iterate $(\tilde{X}, \tilde{y}, \tilde{S})$ with $\tilde{X} \succ 0$ and $\tilde{S} \succ 0$ (not necessarily feasible), a (damped) Newton-like step is taken (to ensure that X and S components remain positive definite) to obtain the next iterate. The same procedure is then repeated at this new iterate by appropriately decreasing μ towards 0. For SDP, however, unlike linear programming, Newton's method cannot directly be applied to (4.1) since the residual map takes an iterate $(X, y, S) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$ to a point in $\mathbb{R}^m \times \mathcal{S}^n \times \mathbb{R}^{n \times n}$ (since $XS - \mu I$ is in general not symmetric), which is a space of higher dimension. Many authors have suggested different ways of symmetrizing the third equation in (4.1) so that the residual lies in \mathcal{S}^n . Todd [60] studies twenty different search directions for SDP.

Next, we introduce some notation that we will use throughout this paper. We

will define $\mathcal{A} : \mathcal{S}^n \rightarrow \mathbb{R}^m$ by

$$\mathcal{A}U := (A_i \bullet U)_{i=1}^m, \quad (4.3)$$

with adjoint $\mathcal{A}^* : \mathbb{R}^m \rightarrow \mathcal{S}^n$; then

$$\mathcal{A}^*y = \sum_{i=1}^m y_i A_i. \quad (4.4)$$

We will also use the following notation introduced by Alizadeh, Haeberly, and Overton [5]:

$$(P \odot Q)K := \frac{1}{2}(PKQ^T + QKP^T),$$

where $P, Q \in \mathbb{R}^{n \times n}$ and $K \in \mathcal{S}^n$, and we will regard it as an operator from \mathcal{S}^n to itself. The adjoint operator is defined as usual by $\mathcal{E}^*U \bullet V = U \bullet \mathcal{E}V$ for all $U, V \in \mathcal{S}^n$, and it is easy to see that

$$Q \odot P = P \odot Q, \quad (P \odot Q)^* = P^T \odot Q^T,$$

so that $P \odot Q$ is self-adjoint if P and Q are symmetric. If moreover P and Q are positive definite, then

$$(P \odot Q)U \bullet U = \text{Trace}(PUQU) = \text{Trace}(P^{1/2}UQ^{1/2}Q^{1/2}UP^{1/2}) = \|P^{1/2}UQ^{1/2}\|_F^2,$$

so that $P \odot Q$ is also positive definite. If P is nonsingular,

$$(P \odot P)^{-1} = P^{-1} \odot P^{-1},$$

but there is no simple expression for $(P \odot Q)^{-1}$ in general. Note that $I \odot I$ is the identity operator. Occasionally, we will extend the domain of $P \odot Q$ to $\mathbb{R}^{n \times n}$, defining

$$(P \odot Q)Z := \frac{1}{2}(PZQ^T + QZ^T P^T)$$

for $Z \in \mathbb{R}^{n \times n}$.

In this paper, we will consider the Monteiro-Zhang family of search directions [42, 74, 46, 43] which uses the following symmetrization scheme. Let $P \in \mathbb{R}^{n \times n}$ be a nonsingular matrix. In (4.1), replace $XS = \mu I$ by

$$H_P(XS) = H_P(\mu I) = \mu I, \quad (4.5)$$

where $H_P := P \odot P^{-T}$ so that for $U \in \mathbb{R}^{n \times n}$,

$$H_P(U) = \frac{1}{2} (PUP^{-1} + P^{-T}U^T P^T). \quad (4.6)$$

Pre- and post-multiplying (4.5) by P^T and P yields

$$\frac{1}{2} (MXS + SXM) = \mu M, \quad (4.7)$$

where $M := P^T P$. Different choices of M give rise to different search directions in the Monteiro-Zhang family. For example, $M = I$ gives the AHO direction suggested by Alizadeh, Haeberly and Overton [5]. The H..K..M direction, which was independently introduced by Helmberg, Rendl, Vanderbei and Wolkowicz [29] and Kojima, Shindoh and Hara [39], and rediscovered from this viewpoint by Monteiro [42], uses $M = S$. The dual H..K..M direction [39, 42] is given by $M = X^{-1}$. Finally, $M = W^{-1}$, where W is the unique positive definite scaling matrix satisfying $WSW = X$, yields the NT direction proposed by Nesterov and Todd [48, 49].

With this symmetrization scheme, Newton's method can now be applied to approximately find a solution to the nonlinear system (4.1). The Newton direction at a given point (X, y, S) with $X \succ 0$ and $S \succ 0$ will be given by the solution of

the following system:

$$\begin{aligned}
\mathcal{A}\Delta X &= r_p, \\
\mathcal{A}^*\Delta y + \Delta S &= R_d, \\
\mathcal{E}\Delta X + \mathcal{F}\Delta S &= R_{EF},
\end{aligned} \tag{4.8}$$

where $r_p := b - \mathcal{A}X$ is the primal residual, $R_d := C - \mathcal{A}^*y - S$ is the dual residual, $R_{EF} = R_{EF}(X, S) := \mu M - 1/2(MXS + SXM)$, and the operators \mathcal{E} and \mathcal{F} are given by

$$\mathcal{E} := S \odot M, \quad \mathcal{F} := MX \odot I. \tag{4.9}$$

Newton's method, however, can also be used to approximate a different feasible point (X', y', S') than the one on the central path, as in the context of target-following methods. It suffices to redefine

$$R_{EF} = 1/2(MX'S' + S'X'M) - 1/2(MXS + SXM).$$

In this case, it is enough to know the product $X'S'$ rather than X' and S' separately.

Assume that \mathcal{E} is nonsingular. The linear operator $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$ maps \mathbb{R}^m to itself. Therefore, it can be represented by an $m \times m$ matrix, called the Schur complement. We find that (4.8) has a unique solution iff the $m \times m$ Schur complement matrix $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$ is nonsingular, and in this case the solution is given by

$$\begin{aligned}
(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)\Delta y &= r_p - \mathcal{A}\mathcal{E}^{-1}(R_{EF} - \mathcal{F}R_d), \\
\Delta S &= R_d - \mathcal{A}^*\Delta y, \\
\Delta X &= \mathcal{E}^{-1}(R_{EF} - \mathcal{F}\Delta S).
\end{aligned} \tag{4.10}$$

We will use the following convention for partitioning a matrix $K \in \mathcal{S}^n$:

$$K = \begin{bmatrix} K_P & K_U^T \\ K_U & K_N \end{bmatrix}, \quad (4.11)$$

where $K_P \in \mathcal{S}^r$, $K_N \in \mathcal{S}^{n-r}$ and $K_U \in \mathbb{R}^{(n-r) \times r}$.

We next discuss primal-dual nondegeneracy in SDP as introduced by Alizadeh, Haeberly and Overton [4], which includes a detailed analysis of the subject. We simply give the definitions here:

Definition 4.2.1 *The primal optimal solution X^* given by (4.2) is nondegenerate if the matrices*

$$B_i := \begin{bmatrix} (A_i)_P & (A_i)_U^T \\ (A_i)_U & 0 \end{bmatrix}, \quad i = 1, \dots, m \quad (4.12)$$

are linearly independent in \mathcal{S}^n .

Definition 4.2.2 *The dual optimal solution (y^*, S^*) , where S^* is given by (4.2), is nondegenerate if the matrices given by the first $n - s$ rows and columns of A_i , $i = 1, \dots, m$, span \mathcal{S}^{n-s} .*

For the purposes of this paper, it turns out that we actually need a somewhat weaker condition than dual nondegeneracy. It will be sufficient if the matrices

$$Y_i := (A_i)_P, \quad i = 1, \dots, m \quad (4.13)$$

span \mathcal{S}^r . Clearly, this condition will be satisfied if dual nondegeneracy condition is satisfied at (y^*, S^*) . We will call this condition *weak dual nondegeneracy*. Weak

dual nondegeneracy coincides with dual nondegeneracy if strict complementarity holds.

Similarly to weak dual nondegeneracy, we say that *weak primal nondegeneracy* holds at X^* if the matrices obtained from the A_i by replacing the bottom right $s \times s$ block by the zero matrix are linearly independent in \mathcal{S}^n . Again, weak primal nondegeneracy is implied by primal nondegeneracy and it coincides with primal nondegeneracy under strict complementarity.

The following SDP instance illustrates that weak primal nondegeneracy and weak dual nondegeneracy are indeed weaker requirements than primal and dual nondegeneracy.

Example 4.2.1 Consider the following SDP instance with $m = 3$, $n = 3$, $b = [1 \ 0 \ 0]^T$:

$$C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

This instance satisfies Assumptions 4.2.1 and 4.2.2 and has the unique optimal solution

$$X^* = \text{diag}(1, 0, 0), \quad y^* = [0 \ 0 \ 0]^T, \quad S^* = \text{diag}(0, 0, 1),$$

where diag denotes a diagonal matrix with the corresponding entries on the diagonal. Strict complementarity does not hold. The dual optimal solution is not

nondegenerate since the matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

do not span \mathcal{S}^2 . However, weak dual nondegeneracy is satisfied since the 1×1 matrices $[1]$, $[0]$, $[0]$ clearly do span \mathcal{S}^1 . Similarly, the primal optimal solution is not nondegenerate since the matrices

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

include the zero matrix and hence are not linearly independent in \mathcal{S}^3 . However, weak primal nondegeneracy is satisfied since

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

are linearly independent in \mathcal{S}^3 .

We make the following assumption throughout the rest of the paper when dealing with perturbations of the right-hand side vector b :

Assumption 4.2.3 *The optimal solution (X^*, y^*, S^*) given by (4.2) satisfies primal nondegeneracy and weak dual nondegeneracy.*

The corresponding assumption for perturbations of the cost matrix C is given by:

Assumption 4.2.4 *The optimal solution (X^*, y^*, S^*) given by (4.2) satisfies weak primal nondegeneracy and dual nondegeneracy.*

By [4, Theorem 7], Assumption 4.2.3 implies that (y^*, S^*) is the unique dual optimal solution. By Lemma 4.2.1, X^* has the maximal range space. Consequently, a similar argument as in [4, Theorem 10] reveals that X^* is the only primal optimal solution. It can be shown similarly that (X^*, y^*, S^*) is the unique optimal solution under Assumption 4.2.4. Therefore, Assumption 4.2.3 or 4.2.4 provides weaker sufficient conditions than usual nondegeneracy for the existence of a unique primal-dual optimal solution.

4.2.1 Sensitivity Analysis in SDP

Sensitivity analysis in linear programming has traditionally been associated with an optimal basis. More specifically, the question has been studied from the perspective of how much data perturbation can be introduced before an optimal basis for the original problem is no longer optimal. The presence of multiple optimal bases poses a problem with this approach since each basis might yield different information.

This shortcoming has been observed by several researchers. Adler and Monteiro [2] and Jansen, de Jong, Roos and Terlaky [32] proposed an alternative approach to sensitivity analysis based on the concept of *optimal partition* in linear programming. Here, optimal partition refers to the partition of the components of the primal variable into two sets according to whether or not there exists an optimal primal solution with the corresponding component strictly positive.

The optimal basis approach cannot directly and practically be extended to SDP. However, it is possible to extend the idea of the optimal partition to SDP, as shown by Goldfarb and Scheinberg [18]. Let optimal X^* and S^* be given by (4.2). Let \mathfrak{R}_P and \mathfrak{R}_D denote the range spaces of X^* and S^* , respectively. Let us denote by \mathfrak{R}_N the orthogonal complement of $\mathfrak{R}_P \oplus \mathfrak{R}_D$. Note that $\mathfrak{R}_N = \{0\}$ iff strict complementarity holds. Furthermore, any two of the three subspaces are mutually orthogonal. Then, $(\mathfrak{R}_P, \mathfrak{R}_N, \mathfrak{R}_D)$ have \mathbb{R}^n as their direct sum and will be called the optimal partition.

If we assume that the right-hand side vector b is replaced by $b + \gamma\Delta b$, one possible approach is to determine the range of γ for which the optimal partition remains constant. Since perturbation of b does not affect the dual feasible region, the problem reduces to finding a new primal solution for the perturbed problem whose range is \mathfrak{R}_P . However, such a solution should be given by

$$X = \begin{bmatrix} X_P & 0 \\ 0 & 0 \end{bmatrix}, \quad (4.14)$$

where $X_P \succ 0$. Therefore, the problem above can be formulated as an SDP in the following way:

$$\begin{aligned} \text{(AUX)} \quad & \min_{\gamma, X_P} (\max_{\gamma, X_P}) \quad \gamma \\ & Y_i \bullet X_P = b + \gamma\Delta b_i, \quad i = 1, \dots, m, \\ & X_P \succeq 0, \end{aligned}$$

where the Y_i are given by (4.13). Since X^* is feasible for (P), a change of variable

reduces (AUX) to

$$\begin{aligned}
 \text{(AUX)} \quad \min_{\gamma, U}(\max_{\gamma, U}) \quad & \gamma \\
 & Y_i \bullet U = \gamma \Delta b_i, \quad i = 1, \dots, m, \\
 & U \succeq -\Lambda^*.
 \end{aligned}$$

Goldfarb and Scheinberg [18, Lemma 4.1] prove that the optimal partition does indeed remain the same for the values of γ in the range given by the optimal values of (AUX).

Since the Y_i span \mathcal{S}^r by Assumption 4.2.3, it follows that the linear operator defining the linear constraints of (AUX) is injective. Therefore, either there exists a unique $\tilde{V} \in \mathcal{S}^r$ such that $Y_i \bullet \tilde{V} = \Delta b_i$, $i = 1, \dots, m$, or the range of γ given by the optimal values of (AUX) is clearly $\{0\}$. In the former case, the constraints of (AUX) can be further simplified to

$$\gamma \tilde{V} \succeq -\Lambda^*. \quad (4.15)$$

We will call the minimum and maximum values of γ for which (4.15) is satisfied the *optimal partition bounds*. Note that (4.15) can be reduced to an eigenvalue inequality since it is equivalent to

$$\gamma(\Lambda^*)^{-1/2} \tilde{V} (\Lambda^*)^{-1/2} \succeq -I, \quad (4.16)$$

which can be rewritten as $\lambda_{\min} \left(\gamma(\Lambda^*)^{-1/2} \tilde{V} (\Lambda^*)^{-1/2} \right) \geq -1$, where $\lambda_{\min}(\cdot)$ denotes the minimum eigenvalue function.

We will next symmetrize the eigenvalue bound above by requiring also that the maximum eigenvalue be at most 1. The symmetrized eigenvalue bound can

conveniently be expressed as

$$\left\| \gamma(\Lambda^*)^{-1/2} \tilde{V}(\Lambda^*)^{-1/2} \right\| \leq 1. \quad (4.17)$$

The range of values of γ for which (4.17) continues to hold will be called the *symmetrized partition bounds*. It is easy to verify that if the optimal partition bounds are given by (γ^-, γ^+) , then the symmetrized partition bounds are given by $(-\gamma^s, +\gamma^s)$, where

$$\gamma^s = \min\{|\gamma^-|, \gamma^+\}.$$

Clearly, γ^s is precisely the reciprocal of

$$\left\| (\Lambda^*)^{-1/2} \tilde{V}(\Lambda^*)^{-1/2} \right\|. \quad (4.18)$$

We finally comment on the differences between the optimal partition approaches in LP and SDP. In LP, maintaining the optimal partition is equivalent to maintaining the whole dual optimal set invariant under perturbations of b . A similar motivation continues to hold in SDP iff strict complementarity holds. However, if $\mathfrak{R}_N \neq \{0\}$, then the equivalence no longer holds since, in this case, any matrix with range space in $\mathfrak{R}_P \oplus \mathfrak{R}_N$ can be optimal for the perturbed problem without altering the dual optimal set.

4.2.2 Interior-Point Approach

In [71], Yildırım and Todd developed an interior-point perspective on sensitivity analysis for LP and extended it to SDP. We briefly review the approach in the SDP case in this subsection.

We assume that (X, y, S) is primal-dual strictly feasible and near-optimal (with a small duality gap). After a perturbation in the right-hand side or in the cost matrix is introduced, the goal is to retrieve a feasible and near-optimal point for the perturbed problem by taking a full Newton step at (X, y, S) . A target-following approach is adopted and the Newton step proposed in [71] arises from trying to approximate a feasible point (X', y', S') of the perturbed problem for which $X'S' = XS$. By the discussion following (4.9), this is equivalent to setting $R_{EF} = 0$ in (4.8). Basically, the motivation for this particular choice stems from the fact that $XS \approx 0$. Therefore, it is meaningful to approximate a point with the same product in order to obtain a near-optimal solution. Furthermore, this choice guarantees a lower duality gap than that of the original iterate if the search direction used satisfies a technical property. The interior-point bound then is defined to be the range of perturbations for which such a Newton step can successfully be taken.

We next present the general results obtained in [71] using the Monteiro-Zhang family of search directions for perturbations of the right-hand side vector and the cost matrix. The reader is referred to [71] for the corresponding proofs.

Proposition 4.2.1 *Assume that (X, y, S) is a strictly feasible point for (P) and (D) and let \mathcal{E} and \mathcal{F} as in (4.8) be given by (4.9). Assume that $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$ is nonsingular. Let the right-hand side vector b be replaced by $b + \gamma\Delta b$, where $\gamma \in \mathbb{R}$ and $\Delta b \in \mathbb{R}^m$. Suppose that a Newton step is taken from (X, y, S) to regain feasibility for the perturbed problem using $R_{EF} = 0$ in (4.8). Then a full Newton step can be taken and the resulting iterate will be feasible for the new problem if, and only if, γ*

satisfies the following:

$$|\gamma| \leq \min\{a, b\}, \quad (4.19)$$

where a is the reciprocal of

$$-\lambda_{\min} \left(X^{-\frac{1}{2}} (\mathcal{E}^{-1} \mathcal{F} \mathcal{A}^* [(\mathcal{A} \mathcal{E}^{-1} \mathcal{F} \mathcal{A}^*)^{-1} \Delta b]) X^{-\frac{1}{2}} \right)$$

(or $+\infty$ if this quantity is negative) and b that of

$$\lambda_{\max} \left(S^{-\frac{1}{2}} (\mathcal{A}^* [(\mathcal{A} \mathcal{E}^{-1} \mathcal{F} \mathcal{A}^*)^{-1} \Delta b]) S^{-\frac{1}{2}} \right)$$

(or, again, $+\infty$ if this quantity is negative). Moreover, the duality gap of the new iterate will be at most $X \bullet S$ if $\mathcal{E}^{-1} \mathcal{F}$ is positive definite.

Proposition 4.2.2 *Assume that (X, y, S) is a strictly feasible pair of points for (P) and (D) and let \mathcal{E} and \mathcal{F} as in (4.8) be given by (4.9). Assume that $\mathcal{A} \mathcal{E}^{-1} \mathcal{F} \mathcal{A}^*$ is nonsingular. Let the cost matrix C be replaced by $C + \gamma \Delta C$, where $\gamma \in \mathbb{R}$ and $\Delta C \in \mathcal{S}^n$. Suppose that a Newton step is taken from (X, y, S) to regain feasibility for the perturbed problem using $R_{EF} = 0$ in (4.8). Then a full Newton step can be taken and the resulting iterate will be feasible for the new problem if, and only if, γ satisfies the following:*

$$|\gamma| \leq \min\{a, b\}, \quad (4.20)$$

where a is the reciprocal of

$$\lambda_{\max} \left(X^{-\frac{1}{2}} (\mathcal{E}^{-1} \mathcal{F} \Delta C - \mathcal{E}^{-1} \mathcal{F} \mathcal{A}^* (\mathcal{A} \mathcal{E}^{-1} \mathcal{F} \mathcal{A}^*)^{-1} \mathcal{A} \mathcal{E}^{-1} \mathcal{F} \Delta C) X^{-\frac{1}{2}} \right)$$

(or $+\infty$ if this quantity is negative) and b that of

$$-\lambda_{\min} \left(S^{-\frac{1}{2}} (\Delta C - \mathcal{A}^* (\mathcal{A} \mathcal{E}^{-1} \mathcal{F} \mathcal{A}^*)^{-1} \mathcal{A} \mathcal{E}^{-1} \mathcal{F} \Delta C) S^{-\frac{1}{2}} \right)$$

(or, again, $+\infty$ if this quantity is negative). Moreover, the duality gap of the new iterate will be at most $X \bullet S$ if $\mathcal{E}^{-1}\mathcal{F}$ is positive definite.

The operator $\mathcal{E}^{-1}\mathcal{F}$ is positive definite for the H..K..M, dual H..K..M, and NT directions [60]. Shida, Shindoh and Kojima [57] and Todd, Toh and Tütüncü [62] show that the AHO direction enjoys this property if $XS + SX$ is positive semidefinite, while Monteiro and Zanjácomo [45] show that the positive definiteness property also holds for the AHO direction if the iterate lies in a suitable central path neighborhood.

4.3 Primal-Dual Symmetry

Although (P) and (D) do not look symmetric, either one can actually be formulated in the form of the other. We briefly review this reformulation in this section.

Let \hat{X} satisfy the primal equality constraints. Then, (P) can be reformulated as

$$(P) \quad \min_X C \bullet X, \quad \text{s.t.} \quad X \in \hat{X} + \Sigma, \quad X \succeq 0, \quad (4.21)$$

where Σ is the null space of \mathcal{A} .

Let $G_1, \dots, G_p \in \mathcal{S}^n$ be a basis for Σ , where $p = n^2 - m$. Thus, the constraint of (P) involving Σ can be written as $\sum_{i=1}^m w_i G_i + X = \hat{X}$, where $w \in \mathbb{R}^p$. It follows then that $C \bullet X = C \bullet \hat{X} - \sum_{i=1}^p w_i C \bullet G_i$ for feasible X . Therefore, defining $\hat{b}_i := C \bullet G_i$, we obtain the following equivalent formulation of (P):

$$(P1) \quad \max_{w, X} \hat{b}^T w \quad \text{s.t.} \quad \sum_{i=1}^m w_i G_i + X = \hat{X}, \quad X \succeq 0, \quad (4.22)$$

which is exactly in the form of (D). The dual of (P1) is given by

$$(D1) \quad \min_S \hat{X} \bullet S \text{ s.t. } G_i \bullet S = \hat{b}_i, \quad i = 1, \dots, p, \quad S \succeq 0. \quad (4.23)$$

It is not hard to verify that (D) is actually equivalent to (D1), which is exactly in the same form as (P).

For a search direction, *the dual direction* is given by applying the search direction to the alternative formulation given by (D1) and (P1), i.e., by interchanging the roles of the primal and dual problems. We will say that a search direction is *primal-dual symmetric* if the ΔX and ΔS components of the search direction computed at (X, y, S) for (P) and (D) (assuming it exists) coincide with those of the dual search direction computed at (S, w, X) for (D1) and (P1). Todd [60] shows that the AHO and NT directions are primal-dual symmetric whereas the H..K..M and the dual H..K..M are the dual directions of one another.

If (P) and (D) satisfy Assumptions 4.2.1 and 4.2.2, then (D1) and (P1) also satisfy the same assumptions. Furthermore, the two formulations given above possess a nice symmetric nondegeneracy relationship. More precisely, if an optimal solution (X^*, y^*, S^*) of (P) and (D) satisfies weak primal nondegeneracy and dual nondegeneracy (Assumption 4.2.4), then the optimal solution (S^*, w^*, X^*) of (D1) and (P1) satisfies primal nondegeneracy and weak dual nondegeneracy, which is precisely Assumption 4.2.3. To see this, note that if weak primal nondegeneracy holds for (P) and (D), then we need to show that the bottom right $s \times s$ blocks of the G_i span \mathcal{S}^s . Otherwise, there exists a nonzero $P \in \mathcal{S}^s$ such that P is orthogonal

to the space spanned by the corresponding parts of the G_i . Therefore,

$$G_i \bullet \begin{bmatrix} 0 & 0 \\ 0 & P \end{bmatrix} = 0, \quad i = 1, \dots, p.$$

However, the space spanned by the G_i is precisely the null space of \mathcal{A} by definition of the G_i . Therefore, the matrix above lies in the range space of \mathcal{A}^* , i.e., there exists y_i , $i = 1, \dots, m$, such that

$$\begin{bmatrix} 0 & 0 \\ 0 & P \end{bmatrix} = \sum_{i=1}^m y_i A_i.$$

The weak primal nondegeneracy of (P) and (D) implies that $y = 0$. Therefore, $P = 0$, contradicting our assumption and proving weak dual nondegeneracy for (D1) and (P1). The other relationship is proved similarly.

The arguments above, along with those corresponding to Section 3.3, reveal that for any given search direction, a result concerning perturbations of the right-hand side vector for a problem instance satisfying Assumptions 4.2.1–4.2.3 directly carries over to perturbations of the cost matrix using the dual search direction if the problem instance satisfies Assumptions 4.2.1, 4.2.2 and 4.2.4. Therefore, if a search direction is primal-dual symmetric, any result for perturbations of the right-hand side vector also holds for perturbations of the cost matrix under the corresponding assumptions.

Based on this observation, we will consider only perturbations of the right-hand side vector b and state the corresponding results for perturbations of the cost matrix C as corollaries.

4.4 The Central Path

In this section, we start analyzing the interior-point bound outlined in Section 4.2.2 on the central path using the Nesterov-Todd (NT) direction [48, 49]. Our aim is to compare this bound as points on the central path converge to the optimal solutions to the symmetrized partition bound of Section 4.2.1. We then show the coincidence of all search directions in the Monteiro-Zhang family on the central path, thereby extending our analysis to the whole family of search directions.

The NT direction uses $M = W^{-1}$ in (4.9), where W is the unique scaling matrix given by

$$W = X^{\frac{1}{2}}(X^{\frac{1}{2}}SX^{\frac{1}{2}})^{-\frac{1}{2}}X^{\frac{1}{2}} \quad (4.24)$$

so that $WSW = X$. Therefore, the operators \mathcal{E} and \mathcal{F} as in (4.9) are given by

$$\mathcal{E} := S \odot W^{-1}, \quad \mathcal{F} := W^{-1}X \odot I. \quad (4.25)$$

An alternative formulation of the NT direction [62], in which case \mathcal{E} does not need to be inverted, is given by

$$\mathcal{E} = I \odot I, \quad \mathcal{F} = W \odot W. \quad (4.26)$$

Furthermore, Yıldırım and Todd [71] showed that the target-following approach can be generalized using this alternative formulation if R_{EF} as in (4.8) satisfies

$$(W^{-1} \odot S)R_{EF} = W^{-1}(X'S' - XS) + (S'X' - SX)W^{-1}. \quad (4.27)$$

At a given iterate (X, y, S) , if we target the point with $X'S' = XS$, we again have $R_{EF} = 0$ by (4.27).

Before we start the analysis of the NT direction, we introduce some tools that will be useful. For a given matrix $U \in \mathcal{S}^n$, we say that a k -dimensional subspace $\Sigma \subset \mathbb{R}^n$ is an invariant subspace for U if $U\Sigma \subset \Sigma$. Clearly, a subspace spanned by one or more eigenvectors of U is an invariant subspace for U . We start with the following theorem, where $\lambda(\cdot)$ denotes the set of eigenvalues. We refer the reader to [20] for a proof and further details.

Theorem 4.4.1 *Let $U \in \mathcal{S}^n$. Suppose $Q \in \mathbb{R}^{n \times n}$ is orthogonal and that it is partitioned as*

$$Q = [Q_1 \quad Q_2],$$

where $Q_1 \in \mathbb{R}^{n \times r}$ is such that the range of Q_1 is an invariant subspace for U . Then

$$Q^T U Q = D = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}, \quad (4.28)$$

where $D_1 \in \mathcal{S}^r$ and $D_2 \in \mathcal{S}^{n-r}$. Moreover, $\lambda(U) = \lambda(D_1) \cup \lambda(D_2)$.

Although eigenvalues of a symmetric matrix vary continuously, it is well-known that eigenvectors can be very sensitive to small perturbations if they are associated with repeated eigenvalues, in which case eigenvectors are not necessarily unique. However, if such eigenvalues are separated from the rest of the spectrum, then the invariant subspace spanned by the corresponding eigenvectors varies continuously. The appropriate measure of separation between the eigenvalues of $U \in \mathcal{S}^n$ and $V \in \mathcal{S}^n$ is given by

$$\text{sep}(U, V) := \min_{\lambda \in \lambda(U), \mu \in \lambda(V)} |\lambda - \mu|.$$

The following theorem is due to Stewart [58].

Theorem 4.4.2 *Suppose $U \in \mathcal{S}^n$ and $E \in \mathcal{S}^n$. Let $Q = [Q_1 \ Q_2]$ be an orthogonal matrix such that the range of $Q_1 \in \mathbb{R}^{n \times r}$ is an invariant subspace for U . Partition the matrices $Q^T U Q$ and $Q^T E Q$ as follows:*

$$Q^T U Q = D = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}, \quad Q^T E Q = \begin{bmatrix} E_{11} & E_{21}^T \\ E_{21} & E_{22} \end{bmatrix}.$$

If $\delta := \text{sep}(D_1, D_2) > 0$ and $\|E\| \leq \delta/5$, then there exists a matrix $P \in \mathbb{R}^{(n-r) \times r}$ with

$$\|P\| \leq \frac{4}{\delta} \|E_{21}\|$$

such that the columns of $\hat{Q}_1 := (Q_1 + Q_2 P)(I + P^T P)^{-1/2}$ define an orthonormal basis for a subspace that is invariant for $U + E$.

If $Q \in \mathbb{R}^{n \times n}$ is an orthogonal matrix, the following theorem shows that the singular value decompositions of its submatrices are highly correlated. The reader is again referred to [20] and the references therein for a proof and further details.

Theorem 4.4.3 (CS Decomposition) *Let $Q \in \mathbb{R}^{n \times n}$ be an orthogonal matrix partitioned as*

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix},$$

where $Q_{11} \in \mathbb{R}^{r \times r}$. Assume that $r < n - r$. Then there exist orthogonal matrices

$$U = \begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix}$$

partitioned in accordance with Q such that

$$Q = U \begin{bmatrix} C & S & 0 \\ S & -C & 0 \\ 0 & 0 & I \end{bmatrix} V^T,$$

where $C \in \mathbb{R}^{r \times r}$ and $S \in \mathbb{R}^{r \times r}$ are diagonal matrices all of whose entries lie in $[0, 1]$.

We now specialize the general bounds given by Proposition 4.2.1 for perturbations of b to the NT direction. Suppose that b is replaced by $b + \gamma \Delta b$. In this case, Yıldırım and Todd [71] showed that the interior-point bound evaluated at (X, y, S) reduces to a convenient norm bound given by

$$\left\| \gamma X^{-\frac{1}{2}} (\mathcal{F} \mathcal{A}^* [(\mathcal{A} \mathcal{F} \mathcal{A}^*)^{-1} \Delta b]) X^{-\frac{1}{2}} \right\| \leq 1, \quad (4.29)$$

where $\mathcal{F} = W \odot W$. Therefore, we will call the reciprocal of

$$\left\| X^{-\frac{1}{2}} (\mathcal{F} \mathcal{A}^* [(\mathcal{A} \mathcal{F} \mathcal{A}^*)^{-1} \Delta b]) X^{-\frac{1}{2}} \right\| \quad (4.30)$$

the NT bound and investigate its asymptotic relationship with the symmetrized partition bound given by the reciprocal of (4.18).

If $(X, y, S) = (X(\mu), y(\mu), S(\mu))$ then $W = \mu^{-1/2} X$ by (4.24). Consequently, (4.30) can further be simplified as

$$\left\| \mathcal{G}^{1/2} \mathcal{A}^* [(\mathcal{A} \mathcal{G} \mathcal{A}^*)^{-1} \Delta b] \right\|, \quad (4.31)$$

where $\mathcal{G} := X \odot X$.

We assume that X (we drop the dependence on μ to simplify the notation) is given by

$$X = \begin{bmatrix} \Lambda^* & \\ & 0 \end{bmatrix} + \begin{bmatrix} E_P & E_U^T \\ E_U & E_N \end{bmatrix}, \quad (4.32)$$

where the following bounds are satisfied:

$$E_P = O(\mu^\kappa), E_U = O(\mu^\beta), E_N = O(\mu^\alpha), \quad \text{where } \kappa > 0, \beta > 0, 0 < \alpha \leq 1. \quad (4.33)$$

Note that $X \succ 0$ implies that we can assume

$$\beta \geq \alpha/2, \quad (4.34)$$

since $X_{ij}^2 < X_{ii} X_{jj}$ for $i \neq j$. Luo, Sturm and Zhang [40] show that

$$\|X(\mu) - X^*\| = O(\mu), \quad \|S(\mu) - S^*\| = O(\mu) \quad (4.35)$$

if strict complementarity holds. Therefore, in this case, we can take $\alpha = \beta = \kappa = 1$ in (4.33). In contrast to the case for LP, the relationship (4.35) does not hold in the absence of strict complementarity. In this case, Goldfarb and Scheinberg [17] argue that both of the distances in (4.35) can be $\Omega(\mu^\nu)$, where $\nu < 1$.

Since $\text{sep}(\Lambda^*, 0) = \lambda_{\min}(\Lambda^*)$, we can apply Theorem 4.4.2 to X . By (4.32) and (4.33), for μ sufficiently small, X can then be decomposed as

$$X = QDQ^T = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} D_B & \\ & D_N \end{bmatrix} \begin{bmatrix} Q_{11}^T & Q_{21}^T \\ Q_{12}^T & Q_{22}^T \end{bmatrix}, \quad (4.36)$$

where $D_B \in \mathcal{S}^r$, $D_N \in \mathcal{S}^{n-r}$ and Q is orthogonal. By Theorem 4.4.2, we have

$$\begin{bmatrix} Q_{11} \\ Q_{21} \end{bmatrix} = \begin{bmatrix} I \\ P \end{bmatrix} (I + P^T P)^{-1/2},$$

where $P = O(\|E_U\|) = O(\mu^\beta)$ by (4.33). Therefore, the eigenvalues of $I + P^T P$ are $1 + O(\mu^{2\beta})$. Using the Taylor approximation given by

$$\frac{1}{\sqrt{1+x}} = 1 - \frac{1}{2}x + O(x^2), \quad (4.37)$$

we find that the eigenvalues of $(I + P^T P)^{-1/2}$ are also $1 + O(\mu^{2\beta})$ for μ sufficiently small. Therefore, $Q_{11} = I + O(\mu^{2\beta})$ and $Q_{21} = O(\mu^\beta)$. Furthermore, $Q_{12} = O(\mu^\beta)$ since $\|Q_{12}\| = \|Q_{21}\|$ by Theorem 4.4.3. The following lemma will be useful for the analysis later.

Lemma 4.4.1 *If Q is as in (4.36), then Q_{22} is nonsingular for μ sufficiently small.*

Proof: By Theorem 4.4.3, $C = I + O(\mu^{2\beta})$ indicating that all the singular values of Q_{22} are bounded away from 0. \square

Note that (4.33) implies that X has r eigenvalues of order $\Theta(1)$ for sufficiently small μ . Moreover, by using the block equality in (4.36), it follows that

$$D_N = Q_{12}^T(\Lambda^* + E_P)Q_{12} + Q_{22}^T E_U Q_{12} + Q_{12}^T E_U^T Q_{22} + Q_{22}^T E_N Q_{22} = O(\mu^\alpha), \quad (4.38)$$

by (4.33), (4.34) and the arguments following (4.37). By Theorem 4.4.1, $n - r$ of the eigenvalues of X then are of order $O(\mu^\alpha)$ for μ small enough. We stress that this is actually a sharper bound than the usual eigenvalue bound given by

$$|\lambda_k(A + E) - \lambda_k(A)| \leq \|E\|, \quad (4.39)$$

where $\lambda_k(\cdot)$ denotes the k th largest eigenvalue.

Similarly, we get the following expression for D_B :

$$D_B = Q_{11}^T(\Lambda^* + E_P)Q_{11} + Q_{11}^T E_U^T Q_{21} + Q_{21}^T E_U Q_{11} + Q_{21}^T E_N Q_{21}. \quad (4.40)$$

By (4.33) and the arguments following (4.37), we then have

$$D_B = \Lambda^* + O(\mu^\kappa) + O(\mu^{2\beta}). \quad (4.41)$$

Since $XS = \mu I$, the product of the eigenvalues of X and S are exactly equal to μ . Since the eigenvalues of S are bounded, it follows that the remaining $n - r$ eigenvalues of X are $\Omega(\mu)$. We then conclude that

$$\text{the eigenvalues of } D_B \text{ are } \Theta(1) \text{ and those of } D_N \text{ are } O(\mu^\alpha) \text{ and } \Omega(\mu). \quad (4.42)$$

We briefly discuss the concept of *scale-invariance*. Given (P) and (D), if we apply a change of variable in (P) such that X is replaced by $\hat{X} = Q^T X Q$, where Q is an orthogonal matrix in $\mathbb{R}^{n \times n}$, (P) transforms to

$$\begin{aligned} (\hat{\text{P}}) \quad & \min_{\hat{X}} \hat{C} \bullet \hat{X} \\ & \hat{A} \hat{X} = b, \\ & \hat{X} \succeq 0, \end{aligned}$$

where $\hat{C} := Q^T C Q$, and \hat{A} and \hat{A}^* are defined from $\{\hat{A}_i := Q^T A_i Q\}$ as in (4.3) and (4.4). The dual of this problem is

$$\begin{aligned} (\hat{\text{D}}) \quad & \max_{\hat{y}, \hat{S}} b^T \hat{y} \\ & \hat{A}^* \hat{y} + \hat{S} = \hat{C}, \\ & \hat{S} \succeq 0, \end{aligned}$$

which is exactly the transformation of (D) with (y, S) replaced by $(\hat{y} := y, \hat{S} := Q^T S Q)$. If (X, y, S) is a strictly feasible point for (P) and (D), then $(\hat{X}, \hat{y}, \hat{S}) = (Q^T X Q, y, Q^T S Q)$ is a strictly feasible point for $(\hat{\text{P}})$ and $(\hat{\text{D}})$.

A method for defining a search direction for SDP is said to be *Q-scale-invariant* if the direction at any iterate is the same as would result from scaling the problem and iterate by an arbitrary orthogonal matrix Q , using the method to determine

the direction for the scaled problem, and then scaling back. Todd [60] shows that the AHO, H..K..M, dual H..K..M and NT directions are all Q -scale invariant.

Furthermore, Yildirim and Todd [71] show that the Schur complement matrix given by $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$ is invariant under scaling, i.e., $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^* = \hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{\mathcal{A}}^*$.

We now apply this transformation to (P) using Q as given by (4.36). Since the NT direction is Q -scale invariant and since the right-hand side vector b is invariant under this transformation, we can evaluate the NT bound for (\hat{P}) instead of (P) and get the same interior-point bound. With this transformation,

$$\begin{aligned} \hat{A}_i &= \begin{bmatrix} Q_{11}^T & Q_{21}^T \\ Q_{12}^T & Q_{22}^T \end{bmatrix} \begin{bmatrix} (A_i)_P & (A_i)_U^T \\ (A_i)_U & (A_i)_N \end{bmatrix} \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \\ &= \begin{bmatrix} (A_i)_P + \xi_1 & (A_i)_U^T Q_{22} + \xi_2^T \\ Q_{22}^T (A_i)_U + \xi_2 & Q_{22}^T (A_i)_N Q_{22} + \xi_3 \end{bmatrix}, \end{aligned} \quad (4.43)$$

where ξ_1 , ξ_2 and ξ_3 are matrices of appropriate dimensions each of which is $O(\mu^\beta)$. We claim that the matrices \hat{A}_i also satisfy Assumption 4.2.4 provided that μ is sufficiently small. Indeed, it follows from (4.43) that the submatrices $(\hat{A}_i)_P$ span \mathcal{S}^r . For primal nondegeneracy, we use Lemma 4.4.1 to show that the relevant parts of the \hat{A}_i are linearly independent in \mathcal{S}^n since Q_{22} is nonsingular for μ sufficiently small.

We now reconsider (4.31). First, we study the nontrivial case, i.e., when Δb is in the right space. By Assumption 4.2.4, there exists a unique $\tilde{V} \in \mathcal{S}^r$ such that

$$\Delta b_i = (A_i)_P \bullet \tilde{V}, \quad i = 1, \dots, m. \quad (4.44)$$

We find it easier to express (4.31) in standard matrix-vector form by using sym-

metrized Kronecker products. We define the linear operator \mathbf{svec} , which maps $U \in \mathcal{S}^n$ to \mathbb{R}^{n^2} , as follows:

$$\mathbf{svec}(U) := (U_{11}, \sqrt{2}U_{21}, \dots, \sqrt{2}U_{n1}, U_{22}, \sqrt{2}U_{32}, \dots, \sqrt{2}U_{n2}, \dots, U_{nn})^T. \quad (4.45)$$

Note that \mathbf{svec} is an isometry between \mathcal{S}^n and \mathbb{R}^{n^2} since for every $U, V \in \mathcal{S}^n$,

$$U \bullet V = \mathbf{svec}(U)^T \mathbf{svec}(V). \quad (4.46)$$

The symmetric Kronecker product of two matrices $U, V \in \mathbb{R}^{n \times n}$ is an $n^2 \times n^2$ matrix, whose action on a vector $\mathbf{svec}(Z) \in \mathbb{R}^{n^2}$ where $Z \in \mathcal{S}^n$ is defined by

$$(U \otimes_s V) \mathbf{svec}(Z) := \mathbf{svec}((U \odot V)Z) = \frac{1}{2} \mathbf{svec}(UZV^T + VZU^T). \quad (4.47)$$

We also define another linear operator, called \mathbf{vec} , which maps $\mathbb{R}^{k \times n}$ to \mathbb{R}^{nk} . For $P \in \mathbb{R}^{k \times n}$, $\mathbf{vec}(P)$ is defined by

$$\mathbf{vec}(P) := (P_{11}, P_{21}, \dots, P_{k1}, P_{12}, \dots, P_{k2}, \dots, P_{kn})^T. \quad (4.48)$$

Note that \mathbf{vec} is also an isometry between $\mathbb{R}^{k \times n}$ and \mathbb{R}^{nk} with respect to their standard inner products. The Kronecker product of two matrices $U \in \mathbb{R}^{n \times n}$ and $V \in \mathbb{R}^{k \times k}$ is an $nk \times nk$ matrix whose (i, j) block is given by $U_{ij}V$. Its action on a vector $\mathbf{vec}(Z) \in \mathbb{R}^{nk}$ where $Z \in \mathbb{R}^{k \times n}$ (cf. (4.47)) is given by

$$(U \otimes V) \mathbf{vec}(Z) := \mathbf{vec}(VZU^T). \quad (4.49)$$

Although we define the operator \mathbf{svec} as in (4.45), any permutation on the right-hand side would clearly preserve the isometry between \mathcal{S}^n and \mathbb{R}^{n^2} . Therefore, we

introduce yet another linear map \mathbf{sv} , which is basically equivalent to the \mathbf{svec} operation applied in a different order. For any $U \in \mathcal{S}^n$, \mathbf{sv} is defined by

$$\mathbf{sv}(U) := (\mathbf{svec}(U_P)^T, \sqrt{2}\mathbf{vec}(U_U^T)^T, \mathbf{svec}(U_N)^T)^T. \quad (4.50)$$

Similarly, we denote the symmetric Kronecker product resulting from \mathbf{sv} of U and $V \in \mathbb{R}^{n \times n}$ by $U \otimes_{sv} V$, where the definition is similar to (4.47) with \mathbf{svec} replaced by \mathbf{sv} .

We can rewrite (4.31) defining the NT bound as

$$\|\mathcal{G}^{1/2} \mathcal{A}^* w\|, \quad (4.51)$$

where w satisfies

$$\mathcal{A} \mathcal{G} \mathcal{A}^* w = \Delta b. \quad (4.52)$$

Defining $A \in \mathbb{R}^{m \times n^2}$ as

$$A := \begin{bmatrix} \mathbf{sv}(A_1)^T \\ \vdots \\ \mathbf{sv}(A_m)^T \end{bmatrix}, \quad (4.53)$$

we can easily verify that (4.52) can be rewritten in standard matrix-vector form as

$$A (X \otimes_{sv} X) A^T w = \Delta b. \quad (4.54)$$

Note that A has full row rank by Assumption 4.2.2 and $X \otimes_{sv} X$ is positive definite. Therefore, w is uniquely defined by (4.54). We also partition A in accordance with (4.50) as

$$A = [B \quad N_1 \quad N_2],$$

where B , N_1 and N_2 have $r^{\bar{2}}$, $r(n-r)$ and $(n-r)^{\bar{2}}$ columns, respectively. By Assumption 4.2.4, B has full column rank and $[B \ N_1]$ has full row rank. By (4.44), $\Delta b = B\tilde{v}$, where $\tilde{v} = \mathbf{svec}(\tilde{V})$.

We now apply the NT bound to the transformed problem (\hat{P}) . We will denote the corresponding matrices with a hat. For (\hat{P}) , (4.54) is given by

$$\hat{A}(D \otimes_{sv} D) \hat{A}^T w = \Delta b = \hat{B}\tilde{v} + u, \quad (4.55)$$

where

$$u := B\tilde{v} - \hat{B}\tilde{v} = O(\mu^\beta) \quad (4.56)$$

by (4.43). Furthermore, \hat{A} can be partitioned similarly to A as

$$\hat{A} = \begin{bmatrix} \hat{B} & \hat{N}_1 & \hat{N}_2 \end{bmatrix}. \quad (4.57)$$

It follows from (4.43) and the argument following it that \hat{B} has full column rank and $[\hat{B} \ \hat{N}_1]$ has full row rank for μ sufficiently small. Note that D is block diagonal. The next lemma will show that $D \otimes_{sv} D$ will also be a block diagonal matrix in this case.

Lemma 4.4.2 *If D is given by*

$$D = \begin{bmatrix} D_B & \\ & D_N \end{bmatrix},$$

where $D_B \in \mathcal{S}^r$ and $D_N \in \mathcal{S}^{n-r}$, then $D \otimes_{sv} D$ is given by

$$D \otimes_{sv} D = \begin{bmatrix} D_B \otimes_s D_B & & \\ & D_N \otimes D_B & \\ & & D_N \otimes_s D_N \end{bmatrix}. \quad (4.58)$$

Proof: Due to the partition given by (4.50), it suffices to consider the action of $D \otimes_{sv} D$ on the vectors obtained from the matrices of the form

$$U_1 = \begin{bmatrix} U_P \\ 0 \end{bmatrix}, \quad U_2 = \begin{bmatrix} & U_U^T \\ U_U & \end{bmatrix}, \quad U_3 = \begin{bmatrix} 0 \\ & U_N \end{bmatrix}.$$

Let us consider each form in turn:

$$\begin{aligned} (D \otimes_{sv} D) \mathbf{sv}(U_1) &= \begin{bmatrix} \mathbf{svec}(D_B U_P D_B) \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} D_B \otimes_s D_B \\ 0 \\ 0 \end{bmatrix} \mathbf{svec}(U_P). \\ (D \otimes_{sv} D) \mathbf{sv}(U_2) &= \begin{bmatrix} 0 \\ \sqrt{2} \mathbf{svec}(D_B U_U^T D_N) \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ D_N \otimes D_B \\ 0 \end{bmatrix} \sqrt{2} \mathbf{svec}(U_U^T). \\ (D \otimes_{sv} D) \mathbf{sv}(U_3) &= \begin{bmatrix} 0 \\ 0 \\ \mathbf{svec}(D_N U_N D_N) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ D_N \otimes_s D_N \end{bmatrix} \mathbf{svec}(U_N). \end{aligned}$$

□

By a property of Kronecker products, Lemma 4.4.2 allows us to obtain bounds on the eigenvalues of each block in (4.58). Let U and V be two matrices in \mathcal{S}^n with $\lambda(U) = \{\lambda_1, \dots, \lambda_n\}$ and $\lambda(V) = \{\mu_1, \dots, \mu_n\}$. Then, the eigenvalues of $U \otimes V$ are given by $\{\lambda_i \mu_j : 1 \leq i, j \leq n\}$ and those of $U \otimes_s U$ by $\{\lambda_i \lambda_j : 1 \leq i \leq j \leq n\}$. Therefore, if $D \otimes_{sv} D$ is written as

$$D \otimes_{sv} D = \begin{bmatrix} D_1 & & \\ & D_2 & \\ & & D_3 \end{bmatrix},$$

then D_1 has r^2 eigenvalues which are $\Theta(1)$; D_2 has $r(n-r)$ eigenvalues which are $O(\mu^\alpha)$ and $\Omega(\mu)$; and D_3 has $(n-r)^2$ eigenvalues which are $O(\mu^{2\alpha})$ and $\Omega(\mu^2)$ by (4.42). Consequently, we get

$$\|D_1\| = O(1), \quad \|D_2\| = O(\mu^\alpha), \quad \|D_3\| = O(\mu^{2\alpha}), \quad (4.59)$$

as well as

$$\|D_1^{-1}\| = O(1), \quad \|D_2^{-1}\| = O(1/\mu), \quad \|D_3^{-1}\| = O(1/\mu^2). \quad (4.60)$$

By (4.55) and (4.57), we aim to find w which satisfies

$$\begin{bmatrix} \hat{B} & \hat{N} \end{bmatrix} \begin{bmatrix} D_{\hat{B}} \\ D_{\hat{N}} \end{bmatrix} \begin{bmatrix} \hat{B}^T \\ \hat{N}^T \end{bmatrix} w = \hat{B}\tilde{v} + u, \quad (4.61)$$

where $\hat{N} := [\hat{N}_1 \quad \hat{N}_2]$, $D_{\hat{B}} := D_1$ and $D_{\hat{N}}$ is block diagonal with D_2 and D_3 as the diagonal blocks. Note that \hat{B} has full column rank and $\hat{B} = B + O(\mu^\beta)$. Therefore, if we choose $R \in \mathbb{R}^{m \times (m-r^2)}$ with columns a basis for the null space of B^T so that the augmented matrix $[B \quad R]$ is nonsingular, then for all sufficiently small μ , $[\hat{B} \quad R]$ is nonsingular and its inverse, which we denote by Z , is uniformly bounded. Therefore, premultiplying (4.61) by Z yields

$$\begin{bmatrix} I & \tilde{N}_U \\ 0 & \tilde{N}_L \end{bmatrix} \begin{bmatrix} D_{\hat{B}} \\ D_{\hat{N}} \end{bmatrix} \begin{bmatrix} I & 0 \\ \tilde{N}_U^T & \tilde{N}_L^T \end{bmatrix} \tilde{w} = \begin{bmatrix} \tilde{v} \\ 0 \end{bmatrix} + \begin{bmatrix} \tilde{u}_1 \\ \tilde{u}_2 \end{bmatrix}, \quad (4.62)$$

where $\tilde{w} := Z^{-T}w$, \tilde{u}_1 and \tilde{u}_2 are the appropriate partitions of $\tilde{u} := Zu$, and \tilde{N}_U and \tilde{N}_L those of $\tilde{N} := Z\hat{N}$. Since \hat{A} has full row rank, so does $Z\hat{A}$, therefore \tilde{N}_L has full row rank. However, we can get more information using primal nondegeneracy: Since

$[\hat{B} \quad \hat{N}_1]$ has full row rank, so does $Z[\hat{B} \quad \hat{N}_1]$. Consequently, if $\tilde{N}_L =: [\tilde{N}_{L1} \quad \tilde{N}_{L2}]$, where \tilde{N}_{L1} has $r(n-r)$ columns, then \tilde{N}_{L1} has full row rank. If we partition \tilde{w} similarly as \tilde{w}_1 and \tilde{w}_2 , (4.62) can be decomposed into two equations given by

$$(D_{\hat{B}} + FF^T) \tilde{w}_1 + FG^T \tilde{w}_2 = \tilde{v} + \tilde{u}_1, \quad (4.63)$$

$$GF^T \tilde{w}_1 + GG^T \tilde{w}_2 = \tilde{u}_2, \quad (4.64)$$

where

$$F := \tilde{N}_U D_{\hat{N}}^{1/2}, \quad G := \tilde{N}_L D_{\hat{N}}^{1/2}. \quad (4.65)$$

Note that G has full row rank since \tilde{N}_L does. Moreover, by (4.59), F and G satisfy

$$\|F\| = O(\mu^{\alpha/2}), \quad \|G\| = O(\mu^{\alpha/2}). \quad (4.66)$$

We can write \tilde{w}_2 in terms of \tilde{w}_1 using (4.64):

$$\tilde{w}_2 = (GG^T)^{-1} \tilde{u}_2 - (GG^T)^{-1} GF^T \tilde{w}_1. \quad (4.67)$$

We can now combine (4.67) with (4.63) to obtain

$$(D_{\hat{B}} + F(I - P_G)F^T) \tilde{w}_1 = \tilde{v} + \tilde{u}_1 - FG^T (GG^T)^{-1} \tilde{u}_2, \quad (4.68)$$

where P_G is the orthogonal projection matrix onto the range space of G^T . Therefore, $I - P_G$ is the orthogonal projection matrix onto the null space of G . The following lemma will be useful for the analysis.

Lemma 4.4.3 *We have*

$$\|G^T (GG^T)^{-1} \tilde{u}_2\| = O(\mu^{\beta-1/2}). \quad (4.69)$$

Proof: Note that the left-hand side of (4.69) is precisely the optimal value of the least squares problem given by $\min_x \{\|x\| : Gx = \tilde{u}_2\}$. Therefore, it suffices to manufacture a feasible solution which satisfies the given upper bound. However, $\tilde{N}_L = [\tilde{N}_{L1} \quad \tilde{N}_{L2}]$ and \tilde{N}_{L1} has full row rank by the argument following (4.62). Therefore, there exists $q \in \mathbb{R}^{r(n-r)}$ with $\|q\| = O(\mu^\beta)$ such that $\tilde{u}_2 = \tilde{N}_{L1} q$. It follows then that

$$\tilde{u}_2 = \begin{bmatrix} \tilde{N}_{L1} & \tilde{N}_{L2} \end{bmatrix} \begin{bmatrix} q \\ 0 \end{bmatrix} = \tilde{N}_L D_{\tilde{N}}^{1/2} D_{\tilde{N}}^{-1/2} \begin{bmatrix} q \\ 0 \end{bmatrix} = G \begin{bmatrix} D_2^{-1/2} q \\ 0 \end{bmatrix}.$$

Therefore, a feasible solution for the least squares problem is given by

$$\tilde{x} = \begin{bmatrix} D_2^{-1/2} q \\ 0 \end{bmatrix}$$

and $\|\tilde{x}\| = O(\mu^{\beta-1/2})$ by (4.60) together with the fact that $\|q\| = O(\mu^\beta)$. \square

We briefly review the Neumann lemma now [20]. Let U be an invertible matrix and let V satisfy $\|U^{-1}V\| \leq 1/2$. Then, $I + U^{-1}V$ is invertible with $\|(I + U^{-1}V)^{-1}\| \leq 2$. Moreover $U + V$ is invertible and given by

$$(U + V)^{-1} = U^{-1} - U^{-1}V(I + U^{-1}V)^{-1}U^{-1}. \quad (4.70)$$

We now apply this result to (4.68) with $U := D_{\hat{B}}$ and $V := F(I - P_G)F^T$. Note that $\|U^{-1}\| = O(1)$ by (4.60) and $\|V\| = O(\mu^\alpha)$ by (4.66) and the fact that $\|I - P_G\| \leq 1$. It follows from (4.70) that $\|(U + V)^{-1}\| = O(1)$. The right-hand side of (4.68) is given by $\tilde{v} + O(\mu^\beta) + O(\mu^{\alpha/2+\beta-1/2})$ by (4.56) and Lemma 4.4.3. Therefore, we obtain $\tilde{w}_1 = O(1)$ by (4.68) for μ sufficiently small if

$$\alpha + 2\beta > 1. \quad (4.71)$$

We will assume hereafter that α and β satisfy (4.71).

Let us reconsider (4.51). We will apply the **sv** operation to the matrix whose operator norm we need to evaluate:

$$\mathbf{sv}(\mathcal{G}^{1/2} \mathcal{A}^* w) = (X^{1/2} \otimes_{sv} X^{1/2}) A^T w.$$

By the scale-invariance of the NT bound, we can alternatively use the transformed problem to evaluate it. Therefore, we need to compute

$$\begin{aligned} (D^{1/2} \otimes_{sv} D^{1/2}) \hat{A}^T w &= \begin{bmatrix} D_{\hat{B}}^{1/2} & \\ & D_{\hat{N}}^{1/2} \end{bmatrix} \begin{bmatrix} I & 0 \\ \tilde{N}_U^T & \tilde{N}_L^T \end{bmatrix} \begin{bmatrix} \tilde{w}_1 \\ \tilde{w}_2 \end{bmatrix} \\ &= \begin{bmatrix} D_{\hat{B}}^{1/2} \tilde{w}_1 \\ F^T \tilde{w}_1 + G^T \tilde{w}_2 \end{bmatrix}. \end{aligned} \quad (4.72)$$

For the top part, we use (4.68):

$$D_{\hat{B}}^{1/2} \tilde{w}_1 = D_{\hat{B}}^{-1/2} \tilde{v} + D_{\hat{B}}^{-1/2} (\tilde{u}_1 - FG^T(GG^T)^{-1} \tilde{u}_2 - F(I - P_G)F^T \tilde{w}_1). \quad (4.73)$$

Using (4.60) together with the arguments following (4.70), we conclude that

$$D_{\hat{B}}^{1/2} \tilde{w}_1 = D_{\hat{B}}^{-1/2} \tilde{v} + O(\mu^\beta) + O(\mu^{\alpha/2+\beta-1/2}) + O(\mu^\alpha) = D_{\hat{B}}^{-1/2} \tilde{v} + O(\mu^\tau), \quad (4.74)$$

where

$$\tau := \min\{\beta, \alpha/2 + \beta - 1/2, \alpha\} = \min\{\alpha, \alpha/2 + \beta - 1/2\} > 0 \quad (4.75)$$

since $\alpha \leq 1$. However, if we apply the **svec** operation to (4.18) defining the symmetrized partition bound, we get

$$\mathbf{svec} \left((\Lambda^*)^{-1/2} \tilde{V} (\Lambda^*)^{-1/2} \right) = ((\Lambda^*)^{-1/2} \otimes_s (\Lambda^*)^{-1/2}) \tilde{v}. \quad (4.76)$$

Since $D_{\hat{B}} = D_B \otimes_s D_B$, and $D_B \rightarrow \Lambda^*$ as $\mu \downarrow 0$ by (4.41), we conclude that the top part of (4.72) which is given by (4.74) converges to (4.76) since $\tau > 0$ by (4.71). To determine the rate of convergence, let us define

$$\eta := \min\{\kappa, 2\beta\} > 0. \quad (4.77)$$

By (4.41), $D_B = \Lambda^* + O(\mu^\eta)$, which implies that the eigenvalues of D_B are given by those of Λ^* perturbed by $O(\mu^\eta)$ terms. Therefore, the eigenvalues of $D_{\hat{B}}$ are also given by those of $\Lambda^* \otimes_s \Lambda^*$ plus error terms given by $O(\mu^\eta)$. Finally, an application of (4.37) to the eigenvalues of $D_{\hat{B}}^{-1/2}$ yields

$$D_{\hat{B}}^{-1/2} = (\Lambda^*)^{-1/2} \otimes_s (\Lambda^*)^{-1/2} + O(\mu^\eta). \quad (4.78)$$

Using (4.74), we then conclude that

$$D_{\hat{B}}^{1/2} \tilde{w}_1 = ((\Lambda^*)^{-1/2} \otimes_s (\Lambda^*)^{-1/2}) \tilde{v} + O(\mu^\eta) + O(\mu^\tau). \quad (4.79)$$

Under strict complementarity, α , β and κ can all be set to 1 and hence, $\tau = \eta = 1$ by (4.75) and (4.77).

Let us now examine the lower part of (4.72). Since $\tilde{w}_1 = O(1)$ under the condition (4.71) for μ sufficiently small, we have $\|F^T \tilde{w}_1\| = O(\mu^{\alpha/2})$ using (4.66). By (4.67), we can write the second term as

$$G^T \tilde{w}_2 = G^T (GG^T)^{-1} \tilde{u}_2 - P_G F^T \tilde{w}_1. \quad (4.80)$$

By Lemma 4.4.3, the first term in (4.80) is $O(\mu^{\beta-1/2})$. The second term is $O(\mu^{\alpha/2})$ since $\|P_G\| \leq 1$. Consequently, the lower part of (4.72) is $O(\mu^\zeta)$, where

$$\zeta = \min\{\alpha/2, \beta - 1/2\}. \quad (4.81)$$

Once again, we can take $\zeta = 1/2$ under strict complementarity. Therefore, we conclude that the lower part tends to 0 as $\mu \downarrow 0$ if

$$\beta > \frac{1}{2}. \quad (4.82)$$

Note that the condition (4.82) is sufficient to have $\tau > 0$ since (4.71) is then automatically satisfied. Therefore, if the condition (4.82) is satisfied, we conclude that the interior-point bound evaluated on the central path using the NT direction converges to the symmetrized partition bound given by the reciprocal of (4.18) for perturbations Δb in the range space of B .

We now comment on the rate of convergence. Note that the 2-norm of a symmetric matrix is given by the maximum of its eigenvalues in absolute value. We can use the eigenvalue inequality (4.39) in conjunction with the norm inequality $\|E\|_2 \leq \|E\|_F$ for any $E \in \mathbb{R}^{n \times n}$. By the preceding discussion, we then conclude that the rate of convergence can be bounded by μ^ψ , where

$$\psi := \min\{\eta, \tau, \zeta\} = \min\{\kappa, \alpha/2, \beta - 1/2\} > 0, \quad (4.83)$$

where we used (4.75), (4.77), (4.81) and (4.34). Under strict complementarity, ψ can be taken to be $1/2$.

We next consider the case when Δb is not in the range space of B . In this case, both the optimal partition and the symmetrized partition bounds are 0. We can write $\Delta b = B\tilde{v} + u$, where $B\tilde{v}$ is the orthogonal projection of Δb onto the range space of B . It follows then that u lies in the range space of R below (4.61) and $\|u\|$ is bounded away from 0. By (4.56), $\Delta b = \hat{B}\tilde{v} + u + O(\mu^\beta)$. Now $u = Rd$ for some d ,

and thus $\|d\|$ is also bounded away from 0. However, \tilde{u}_2 as in (4.62) then is equal to $d + O(\mu^\beta)$. Hence $\|\tilde{u}_2\|$ is bounded away from 0. By (4.64), we have

$$\tilde{N}_L D_{\tilde{N}}^{1/2} (F^T \tilde{w}_1 + G^T \tilde{w}_2) = \tilde{u}_2. \quad (4.84)$$

However, the term in parentheses is precisely the lower part of (4.72). Let us denote that term by p and let $q := D_{\tilde{N}}^{1/2} p$. Since \tilde{u}_2 is nonzero, $\|q\| \geq \nu$, where $\nu > 0$ is the norm of the least-squares solution of $\tilde{N}_L x = \tilde{u}_2$. Therefore,

$$\nu \leq \|q\| \leq \left\| D_{\tilde{N}}^{1/2} \right\| \|p\|,$$

which implies that

$$\|p\| \geq \frac{\nu}{\left\| D_{\tilde{N}}^{1/2} \right\|}.$$

Consequently, $\|p\| \rightarrow \infty$ as $\mu \downarrow 0$ since $\left\| D_{\tilde{N}}^{1/2} \right\| = O(\mu^{\alpha/2})$ by (4.59). This argument reveals that the NT bound given by the reciprocal of the matrix norm tends to 0 as desired as μ tends to 0, indeed with convergence rate $O(\mu^{\alpha/2})$.

Therefore, we have proved the following theorem:

Theorem 4.4.4 *Let (P) and (D) satisfy Assumptions 4.2.1–4.2.3. Assume that $(X, y, S) := (X(\mu), y(\mu), S(\mu))$ is such that X and the primal optimal solution X^* satisfy (4.32) and (4.33) with $\beta > 1/2$. Let the right-hand side vector b be replaced by $b + \gamma \Delta b$, where $\gamma \in \mathbb{R}$ and $\Delta b \in \mathbb{R}^m$. Then the NT bound evaluated at (X, y, S) converges to the symmetrized partition bound as $\mu \downarrow 0$ and, in the nontrivial case when these bounds are nonzero, the rate of convergence is bounded above by $O(\mu^\psi)$, where ψ is defined by (4.83).*

The NT direction is primal-dual symmetric. By the primal-dual symmetry outlined in Section 4.3, we conclude that the NT bound for perturbations of the cost matrix C also converges to the symmetrized partition bound if Assumption 4.2.4 holds (i.e., weak primal nondegeneracy and dual nondegeneracy) and the requirements on X and X^* are replaced by the corresponding ones on S and S^* , which we call the dual requirements.

Corollary 4.4.1 *Let (P) and (D) satisfy Assumptions 4.2.1, 4.2.2 and 4.2.4. Assume that $(X, y, S) := (X(\mu), y(\mu), S(\mu))$ is such that S and the dual optimal solution (y^*, S^*) satisfy the dual requirements of (4.32) and (4.33) with $\beta > 1/2$. Let the cost matrix C be replaced by $C + \gamma\Delta C$, where $\gamma \in \mathbb{R}$ and $\Delta C \in \mathcal{S}^n$. Then the NT bound evaluated at (X, y, S) converges to the symmetrized partition bound as $\mu \downarrow 0$ and, in the nontrivial case when these bounds are nonzero, the rate of convergence is bounded above by $O(\mu^\psi)$, where ψ is defined by (4.83).*

So far, we have only considered the NT direction. In the next theorem, we prove that this nice correspondence between the interior-point bounds and the symmetrized partition bounds continues to hold for any direction in the Monteiro-Zhang family for iterates on the central path.

Theorem 4.4.5 *Let (P) , (D) , $(X, y, S) := (X(\mu), y(\mu), S(\mu))$ and the optimal solution (X^*, y^*, S^*) satisfy the assumptions given in Theorem 4.4.4 for perturbations of b or those given in Corollary 4.4.1 for perturbations of C . Then the interior-point bound evaluated at (X, y, S) using any direction in the Monteiro-Zhang family converges to the symmetrized partition bound as $\mu \downarrow 0$.*

Proof: The interior-point bounds are given by Propositions 4.2.1 and 4.2.2 respectively for perturbations of b and C , where the operators \mathcal{E} and \mathcal{F} are given by (4.9) with $M \succ 0$. We claim that it suffices to show

$$\mathcal{E}^{-1}\mathcal{F} = W \odot W \tag{4.85}$$

for iterates on the central path. Indeed, if (4.85) holds, then the matrices whose minimum and maximum eigenvalues define the interior-point bounds actually coincide in both Propositions 4.2.1 and 4.2.2 and both bounds can be expressed as norm bounds on the corresponding matrices. This is precisely the NT bound.

We now prove (4.85). For any $U \in \mathcal{S}^n$, we need to show that

$$\mathcal{E}^{-1}\mathcal{F}U = (W \odot W)U = WUW, \tag{4.86}$$

or, since \mathcal{E} is invertible, that $\mathcal{F}U = \mathcal{E}(WUW)$. But $WUW = \mu^{-1}XUX$, and

$$\mathcal{E}(\mu^{-1}XUX) = \frac{1}{2}\mu^{-1}(SXUXM + MXUXS) = \frac{1}{2}(UXM + MXU) = \mathcal{F}U,$$

where we used $XS = \mu I$. □

4.5 Central Path Neighborhood

In Section 4.4, we have shown that the interior-point bounds evaluated on the central path asymptotically coincide with the symmetrized partition bounds. In this section, we aim to extend the same asymptotic result to iterates lying in a very narrow neighborhood of the central path using the Nesterov-Todd direction.

The primal-dual path-following algorithms generate iterates that “closely” follow the central path. One of the proximity measures that is used in short-step path-following algorithms is given by

$$\mathcal{N}_F(\theta) := \{(X, y, S) \in \mathcal{F}_P \times \mathcal{F}_D : \left\| X^{\frac{1}{2}} S X^{\frac{1}{2}} - \mu I \right\|_F \leq \theta \mu\}, \quad (4.87)$$

where \mathcal{F}_P and \mathcal{F}_D respectively denote the set of primal and dual strictly feasible points, $\theta \in [0, 1)$ is a parameter and $\mu := \mu(X, S) = (X \bullet S)/n$ is the duality measure. This is called the narrow neighborhood.

We first discuss the possibility of extending the results of the previous section to iterates lying in a narrow neighborhood.

Example 4.5.1 Consider the following SDP instance with $m = 1$, $n = 2$, $b = 1$ and

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

The unique optimal solution is given by

$$X^* = \text{diag}(1, 0), \quad y^* = 0, \quad S^* = \text{diag}(0, 1).$$

The optimal solution is primal and dual nondegenerate and strict complementarity holds. The points on the central path are given by $X(\mu) = \text{diag}(1, \mu)$, $y(\mu) = -\mu$, $S(\mu) = \text{diag}(\mu, 1)$. Assume that the right-hand side vector is replaced by $1 + \gamma$. The partition bounds are $(-1, +\infty)$ and the symmetrized partition bounds are $(-1, 1)$. Consider the following feasible point with

$$X = \begin{bmatrix} 1 & \sqrt{\mu/3} \\ \sqrt{\mu/3} & \mu \end{bmatrix}$$

and $(y, S) = (y(\mu), S(\mu))$, which satisfies $(X \bullet S)/2 = \mu$. Furthermore, $(X, y, S) \in \mathcal{N}_F(\sqrt{2/3})$. However, the NT bound evaluated at (X, y, S) does not converge to the desired symmetrized partition bound $(-1, 1)$. Numerical evidence suggests that the bound converges approximately to the interval $(-.9, .9)$ as $\mu \downarrow 0$. For such a point,

$$\left\| X^{\frac{1}{2}} S X^{\frac{1}{2}} - \mu I \right\|_F = \Theta(\mu).$$

□

As Example 4.5.1 illustrates, the interior-point bound does not necessarily converge to the symmetrized partition bound even when the iterates lie in a narrow neighborhood. However, in the same example, if we let

$$X := X(\beta) = \begin{bmatrix} 1 & (1/\sqrt{3})\mu^\beta \\ (1/\sqrt{3})\mu^\beta & \mu \end{bmatrix}$$

with $\beta \in (1/2, 1]$, and fix (y, S) at $(y(\mu), S(\mu))$, then (X, y, S) satisfies

$$\left\| X^{\frac{1}{2}} S X^{\frac{1}{2}} - \mu I \right\|_F = O(\mu^{1/2+\beta})$$

and the NT bound evaluated at $(X(\beta), y, S)$ does converge to the desired symmetrized partition bound $(-1, 1)$ as $\mu \downarrow 0$.

This observation suggests the possibility of extending the convergence results of the previous section to iterates satisfying the condition

$$\left\| X^{\frac{1}{2}} S X^{\frac{1}{2}} - \mu I \right\|_F = O(\mu^\sigma) \quad \text{where } \sigma > 1. \quad (4.88)$$

Alternatively, condition (4.88) can be stated as

$$X^{\frac{1}{2}} S X^{\frac{1}{2}} = \mu I + E \quad \text{where } E = O(\mu^\sigma). \quad (4.89)$$

Henceforth, we assume that (X, y, S) is a primal-dual strictly feasible iterate, with a sufficiently small duality gap, satisfying (4.89). Note that the corresponding scaling matrix W is given by

$$W = X^{\frac{1}{2}} \left(X^{\frac{1}{2}} S X^{\frac{1}{2}} \right)^{-\frac{1}{2}} X^{\frac{1}{2}} = X^{\frac{1}{2}} (\mu I + E)^{-\frac{1}{2}} X^{\frac{1}{2}}. \quad (4.90)$$

Let us define

$$\bar{W} := \sqrt{\mu} W. \quad (4.91)$$

Note that $\bar{W} = X$ on the central path. If we denote the eigenvalues of E in (4.90) by λ_i , $i = 1, \dots, n$ where $\lambda_i = O(\mu^\sigma)$ by (4.89), the Taylor approximation given by (4.37) can be applied to the eigenvalues of $(\mu I + E)^{-1/2}$ to obtain

$$\bar{W} = X^{\frac{1}{2}} (I + F) X^{\frac{1}{2}}, \quad (4.92)$$

where $F = O(\mu^{\sigma-1})$.

Assuming the right-hand side vector b is replaced by $b + \gamma \Delta b$, the NT bound at (X, y, S) given by the reciprocal of (4.30) is equivalent to the reciprocal of

$$\left\| X^{-\frac{1}{2}} \left(\bar{\mathcal{F}} \mathcal{A}^* \left[(\mathcal{A} \bar{\mathcal{F}} \mathcal{A}^*)^{-1} \Delta b \right] \right) X^{-\frac{1}{2}} \right\|, \quad (4.93)$$

where $\bar{\mathcal{F}} := \bar{W} \odot \bar{W}$.

Throughout this section, we assume that \bar{W} defined as (4.91) satisfies the assumptions (4.32) and (4.33) imposed on X in the previous section. Therefore, $\bar{W} \rightarrow X^*$ as $\mu \downarrow 0$. By repeating the analysis, we can extend the results proved there to iterates satisfying (4.88) if we can show that $X^{-1/2}$ is “close” to $\bar{W}^{-1/2}$ for such iterates.

Using (4.92), it follows that

$$\bar{W}^{-1} = X^{-\frac{1}{2}}(I + F)^{-1}X^{-\frac{1}{2}} = \mu^{-1/2}X^{-\frac{1}{2}}(\mu I + E)^{1/2}X^{-\frac{1}{2}} \quad (4.94)$$

by definitions of E and F . Once again, an application of the Taylor approximation given by

$$\sqrt{1+x} = 1 + \frac{1}{2}x + O(x^2) \quad (4.95)$$

to the eigenvalues of $\mu^{-1/2}(\mu I + E)^{1/2}$ yields

$$\bar{W}^{-1} = X^{-\frac{1}{2}}(I + H)X^{-\frac{1}{2}}, \quad (4.96)$$

where $H = O(\mu^{\sigma-1})$.

If $A \in \mathcal{S}^n$ is given by $A = BB^T$ with $B \in \mathbb{R}^{n \times n}$ nonsingular, then

$$(B^{-1}A^{1/2})(A^{1/2}B^{-T}) = I. \quad (4.97)$$

Consequently, $B^{-1}A^{1/2} =: Q$ is an orthogonal matrix and $A^{1/2} = BQ$. Applying this to (4.96) with $A := \bar{W}^{-1}$ and $B := X^{-\frac{1}{2}}(I + H)^{1/2}$ yields

$$\bar{W}^{-1/2} = X^{-\frac{1}{2}}(I + H)^{1/2}Q \quad \text{or} \quad X^{-\frac{1}{2}} = \bar{W}^{-1/2}Q^T(I + H)^{-1/2}. \quad (4.98)$$

A final application of the Taylor approximation (4.37) to the eigenvalues of $(I + H)^{-1/2}$ gives

$$X^{-\frac{1}{2}} = \bar{W}^{-1/2}Q^T(I + J), \quad (4.99)$$

where $J = O(\mu^{\sigma-1})$. Substituting this into (4.93), which defines the NT bound, we obtain

$$\left\| (I + J)Q\bar{W}^{-1/2} \left(\bar{\mathcal{F}}\mathcal{A}^* \left[(\mathcal{A}\bar{\mathcal{F}}\mathcal{A}^*)^{-1} \Delta b \right] \right) \bar{W}^{-1/2}Q^T(I + J) \right\|. \quad (4.100)$$

However, the analysis of the previous section in conjunction with the facts that $J = O(\mu^{\sigma-1})$ and that the operator norm is invariant under orthogonal transformations implies that the NT bound converges to the desired symmetrized partition bound as $\mu \downarrow 0$. We have thus proved the following theorem.

Theorem 4.5.1 *Let (P) and (D) satisfy Assumptions 4.2.1–4.2.3. Assume that (X, y, S) is a primal-dual strictly feasible iterate satisfying (4.88) such that the corresponding \bar{W} as in (4.91) satisfies the requirements (4.32) and (4.33) of Theorem 4.4.4 imposed on $X(\mu)$ with $\beta > 1/2$. Let the right-hand side vector b be replaced by $b + \gamma \Delta b$, where $\gamma \in \mathbb{R}$ and $\Delta b \in \mathbb{R}^m$. Then the NT bound evaluated at (X, y, S) converges to the symmetrized partition bound as $\mu \downarrow 0$.*

Section 4.4 implies that we can formulate perturbations of the cost matrix C as perturbations of the right-hand side using the primal-dual symmetry. However, due to the interchange of roles between the primal and dual variables, the scaling matrix for the alternative formulation (D1) and (P1) is given by

$$\tilde{W} := W(S, X) = S^{\frac{1}{2}}(S^{\frac{1}{2}}XS^{\frac{1}{2}})^{-\frac{1}{2}}S^{\frac{1}{2}}, \quad (4.101)$$

so that $\tilde{W}X\tilde{W} = S$. It follows then that $\tilde{W} = W^{-1}$. Consequently, the assumptions on \bar{W} and X^* should be replaced by the corresponding assumptions on $\sqrt{\mu} \tilde{W}$ and S^* , or equivalently by $\sqrt{\mu} W^{-1}$ and S^* . We then have the following corollary for perturbations of the cost matrix.

Corollary 4.5.1 *Let (P) and (D) satisfy Assumptions 4.2.1, 4.2.2 and 4.2.4. Assume that (X, y, S) is a primal-dual strictly feasible iterate satisfying (4.88) such*

that the corresponding $\sqrt{\mu} W^{-1}$ satisfies the dual requirements of (4.32) and (4.33) of Corollary 4.4.1 imposed on $S(\mu)$ with $\beta > 1/2$. Let the cost matrix C be replaced by $C + \gamma C$, where $\gamma \in \mathbb{R}$ and $C \in \mathcal{S}^n$. Then, the NT bound evaluated at (X, y, S) converges to the symmetrized partition bound as $\mu \downarrow 0$.

4.6 Conclusion

We have studied the asymptotic behavior of the bounds arising from the interior-point perspective on sensitivity analysis in SDP developed in [71] in comparison with the optimal partition bounds of [18]. For perturbations of the right-hand side vector and the cost matrix, we have shown that the interior-point bounds evaluated at the iterates on the central path using the Monteiro-Zhang family of search directions converge to the symmetrized version of the optimal partition bounds under appropriate nondegeneracy assumptions. Furthermore, the two bounds asymptotically coincide even in the absence of strict complementarity if the distance between the point on the central path and the analytic center satisfies a certain requirement. We have then extended the same relationship to a central path neighborhood using the Nesterov-Todd direction.

The computational cost of evaluating the interior-point bound is the same as that of a single interior-point iteration. The sufficient conditions for the asymptotic coincidence of the two bounds outlined in this paper require that the iterates be in close proximity of the central path. In practical implementations, this can be achieved by taking a few centering steps to get closer to the central path when

the duality gap is sufficiently small. Our limited computational tests suggest that the interior-point bounds tend to approximate the symmetrized optimal partition bounds fairly well even when the iterates are not too close to the central path – especially if the Nesterov-Todd direction is used. Although the interior-point bounds using any member of the Monteiro-Zhang family coincide on the central path, our test problems indicate a much more rapid deterioration of the interior-point bounds using the AHO, H.K.M and the dual H.K.M directions than those using the NT direction. Furthermore, the fact that the interior-point bound using the NT direction reduces to a convenient norm bound as opposed to two eigenvalue bounds for most of the other directions in the Monteiro-Zhang family makes the analysis somewhat simpler. This observation in conjunction with the computational results is the reason why we considered the NT direction in further detail in this paper.

Finally, we note that the cost of computing the interior-point bound can in practice be lowered by evaluating it at the penultimate iterate rather than the final iterate generated during the course of the interior-point algorithm. In this case, the evaluation and factorization of the Schur complement matrix, which usually requires the major computational effort in an interior-point iteration, will already be available. The computation of the interior-point bound then reduces to solving a couple of triangular systems followed by a few matrix-matrix products.

Chapter 5

Computational Experiments

In this chapter, we present computational results to illustrate the behavior of the interior-point bounds on randomly generated linear programming (LP) and semidefinite programming (SDP) problems. These results are intended to shed some light on the behavior of our bounds in practice.

The notation used in this chapter will be consistent with the notation of Chapters 2, 3 and 4. The randomly generated LP and SDP instances will be in their respective standard forms.

5.1 Linear Programming

In Chapter 2, we studied the asymptotic behavior of the interior-point bounds for perturbations of the right-hand side and the cost vectors in comparison with the optimal basis (or, equivalently, optimal partition) bounds under nondegeneracy. We then extended our analysis to degenerate LPs in Chapter 3.

In this section, complementary to the theoretical foundation developed in Chapters 2 and 3, we present some computational results to illustrate the performance of the interior-point bounds evaluated at near-optimal solutions of randomly generated LP instances.

5.1.1 Generating Random LP Instances

We have generated random LPs with $m = 200$ and $n = 400$. These values of m and n are chosen in order to make the primal and dual problems look symmetric. Here, by symmetry, we refer to the fact that the dual problem formulated in the primal form as outlined in Section 3.3 has precisely the same dimensions as the original primal problem. The input parameters are the cardinality of the index set \mathcal{B} ($|\mathcal{B}|$), which can vary from 0 to 400, and the dimension of the primal optimal set ($\dim(\Omega_P)$), which is given by $|\mathcal{B}| - \text{rank}(B)$ by Proposition 3.2.1, where B is the submatrix of the coefficient matrix A consisting of those columns corresponding to the indices in \mathcal{B} . These two parameters together determine the dimension of the dual optimal set $\dim(\Omega_D)$ and $\text{rank}(B)$. The ability to control all these parameters allows us to incorporate all scenarios of primal and dual degeneracies into the randomly generated LPs.

Our goal is to generate random LPs with prespecified degeneracy requirements so that the resulting problems will help illustrate the behavior of the interior-point bounds under various degeneracy scenarios. Since our theoretical results pertain to interior-point bounds evaluated at strictly feasible, near-optimal solutions, we start with generating a random, strictly complementary optimal solution (x^*, y^*, s^*) .

The positive components of x^* and s^* are generated independently according to the absolute value of a standard normal distribution and the entries of y^* are chosen independently from the standard normal distribution.

The next step is to generate a feasible direction $(\Delta x, \Delta y, \Delta s)$ at this optimal solution simultaneously with a suitable coefficient matrix A . We say that $(\Delta x, \Delta y, \Delta s)$ is a feasible direction at (x^*, y^*, s^*) if there exists $\alpha > 0$ such that $(x^*, y^*, s^*) + t(\Delta x, \Delta y, \Delta s)$ is feasible for the primal and dual problems for all $t \in [0, \alpha)$. Consequently, the near-optimal strictly feasible solutions at which the interior-point bounds are evaluated are obtained by taking different multiples of this feasible direction from the optimal solution (x^*, y^*, s^*) .

Note that the feasible direction $(\Delta x, \Delta y, \Delta s)$ and the coefficient matrix $A = [B \ N]$ partitioned according to the optimal partition should have the following three properties:

1. $\Delta x_N > 0$ and Δx_B satisfy

$$B\Delta x_B + N\Delta x_N = 0.$$

2. $\Delta y, \Delta s_B > 0$ and Δs_N satisfy

$$B^T \Delta y + \Delta s_B = 0, \quad \text{and} \quad N^T \Delta y + \Delta s_N = 0.$$

3. The matrix B satisfies

$$\text{rank}(B) = |\mathcal{B}| - \dim(\Omega_P), \quad \text{and} \quad \text{rank}(B) = m - \dim(\Omega_D).$$

Based on the input parameters, we compute $\text{rank}(B)$ and generate a square matrix B_1 of size $\text{rank}(B)$ with independent standard normal entries. Next, we choose the

(positive) subvector Δs_B^1 of Δs_B whose size is determined by B_1 according to the absolute value of a standard normal distribution. We compute $\Delta y_1 := -B_1^{-T} \Delta s_B^1$ and then set $\Delta y^T := [\Delta y_1^T, 0]$. We then add linearly dependent columns to B to have a total of $|\mathcal{B}|$ columns. In doing so, we generate an appropriate matrix K with independent standard normal entries such that $\Delta s_B^2 := K^T \Delta s_B^1 > 0$ and then set

$$\Delta s_B = \begin{bmatrix} \Delta s_B^1 \\ \Delta s_B^2 \end{bmatrix}$$

and $B_1 \leftarrow [B_1 \quad B_1 K]$. In order to ensure that B has a total of m rows, we finally add linearly dependent rows to B_1 (if necessary) by first generating a matrix L with independent standard normal entries and then letting LB_1 be the additional rows. We thus obtain B . The next step is the formation of the submatrix N of A . To do so, we first generate Δx_B and $\Delta x_N > 0$ independently from a standard normal distribution. We then form a matrix N with independent standard normal entries and modify it by distributing the residual vector $-B\Delta x_B - N\Delta x_N$ to its columns so that $B\Delta x_B + N\Delta x_N = 0$. Consequently, $(\Delta x, \Delta y, \Delta s)$ and $A = [B \quad N]$ possess the desired properties.

We finally choose b and c to make (x^*, y^*, s^*) feasible and hence optimal. The MATLAB code used to generate the random LP instances having the properties above is given in Appendix A.

Having generated a random LP with the prespecified degeneracy, we obtain a strictly feasible, near-optimal solution by perturbing the known strictly complementary optimal solution in the direction $(\Delta x, \Delta y, \Delta s)$. Next, random perturbations Δb of b and Δc of c with unit length are generated in the correct subspaces so that

the LPs (AUX1) and (AUX2) of Section 3.2 used to determine the optimal partition bounds have nontrivial optimal solutions. This is simply done by setting Δb to a random linear combination (chosen independently from the standard normal distribution) of the columns of B and similarly by setting Δc_B to a random linear combination of the columns of B^T and finally augmenting it by another vector Δc_N with independent standard normal entries in order to obtain Δc . We compute the interior-point bounds evaluated at those near-optimal solutions. For perturbations of the right-hand side vector b , we use (3.15) to compute the interior-point bound at a given iterate. For perturbations of the cost vector c , we use the right-hand side of (3.16). However, we note that the straightforward computation of the right-hand side of (3.16) leads to numerically unstable results. Consequently, we use a QR factorization of the matrix DA^T first and perform the computations accordingly without having to form AD^2A^T explicitly. This strategy produces much more stable and reliable results. The computation of the interior-point bound (3.15), on the other hand, does not seem to be sensitive to the way in which the operations are performed.

Our goal is to compare the interior-point bounds with the optimal solutions to (AUX1) and (AUX2) as well as the optimal solutions to the symmetrized LPs (SA1) and (SA2) of Section 3.2, where the initially generated strictly complementary optimal solution is used to symmetrize the constraints. The resulting LPs are solved using the MATLAB-CPLEX interface written by Nathan Edwards during his Ph.D. studies at Cornell University. Some documentation is available in his Ph.D. thesis [12].

5.1.2 Results

We present our results for various degeneracy scenarios in Tables 5.1, 5.2 and 5.3. Eighteen instances with various levels of primal-dual degeneracy are reported. For each instance, the interior-point bounds are evaluated at two iterates corresponding to each row. DP and DD are the dimensions of the primal and dual optimal sets, respectively. μ is the duality measure of the iterate given by $x^T s/n$, and γ is the parameter of the narrowest wide central-path neighborhood containing the iterate. More precisely, $\gamma := \min_i(x_i s_i)/\mu$. (AUX1) and (AUX2) are the minima of the absolute values of the optimal values of the corresponding minimization and maximization problems (symmetrizations). (SA1) and (SA2) are the optimal values of the symmetrized maximization problems, where the optimal solution (x^*, y^*, s^*) is used to symmetrize the constraints of (AUX1) and (AUX2). Finally, IPB and IPC are the upper interior-point bounds for changes in b and c evaluated at the corresponding iterates using (3.15) and (3.16).

We chose to let $|\mathcal{B}|$ range from 40 to 360 in increments of 40. For each value of $|\mathcal{B}|$ we generated two random LP instances such that the first instance either has a unique primal optimal solution or a unique dual optimal solution. The second instance, on the other hand, has multiple primal and dual optimal solutions such that their respective dimensions are set at an intermediate value within their ranges given by Proposition 3.2.1. Note that the difference $n - m$ is an upper bound on $\dim(\Omega_P)$ in order to satisfy the full row rank assumption on A . Finally, for each instance, we considered two strictly feasible iterates. The first one is very close to optimality with a duality measure on the order of 10^{-6} or 10^{-7} . The second one

has a duality gap on the order of 10^{-3} or 10^{-4} .

The situation considered in Chapter 2 is illustrated by Instance 9, where both the primal and dual problems have a unique nondegenerate optimal solution. As expected, the interior-point bounds converge to the symmetrized optimal partition bounds (SA1) and (SA2) as $\mu \downarrow 0$.

In the presence of degeneracy, Chapter 3 provides two different results. The asymptotic coincidence between the interior-point bounds and the symmetrized optimal partition bounds continues to hold for perturbations of b if the primal optimal solution is unique, as illustrated by Instances 1, 3, 5 and 7, and for perturbations of c if the dual optimal solution is unique, as illustrated by Instances 11, 13, 15 and 17. Observe that the interior-point bounds converge to the symmetrized partition bounds even though γ is very small, which is typical in practical implementations of the interior-point methods. This observation encourages the applicability of our results in practice.

For the remaining degeneracy scenarios, the interior-point bounds lie within a factor of the symmetrized partition bounds as given by Theorem 3.5.1 and Corollary 3.5.1. It is worth noting, however, that the actual ratio seems to be much better than the theoretical worst-case ratios (3.76) and (3.77). In our extensive computational tests, the ratio was never worse than a hundredth although the predicted lower bounds (3.76) and (3.77) are on the order of 10^{-5} in most of the instances. Therefore, the interior-point bounds still provide useful information even for degenerate linear programs.

Finally, we note that although the condition number of AD^2A^T blew up in all of

Table 5.1: Computational Results for Linear Programming – Part 1 ($m = 200, n = 400$)

Ins	$ \mathcal{B} $	DP	DD	μ	γ	(AUX1)	(SA1)	IPB	(AUX2)	(SA2)	IPC
1	40	0	160	1.3 e-6	4.81 e-6	5.472	5.472	5.472	64.163	36.864	3.058
				1.3 e-3	5.29 e-6			6.006			
2	40	20	180	4.4 e-6	8.84 e-5	1133.628	644.564	152.758	58.149	47.928	6.496
				2.2 e-3	8.82 e-5			153.663			
3	80	0	120	1.5 e-6	1.01 e-4	19.574	19.574	19.574	67.039	30.585	3.707
				1.5 e-3	1.02 e-4			19.826			
4	80	40	160	3.9 e-7	1.08 e-4	1755.081	687.031	161.749	499.324	243.944	9.976
				9.7 e-4	1.07 e-4			148.601			
5	120	0	80	7.6 e-7	1.32 e-7	11.882	11.882	11.884	28.685	22.861	2.994
				7.6 e-4	1.32 e-7			13.049			
6	120	60	140	7.6 e-7	8.97 e-5	6447.073	2195.607	311.393	223.653	197.322	25.935
				3.8 e-4	8.73 e-5			311.544			

Table 5.2: Computational Results for Linear Programming – Part 2 ($m = 200, n = 400$)

Ins	$ \mathcal{B} $	DP	DD	μ	γ	(AUX1)	(SA1)	IPB	(AUX2)	(SA2)	IPC
7	160	0	40	5.6 e-7	1.24 e-5	3.412	3.412	3.412	9.335	6.936	1.013
				2.8 e-4	1.24 e-5			3.560			1.014
8	160	80	120	2.1 e-7	1.56 e-4	4863.709	2094.253	101.528	355.701	281.539	47.245
				1.0 e-4	1.51 e-4			102.040			47.239
9	200	0	0	1.1 e-7	7.26 e-5	0.676	0.676	0.677	0.00830	0.00830	0.00829
				1.1 e-4	7.99 e-5			0.778			0.00823
10	200	100	100	6.1 e-7	2.79 e-5	9246.865	4030.466	499.392	69.484	66.870	3.083
				3.0 e-4	2.79 e-5			500.255			3.087
11	240	40	0	6.3 e-8	1.73 e-4	248.922	149.659	11.273	1.073	1.073	1.073
				1.6 e-4	1.97 e-4			11.019			0.911
12	240	120	80	3.0 e-8	1.56 e-5	9658.737	4427.046	674.833	51.501	49.787	15.313
				1.5 e-4	9.13 e-6			673.371			15.230

Table 5.3: Computational Results for Linear Programming – Part 3 ($m = 200, n = 400$)

Ins	$ \mathcal{B} $	DP	DD	μ	γ	(AUX1)	(SA1)	IPB	(AUX2)	(SA2)	IPC
13	280	80	0	3.7 e-7	9.87 e-5	450.587	220.656	9.297	1.561	1.561	1.561
				1.9 e-4	9.88 e-5			9.307			1.493
14	280	140	60	2.5 e-7	1.67 e-5	10520.160	3653.153	378.616	483.922	377.252	32.429
				1.2 e-4	1.66 e-5			377.605			32.416
15	320	120	0	2.2 e-7	2.21 e-5	845.529	295.456	15.259	3.028	3.028	3.028
				1.1 e-4	2.10 e-5			15.260			2.879
16	320	160	40	7.3 e-7	5.12 e-5	12099.260	4067.016	224.887	17.593	15.794	3.397
				3.6 e-4	5.12 e-5			225.798			3.400
17	360	160	0	8.5 e-7	1.14 e-4	1530.485	552.014	87.875	2.636	2.636	2.637
				4.3 e-4	1.20 e-4			87.880			2.715
18	360	180	20	6.7 e-7	1.30 e-4	10578.480	3784.980	217.618	4.951	4.945	1.339
				3.3 e-4	1.29 e-4			216.505			1.348

the degenerate instances as expected, the computation of our bounds still yielded numerically stable results.

5.2 Semidefinite Programming

In Chapter 4, we studied the asymptotic behavior of the interior-point bounds evaluated on the central path using the Monteiro-Zhang family of search directions in comparison with the (symmetrized) optimal partition bounds under appropriate nondegeneracy assumptions and proved that the two bounds asymptotically coincide. We then extended this result to a very narrow central path neighborhood using the Nesterov-Todd (NT) direction.

In this section, we illustrate the behavior of the interior-point bounds for perturbations of the right-hand side and cost parameters using the AHO, H..K..M and NT search directions on randomly generated SDP instances. The computational results are intended to exhibit the asymptotic coincidence with the symmetrized optimal partition bounds and the relative performances of the interior-point bounds using the three search directions evaluated at iterates lying in different central path neighborhoods.

5.2.1 Generating Random SDP Instances

Our theoretical results of Chapter 4 pertain to iterates lying on the central path or in a very narrow neighborhood of it. In practice, however, the “exact” computation of a point on the central path is usually unnecessary and is likely to require extra

computational effort. With this in mind, we then start with generating a random iterate lying exactly on the central path with some prespecified duality measure μ and generate the SDP data (the matrices A_i , the right-hand side vector b and the cost matrix C) so as to make the iterate feasible for the resulting SDP instance given in standard form. This procedure has the advantage of eliminating the extra computational effort of finding a point on the central path.

We apply the following procedure to generate a random iterate (X, y, S) on the central path. First, we need to exercise some care in determining the “small” and “big” eigenvalues of X . In particular, we need some information about the number of each set of eigenvalues of X and S . We also need to make sure that the resulting SDP instance possesses the required nondegeneracy assumptions presented in Chapter 4. Fortunately, Alizadeh, Haeberly and Overton [4] provide necessary conditions on the ranks of X^* and S^* in order for an optimal solution (X^*, y^*, S^*) to be primal and dual nondegenerate.

Theorem 5.2.1 *Let (X^*, y^*, S^*) be primal and dual nondegenerate and optimal for a given SDP instance, with $\text{rank}(X) = r$ and $\text{rank}(S) = s$. Then*

$$n - f(n^{\bar{2}} - m) \leq r \leq f(m) \tag{5.1}$$

and

$$n - f(m) \leq s \leq f(n^{\bar{2}} - m), \tag{5.2}$$

where

$$f(k) = \lfloor h \rfloor, \quad \text{with } h \text{ the positive real root of } h(h+1) = 2k.$$

It is also shown in [4] that the bounds (5.1) and (5.2) are satisfied almost everywhere in the Lebesgue sense for a random SDP instance. Furthermore, strict complementarity and primal-dual nondegeneracy also hold for a random SDP instance almost surely.

The central path converges to the analytic center as $\mu \downarrow 0$. Therefore, the number of small and big eigenvalues of $X(\mu)$ and $S(\mu)$ will be related to the rank of X^* and S^* , respectively, for μ sufficiently small. In our tests, we set $\mu = 10^{-2}$. We then choose the number of big eigenvalues of X to be halfway between its limits given by (5.1). We treat S similarly using (5.2). The big eigenvalues of X and S are generated uniformly in the interval $[1, 2]$. The fact that the product of the eigenvalues of X and S is equal to μ then determines the small eigenvalues of X and S as well. If $r + s < n$, where $r = \text{rank}(X)$ and $s = \text{rank}(S)$, we also need to generate the intermediate eigenvalues of X and S . In doing so, we use the log-normal distribution to make sure that the intermediate eigenvalues of X and S have the same probability distribution. More precisely, if Z is a random variable with the log-normal distribution, then Z and $1/Z$ have the same probability distributions. Therefore, we generate the remaining $n - r - s$ eigenvalues of X according to this distribution and multiply each of them by $\sqrt{\mu}$. The corresponding eigenvalues of S are given by the reciprocal of the eigenvalues of X multiplied by $\sqrt{\mu}$.

We next generate a random orthogonal matrix (by taking the QR factorization of a matrix with independent standard normal entries) to constitute the eigenvectors of X and S . We then generate y whose entries are chosen independently from the standard normal distribution.

In order to compare the performances of the interior-point bounds resulting from the three search directions at feasible iterates lying in different central path neighborhoods, we determine a feasible direction $(\Delta X, \Delta y, \Delta S)$ from (X, y, S) . Recall that a direction $(\Delta X, \Delta y, \Delta S)$ is called a feasible direction from (X, y, S) if there exists $\alpha > 0$ such that $(X, y, S) + t(\Delta X, \Delta y, \Delta S)$ is feasible for all $t \in [0, \alpha)$. We first generate a random symmetric matrix ΔX whose eigenvalues are distributed according to the standard normal distribution. We then choose the matrices A_i , $i = 1, \dots, m$ so that $A_i \bullet \Delta X = 0$ for each i . To do this, we first generate a random symmetric matrix A_i with eigenvalues chosen from a standard normal distribution. Then we replace A_i by $A_i - [(A_i \bullet \Delta X)/(\Delta X \bullet \Delta X)]\Delta X$ so that A_i and ΔX are orthogonal. This is equivalent to projecting A_i orthogonally onto the space spanned by ΔX and subtracting the result from A_i . Next, Δy is chosen randomly from the standard normal distribution and $\Delta S \leftarrow -\sum_{i=1}^m \Delta y_i A_i$.

This choice of $(\Delta X, \Delta y, \Delta S)$ ensures that $(X, y, S) + t(\Delta X, \Delta y, \Delta S)$ will satisfy the linear constraints of the primal and dual problems for any t . The only remaining constraints are the positive semidefiniteness of $X + t\Delta X$ and $S + t\Delta S$. We thus set α to be

$$\alpha = \min\{\max\{t : X + t\Delta X \succeq 0\}, \max\{t : S + t\Delta S \succeq 0\}\} \quad (5.3)$$

so that $(X, y, S) + t(\Delta X, \Delta y, \Delta S)$ will be (strictly) feasible for any $t \in [0, \alpha)$.

Finally, the right-hand side b and the cost matrix C is chosen such that (X, y, S) is feasible for the resulting SDP instance. That is, $b_i \leftarrow A_i \bullet X$ and $C \leftarrow \sum_{i=1}^m y_i A_i + S$. The MATLAB code used to generate the random SDP instances is given in Appendix B.

Next, we solve the SDP instance to “optimality” using SDPT3, a MATLAB-based path-following interior-point code for SDP developed by Toh, Todd and Tütüncü [63]. The code is available at <http://www.math.nus.edu.sg/~mattohkc/>. We initiate the interior-point algorithm from the iterate on the central path, which is an excellent starting point, and solve the SDP instance to optimality. The code is terminated as soon as the duality gap falls below 10^{-12} . Typically, the desired accuracy is achieved only after 6 or 7 iterations.

Let us denote the “optimal” solution given by SDPT3 by (X^*, y^*, S^*) . We obtain an approximate optimal solution $(\tilde{X}, \tilde{y}, \tilde{S})$ by truncating the small eigenvalues of X^* and S^* and setting $\tilde{y} = y$. More specifically, the eigenvalues of X^* and S^* smaller than the square root of the final duality measure are set to 0. This procedure allows us to generate random perturbations Δb and ΔC in the correct spaces as outlined in Sections 1.2.4 and 4.2.1. Basically, we compute the eigenvalue decomposition of \tilde{X} and \tilde{S} . Let Q_P and Q_D denote the matrices consisting of the eigenvectors of \tilde{X} and \tilde{S} , respectively. We then generate a symmetric matrix U by choosing a matrix with independent standard normal entries and adding to it its transpose. We set $\Delta b_i \leftarrow A_i \bullet Q_P U Q_P^T$. Similarly, we generate another random symmetric matrix V and a vector w with independent standard normal entries and set $\Delta C \leftarrow \sum_{i=1}^m w_i A_i + Q_D V Q_D^T$.

We then compute the optimal partition bounds for perturbations of the right-hand side and the cost matrix using the approximate optimal solution $(\tilde{X}, \tilde{y}, \tilde{S})$ and the random perturbations Δb and ΔC .

Finally, we evaluate the interior-point bounds using the three search directions

on three sets of iterates. The first set is obtained by the central path point and those obtained by perturbing it in the feasible direction $(\Delta X, \Delta y, \Delta S)$ using different multiples of α . This enables us to compare the performances of the three search directions on iterates lying in different central path neighborhoods. The second set of iterates is used to observe the quality of the interior-point bounds on iterates with smaller duality gaps. Basically, an appropriate convex combination of the central path point and the near-optimal solution (X^*, y^*, S^*) is chosen to ensure that the duality gap of the resulting iterate is on the order of 10^{-8} . We then evaluate the interior-point bounds on this point as well as on the points obtained from this iterate using the aforementioned feasible direction. The third set of iterates consists of the single point (X^*, y^*, S^*) , the near-optimal solution returned by SDPT3. Since such a point is still strictly feasible, we evaluate the interior-point bounds once more at this point to get an idea about their behavior when the duality gap is on the order of 10^{-12} .

5.2.2 Results

We present our results for only one random SDP instance with $m = 100$ and $n = 20$. The computational results for perturbation of the right-hand side vector are given in Table 5.4, and those for perturbation of the cost matrix are given in Table 5.5. We choose to illustrate the results for only one SDP instance since the other SDP instances produced very similar results. The instance satisfies strict complementarity and primal-dual nondegeneracy.

Each row corresponds to a strictly feasible iterate (X, y, S) . In particular, the

Table 5.4: Computational Results for Semidefinite Programming – Perturbations of b ($m = 100, n = 20$)

Iterate	μ	γ	AHO1	AHO2	HKM1	HKM2	NT1	NT2
1	1.00 e-2	1.0000	2.113704	1.599078	2.113704	1.599078	2.113704	1.599078
2	9.94 e-3	0.9056	2.119567	1.598852	2.114309	1.600411	2.113859	1.600400
3	9.69 e-3	0.5158	2.127617	1.597814	2.126312	1.607442	2.114761	1.606859
4	9.45 e-3	0.1059	2.094671	1.596556	2.156184	1.622831	2.117037	1.618467
5	9.39 e-3	0.0053	0.685585	1.596224	2.166573	1.630814	2.119153	1.626547
6	9.39 e-3	0.0011	0.375809	1.596210	2.167040	1.631235	2.119553	1.627888
7	9.39 e-3	0.0001	0.236148	1.596207	2.167146	1.631331	2.119785	1.628634
8	2.82 e-8	0.8860	2.106860	1.418419	2.033163	1.418420	2.029415	1.418420
9	2.80 e-8	0.8863	2.113602	1.418420	2.034404	1.418420	2.029415	1.418420
10	2.73 e-8	0.5259	2.124318	1.418420	2.048433	1.418420	2.029415	1.418420
11	2.66 e-8	0.1080	2.088043	1.418420	2.079225	1.418420	2.029415	1.418420
12	2.64 e-8	0.0054	0.750970	1.418420	2.090121	1.418420	2.029415	1.418420
13	2.64 e-8	0.0011	0.423538	1.418420	2.090628	1.418420	2.029415	1.418420
14	2.64 e-8	0.0001	0.282813	1.418420	2.090743	1.418420	2.029415	1.418420
15	7.95 e-14	0.5349	1.476841	1.369820	2.101586	1.418419	2.029414	1.418419

Table 5.5: Computational Results for Semidefinite Programming – Perturbations of C ($m = 100, n = 20$)

Iterate	μ	γ	AHO1	AHO2	HKM1	HKM2	NT1	NT2
1	1.00 e-2	1.0000	3.662198	2.402326	3.662198	2.402326	3.662198	2.402326
2	9.94 e-3	0.9056	3.662090	2.402097	3.661353	2.401857	3.661321	2.401849
3	9.69 e-3	0.5158	3.661288	2.400184	3.657212	2.399037	3.656091	2.398762
4	9.45 e-3	0.1059	3.659906	2.393931	3.650903	2.392418	3.639739	2.391092
5	9.39 e-3	0.0053	3.659493	1.973526	3.648827	2.290588	3.590656	2.383296
6	9.39 e-3	0.0011	3.659475	0.876187	3.648732	1.434351	3.565165	2.381257
7	9.39 e-3	0.0001	3.659471	0.173868	3.648711	0.494759	3.544054	2.379869
8	2.82 e-8	0.8860	3.589994	2.333080	3.588104	2.333079	3.587129	2.333079
9	2.80 e-8	0.8863	3.590089	2.333079	3.588316	2.333079	3.587129	2.333079
10	2.73 e-8	0.5259	3.590102	2.333080	3.589494	2.333079	3.587129	2.333079
11	2.66 e-8	0.1080	3.589526	2.333080	3.591194	2.333079	3.587129	2.333079
12	2.64 e-8	0.0054	3.589303	2.333079	3.591674	2.333079	3.587128	2.333079
13	2.64 e-8	0.0011	3.589293	2.333078	3.591695	2.333079	3.587128	2.333079
14	2.64 e-8	0.0001	3.589292	2.333072	3.591700	2.333073	3.587128	2.333079
15	7.95 e-14	0.5349	3.397240	2.330607	3.595173	2.333079	3.587128	2.333079

iterates are numbered consistently in Table 5.4 and Table 5.5. The three sets of iterates described in the previous subsection are presented in three sections in each table. μ denotes the duality measure given by $X \bullet S/n$, and γ denotes the “centrality” measure, i.e., $\gamma := \min \lambda(X^{1/2}SX^{1/2})/\mu$, where $\lambda(\cdot)$ denotes the vector of eigenvalues. Consequently, γ ranges from 1 on the central path to 0 as the iterate approaches the boundary of the feasible region. The remaining columns tabulate the interior-point bounds using the three search directions AHO, H..K..M, and NT. For each direction, two upper bounds (given by Proposition 4.2.2 for perturbations of b and given by Proposition 4.2.2 for perturbations of C) are computed. The upper interior-point bounds are given by the smaller of these two quantities.

For perturbation of the right-hand side, the optimal partition bound is given by $(-2.029415, 1.418419)$. Therefore, the symmetrized partition bound is $(-1.418419, 1.418419)$. The results presented in Table 5.4 confirm that the quality of the interior-point bounds improves (in the sense that they better approximate the symmetrized partition bounds) as the duality measure μ decreases towards 0. For the AHO direction, the interior-point bound at the “optimal solution” is not accurate possibly due to round-off errors. All three search directions yield the same bounds on the central path as expected. The computational results also show that the NT bound is very robust in the sense that it seems very insensitive to the centrality measure. The H..K..M bound is also relatively robust. However, the AHO direction seems to be very sensitive to the centrality measure. The quality of the AHO bound deteriorates as the centrality measure falls below 0.1. Both the H..K..M and the NT directions produce reliable results for iterates with relatively small duality gaps.

This holds for the AHO direction if the iterate is relatively well centered and if the duality gap is not too small.

The interior-point bounds for perturbations of C were computed by reformulating the dual problem in the primal form using the symmetry argument of Section 4.3. We chose this way since the straightforward computation of the interior-point bound leads to significant numerical errors as in the LP case. However, we note that there may be other direct and efficient ways of computing the interior-point bound for perturbations of C for specific search directions without having to put the dual problem into the primal form. The optimal partition bound for perturbation of C is given by $(-3.587128, 2.333079)$, and the symmetrized partition bound is $(-2.333079, 2.333079)$. Similar observations for the computational results of Table 5.5 continue to hold for perturbations of C . In contrast to the case for perturbations of b , the H.K.M bound seems to be sensitive to the centrality measure for relatively large values of the duality measure. Once again, the bounds improve for all directions as the duality measure decreases to 0. As before, the NT bound seems very insensitive to the centrality measure. The quality of the interior-point bounds seems to be less dependent on the centrality measure for smaller values of the duality measure. Finally, the bounds again coincide on the central path.

The observations above for perturbations of the right-hand side and the cost matrices hold for all the random instances we have considered. Based on our numerical experiments, we suggest that the NT direction be used to compute the interior-point bounds since it is very robust with respect to the centrality measure and since it appears to be numerically stable even for very small values of the duality measure.

5.3 Concluding Remarks

In this dissertation, we have developed and analyzed an approach to sensitivity analysis on the right-hand side and cost parameters in LP and SDP, which relies entirely on the tools available from interior-point theory.

For the LP case, we have shown in Chapter 2 that the interior-point bounds resulting from our approach asymptotically coincide with the optimal partition bounds under a nondegeneracy assumption. Chapter 3 is devoted to degenerate LPs. For a specific kind of degeneracy, we have extended the same asymptotic coincidence result of the previous chapter. For general degeneracy, we have shown that our approach still provides provably useful information with a very moderate cost. Chapter 4 dealt with the SDP case. For SDP instances where strict complementarity holds, we have shown that our interior-point bounds evaluated on the central path asymptotically coincide with the optimal partition bounds for all search directions in the Monteiro-Zhang family. Furthermore, the same result continues to hold for iterates lying in an appropriate (but very narrow) central path neighborhood if one uses the Nesterov-Todd direction. We have also proved the same theoretical results for SDP instances where strict complementarity fails as long as the central path converges to the analytic center in a relatively controlled manner. Finally, we reported some computational results on randomly generated LP and SDP instances in this chapter. The results are in agreement with the theoretical findings of the previous chapters. In addition, numerical results suggest that our approach may provide accurate information even when some of the assumptions made in Chapters 3 and

4 are not satisfied.

The proposed interior-point perspective of this dissertation stems from the desire to recover a feasible, near-optimal point for the perturbed problem in a single interior-point iteration starting from a near-optimal solution for the original problem. The interior-point bound then is defined to be the range of a perturbation for which the scheme outlined above can successfully be carried out.

Our interior-point approach, however, does not deal with the case when a perturbation is large enough so that a feasible point cannot be recovered in a single interior-point iteration. A closely related topic is reoptimization, i.e., finding an optimal solution of the perturbed problem using the information obtained from the original problem. Reoptimization using the simplex method is used very often in practice since a basic optimal solution of the original problem usually serves as an excellent starting point for reoptimization of the perturbed problem. Due to the combinatorial nature of the simplex method, a complexity analysis for reoptimization via the simplex method is not possible. Due to reasons similar to those outlined in Chapter 1, interior-point methods have been considered not suitable for reoptimization after a data perturbation. Yildirim and Wright [72] studied this problem for LP and proposed warm-start strategies for reoptimization using interior-point methods. The warm-start strategies are obtained by applying a correction to the iterates generated during the solution of the original problem in order to obtain a feasible and relatively well-centered iterate for the perturbed problem. Their study also includes a complexity analysis of the warm-start schemes.

Yildirim and Wright deal with the case in which the dimension of the problem

data does not change. This situation arises, for example, when linearization methods are used to solve nonlinear problems, as in the sequential linear programming algorithm. Reoptimization is also frequently used when the number of variables or constraints in the problem is increased, and the dimensions of the problem data objects are correspondingly expanded. The latter situation arises in solving subproblems arising from cutting-plane or column-generation algorithms, for example. The reader is also referred to Gondzio [21], Mitchell and Borchers [41] and Gondzio and Vial [22] for consideration of warm-start strategies in a cutting-plane scheme.

The interior-point perspective developed in this dissertation pertains to linear programming and semidefinite programming. However, this perspective can easily be extended to general convex conic programming problems. We would like to study this extension, which might allow us to unify our results in a general framework in the future.

The warm-start strategies for reoptimization developed by Yıldırım and Wright [72] deal only with feasible interior-point methods for linear programming. The computational tests that will help determine the practical usefulness of these strategies are underway. We would like to investigate possible extensions to wider classes of problems, such as convex quadratic programs and linear complementarity problems.

Appendix A

MATLAB Code for Generating a Random LP Instance

```
function [A,b,c,x,y,s,dx,dy,ds] = generate(m,n,k,dimp)

% Input parameters:
% m : number of primal constraints
% n : number of primal variables
% k : number of indices in the basic index set, 0 <= k <= n
% dimp : dimension of the primal optimal set, k-m <= dimp <= n-m

% Output:
% A,b,c : Data of the LP instance satisfying the input parameters
% x,y,s : Strictly complementary optimal solution
```

```
% dx,dy,ds : Feasible direction at (x,y,s)

% The rank of matrix B
rb = k - dimp;

% The dimension of the dual optimal set
dimd = m - rb;

% Generate a random square matrix with rank rb
randn('state',0);
B = randn(rb,rb);

% Now generate a positive perturbation corresponding
% to the basic part of s.
dsb1 = abs(randn(rb,1));
dy = -B'\dsb1;

% Add dimp columns (linearly dependent).
% Want dsb2 to depend on dsb1 and be positive.
if (rb > 0) & (dimp > 0)
    K = randn(rb,dimp);
    dsb2 = K'*dsb1;
```

```

    for j = 1:dimp,
        if (dsb2(j) < 0),
            dsb2(j) = - dsb2(j);
            K(:,j) = - K(:,j);
        end;
    end;

    B = [B, B*K];
    dsb = [dsb1; dsb2];
end;

% Add dimd rows
if (dimd > 0) & (rb > 0)
    B = [B; randn(dimd,rb)*B];
end;

dy = [dy; zeros(dimd,1)];

% Generate A
% Generate N such that there is a positive vector
% in the null space of A
dxn = randn(k,1);
N = randn(m,n-k);
dxn = abs(randn(n-k,1));

```

```
% Want  $B \cdot dx_b + N \cdot dx_n = 0$ .
dx_b = randn(k,1);
last = - N*dx_n - B*dx_b;

% Distribute the ‘‘special’’ column last to the previous ones
sca = (1/(sum(dx_n)))*last;
N = N + sca*ones(1,n-k);

% Therefore,  $dx_n > 0$  and  $dx_b$  are in the null space of A
A = [B N];

% Generate strictly complementary optimal x and s
x = [abs(randn(k,1)); zeros(n-k,1)];
x_b = x(1:k);
s = [zeros(k,1); abs(randn(n-k,1))];
s_n = s(k+1:n);

% Right-hand side
b = B*x_b;

% Cost vector
y = randn(m,1);
c = A'*y + s;
```

```
% Scale dx and ds so that  $x + dx \geq 0$  and  $s + ds \geq 0$   
dx = [dxb; dxn];  
dx = dx/norm((dxb./xb),inf);
```

```
% Compute dsn  
dsn = - N'*dy;  
ds = [dsb; dsn];  
no = norm((dsn./sn),inf);  
ds = ds/no;
```

```
% Scale dy  
dy = dy/no;
```

Appendix B

MATLAB Code for Generating a Random SDP Instance

```
function [X,y,S,dX,dy,dS,A,b,C,alpha] = sdpgenerate(m,n,rl,rh,sl,sh,mu)

% Input parameters:
% m : number of primal constraints
% n : size of primal matrix variable
% rl : lower bound on the rank of optimal X
% rh : upper bound on the rank of optimal X
% sl : lower bound on the rank of optimal S
% sh : upper bound on the rank of optimal S
% mu : duality measure of the iterate on the central path
```

```

% Output:
% X,y,S : Point on the central path
% dX,dy,dS : Feasible direction at (X,y,S)
% A,b,C : Data of the SDP instance
% alpha : Range of feasibility for (dX,dy,dS)

% Initialize the random generator seed
randn('state',0);
rand('state',0);

% Generate a random orthogonal matrix
[q,r] = qr(randn(n));

% Random perturbation dX
dX = q * diag(randn(n,1))*q';

% Generate matrices A{i} such that trace(A{i}dX) = 0
dXdX = sum(sum(dX.*dX));
A = cell(1,m);
for k = 1:m
    G = randn(n);
    [Q,R] = qr(G);
    d = randn(n,1);

```

```
A{k} = Q'*diag(d)*Q;  
A{k} = A{k} - (sum(sum(A{k}.*dX))/dXdX)*dX;  
end;  
  
% Generate y  
y = randn(m,1);  
  
% Generate X and S  
G = randn(n);  
[Q,R] = qr(G);  
  
% Determine number of big eigenvalues of X and S  
r = floor(.5*(rl+rh));  
s = floor(.5*(sl+sh));  
  
% Intermediate eigenvalues of X (if necessary)  
im = [ ];  
if n-r-s > 0  
    im = exp(randn(n-r-s,1));  
end;  
  
% r eigenvalues of X and s eigenvalues of S  
u = [1+rand(r,1)];
```

```
l = [1+rand(s,1)];

% Generate X and S
sm = sqrt(mu);
ex = [u;sm*im;mu./l];
X = Q'*diag(ex)*Q;
es = [mu./u;sm./im;l];
S = Q'*diag(es)*Q;

% Generate dy
dy = randn(m,1);

% Generate b, C and dS
b = zeros(m,1);
C = S;
dS = zeros(n);
for k = 1:m
    b(k) = sum(sum(A{k}.*X));
    C = C + y(k)*A{k};
    dS = dS - dy(k)*A{k};
end;

% Compute X(-0.5) and S(-0.5)
```

```
Xih = Q'*diag(1./sqrt(ex))*Q;  
Sih = Q'*diag(1./sqrt(es))*Q;  
  
% Determine alpha  
temp = Xih*dX*Xih;  
temp2 = Sih*dS*Sih;  
mineval = min(eig(temp));  
mineval2 = min(eig(temp2));  
detr = min(mineval,mineval2);  
if detr >= 0  
    alpha = inf;  
else  
    alpha = 1/(-detr);  
end;
```

Bibliography

- [1] I. Adler and R. D. C. Monteiro. Limiting behavior of the affine scaling continuous trajectories for linear programming problems. *Mathematical Programming*, 50:29–51, 1991.
- [2] I. Adler and R. D. C. Monteiro. A geometric view of parametric linear programming. *Algorithmica*, 8:161–176, 1992.
- [3] F. Alizadeh. Semidefinite programming home page, <http://rutcor.rutgers.edu/~alizadeh/sdppage/papers.html>.
- [4] F. Alizadeh, J.-P. A. Haeberly, and M. L. Overton. Complementarity and non-degeneracy in semidefinite programming. *Mathematical Programming*, 77:111–128, 1997.
- [5] 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 Journal on Optimization*, 8:746–768, 1998.
- [6] E. Anderson and P. Nash. *Linear Programming in Infinite Dimensional Spaces*. John Wiley and Sons, 1987.
- [7] G. P. Barker and D. Carlson. Cones of diagonally dominant matrices. *Pacific J. of Math.*, 57:15–32, 1975.
- [8] G. B. Dantzig. Maximization of a linear function of variables subject to linear inequalities. In T. C. Koopmans, editor, *Activity Analysis of Production and Allocation*, pages 339–347. John Wiley, New York, 1951.
- [9] G. B. Dantzig. *Linear Programming and Extensions*. Princeton University Press, Princeton, New Jersey, 1963.

- [10] I. I. Dikin. Iterative solution of problems of linear and quadratic programming. *Doklady Akademii Nauk SSSR*, 174:747–748, 1967. English Translation: *Soviet Mathematics Doklady*, 1967, Volume 8, pp. 674–675.
- [11] I. I. Dikin. On the speed of an iterative process. *Upravlyaemye Sistemy*, 12:54–60, 1974.
- [12] N. Edwards. *Approximation algorithms for the multi-level facility location problem*. PhD thesis, Cornell University, 2001.
- [13] A. V. Fiacco and G. P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. Wiley, New York, 1968. Reprinted by SIAM Publications, 1990.
- [14] K. R. Frisch. The logarithmic potential method of convex programming. Technical Report, University Institute of Economics, Oslo, Norway, 1955.
- [15] J.-L. Goffin, A. Haurie, and J.-P. Vial. Decomposition and nondifferentiable optimization with the projective algorithm. *Management Science*, 38:284–302, 1992.
- [16] J.-L. Goffin and F.S. Mokhtarian. Primal-dual-infeasible Newton approach for the analytic center deep-cutting plane method. *Journal of Optimization Theory and Applications*, 101:35–58, 1999.
- [17] D. Goldfarb and K. Scheinberg. Interior point trajectories in semidefinite programs. *SIAM Journal on Optimization*, 8:871–886, 1998.
- [18] D. Goldfarb and K. Scheinberg. On parametric semidefinite programming. *Applied Numerical Mathematics: Transactions of IMACS*, 29(3):361–377, March 1999.
- [19] A. J. Goldman and A. W. Tucker. Theory of linear programming. In H. W. Kuhn and A. W. Tucker, editors, *Linear Equalities and Related Systems*, pages 53–97. Princeton University Press, Princeton, N. J., 1956.
- [20] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, third edition, 1996.
- [21] J. Gondzio. Warm start of the primal-dual method applied in the cutting-plane scheme. *Mathematical Programming, Series A*, 83:125–143, 1998.

- [22] J. Gondzio and J.-Ph. Vial. Warm start and epsilon-subgradients in the cutting plane scheme for block-angular linear programs. *Computational Optimization and Applications*, 14(1):17–36, 1999.
- [23] J.-P. Goux. Optimization Online home page, <http://www.optimization-online.org/>.
- [24] H. J. Greenberg. The use of the optimal partition in a linear programming solution for post-optimal analysis. *Operations Research Letters*, 15(4):179–185, 1994.
- [25] H. J. Greenberg. Simultaneous primal-dual right-hand-side sensitivity analysis from a strictly complementary solution of a linear program. *SIAM Journal on Optimization*, 10(2):427–442, 2000.
- [26] H. J. Greenberg, A. G. Holder, C. Roos, and T. Terlaky. On the dimension of the set of rim perturbations for optimal partition invariance. *SIAM Journal on Optimization*, 9:207–216, 1999.
- [27] M. Halicka, E. de Klerk, and C. Roos. On the convergence of the central path in semidefinite optimization. Technical report, Faculty ITS, Delft University of Technology, Delft, The Netherlands, 2001.
- [28] C. Helmberg. Semidefinite programming home page, <http://www.zib.de/helmberg/semidef.html>.
- [29] C. Helmberg, F. Rendl, R. J. Vanderbei, and H. Wolkowicz. An interior-point method for semidefinite programming. *SIAM Journal on Optimization*, 6:342–361, 1996.
- [30] A. J. Hoffman. On approximate solutions of systems of linear inequalities. *Journal of Research of the National Bureau of Standards*, 49:263–265, 1952.
- [31] A. G. Holder, J. F. Sturm, and S. Zhang. Marginal and parametric analysis of the central optimal solution. Technical Report 9801/A, Econometric Institute, Erasmus University, Rotterdam, The Netherlands, 1998. To appear in *Information Systems and Operational Research*.
- [32] B. Jansen, J. J. de Jong, C. Roos, and T. Terlaky. Sensitivity analysis in linear programming: Just be careful! *European Journal of Operational Research*, 101:15–28, 1997.

- [33] L. V. Kantorovich. Mathematical methods of organizing and planning production. *Management Science*, 6:366–422, 1959. Russian original by Publication House of the Leningrad State University, 1939.
- [34] N. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.
- [35] L. G. Khachiyan. A polynomial algorithm for linear programming. *Soviet Math. Doklady*, 20:191–194, 1979. Russian original in *Doklady Akademiia Nauk SSSR*, 244:1093–1096.
- [36] L. G. Khachiyan. Polynomial algorithms in linear programming. *USSR Computational Mathematics and Math. Phys.*, 20:53–72, 1980. Russian original in *Zhurnal Vychisditel’noi Matematiki i Matematicheskoi Fiziki*, 20:51–68.
- [37] W.-J. Kim, C.-K. Park, and S. Park. An ϵ -sensitivity analysis in the primal–dual interior point method. *European Journal of Operational Research*, 116(3):629–639, 1999.
- [38] V. Klee and G. J. Minty. How good is the simplex algorithm? In O. Shisha, editor, *Inequalities – III*, pages 159–175. Academic Press, New York, 1972.
- [39] M. Kojima, S. Shindoh, and S. Hara. Interior point methods for the monotone semidefinite linear complementarity problem in symmetric matrices. *SIAM Journal on Optimization*, 7:86–125, 1997.
- [40] Z.-Q. Luo, J. F. Sturm, and S. Zhang. Superlinear convergence of a symmetric primal–dual path–following algorithm for semidefinite programming. *SIAM Journal on Optimization*, 8:59–81, 1998.
- [41] J. E. Mitchell and B. Borchers. Solving real–world linear ordering problems using a primal–dual interior point cutting plane method. *Annals of Operations Research*, 62:253–276, 1996.
- [42] R. D. C. Monteiro. Primal–dual path–following algorithms for semidefinite programming. *SIAM Journal on Optimization*, 7:663–678, 1997.
- [43] R. D. C. Monteiro. Polynomial convergence of primal–dual algorithms for semidefinite programming based on the Monteiro and Zhang family of directions. *SIAM Journal on Optimization*, 8:797–812, 1998.
- [44] R. D. C. Monteiro and S. Mehrotra. A general parametric analysis approach and its implication to sensitivity analysis in interior point methods. *Mathematical Programming*, 72:65–82, 1996.

- [45] R. D. C. Monteiro and P. R. Zanjácomo. A note on the existence of the Alizadeh–Haeberly–Overton direction for semidefinite programming. *Mathematical Programming*, 78:393–396, 1997.
- [46] R. D. C. Monteiro and Y. Zhang. A unified analysis for a class of path-following primal–dual interior–point algorithms for semidefinite programming. *Mathematical Programming*, 81:281–299, 1998.
- [47] Yu. E. Nesterov and A. S. Nemirovskii. *Interior Point Polynomial Methods in Convex Programming*. SIAM Publications, Philadelphia, 1994.
- [48] Yu. E. Nesterov and M. J. Todd. Self-scaled barriers and interior-point methods for convex programming. *Mathematics of Operations Research*, 22:1–42, 1997.
- [49] Yu. E. Nesterov and M. J. Todd. Primal-dual interior-point methods for self-scaled cones. *SIAM Journal on Optimization*, 8:324–362, 1998.
- [50] M. A. Nunez and R. M. Freund. Condition measures and properties of the central trajectory of a linear program. *Mathematical Programming*, 83(1):1–28, 1998.
- [51] M. A. Nunez and R. M. Freund. Condition-measure bounds on the behavior of the central trajectory of a semi-definite program. Technical report, School of Business and Economics, Chapman University, 333 N Glassell St., Orange, CA 92866, August 1999.
- [52] G. Pataki. On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Mathematics of Operations Research*, 23(2):339–358, 1998.
- [53] J. Renegar. A polynomial-time algorithm, based on Newton’s method, for linear programming. *Mathematical Programming*, 40:59–93, 1988.
- [54] J. Renegar. *A Mathematical View of Interior-Point Methods in Convex Optimization*. MPS/SIAM Series on Optimization 3. SIAM Publications, Philadelphia, 2001.
- [55] C. Roos, T. Terlaky, and J.-Ph. Vial. *Theory and Algorithms for Linear Optimization : An Interior Point Approach*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley and Sons, 1997.
- [56] R. Saigal, L. Vandenbergh, and H. Wolkowicz (Editors). *Handbook of Semidefinite Programming*. Kluwer Academic Publishers, 2000.

- [57] M. Shida, S. Shindoh, and M. Kojima. Existence of search directions in interior-point algorithms for the SDP and the monotone SDLCP. *SIAM Journal on Optimization*, 8:387–396, 1998.
- [58] G. W. Stewart. Error and perturbation bounds for subspaces associated with certain eigenvalue problems. *SIAM Review*, 15(4):727–764, 1973.
- [59] J. F. Sturm and S. Zhang. On sensitivity of central solutions in semidefinite programming. *Mathematical Programming*, 90(2):205–227, 2001.
- [60] M. J. Todd. A study of search directions in primal-dual interior-point methods for semidefinite programming. *Optimization Methods and Software*, 11&12:1–46, 1999.
- [61] M. J. Todd. The many facets of linear programming. Technical report, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY 14853–3801, 2000.
- [62] M. J. Todd, K. C. Toh, and R. H. Tütüncü. On the Nesterov-Todd direction in semidefinite programming. *SIAM Journal on Optimization*, 8(3):769–796, 1998.
- [63] K. C. Toh, M. J. Todd, and R. H. Tütüncü. SDPT3 — a Matlab software package for semidefinite programming. *Optimization Methods and Software*, 11:545–581, 1999.
- [64] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38:49–95, 1996.
- [65] S. A. Vavasis and Y. Ye. Identifying an optimal basis in linear programming. *Annals of Operations Research*, 62:565–572, 1996.
- [66] S. J. Wright. Interior-Point Methods Online home page, <http://www-unix.mcs.anl.gov/otc/interiorpoint/>.
- [67] S. J. Wright. *Primal-Dual Interior-Point Methods*. SIAM Publications, Philadelphia, 1997.
- [68] Y. Ye. *Interior Point Algorithms : Theory and Analysis*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley and Sons, 1997.
- [69] E. A. Yildırım. An interior-point perspective on sensitivity analysis in semidefinite programming. Technical Report 1289, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY 14853–3801, 2001.

- [70] E. A. Yildırım and M. J. Todd. An interior-point approach to sensitivity analysis in degenerate linear programs. Technical Report 1268, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY 14853–3801, 2000.
- [71] E. A. Yildırım and M. J. Todd. Sensitivity analysis in linear programming and semidefinite programming using interior-point methods. *Mathematical Programming*, 90(2):229–261, 2001.
- [72] E. A. Yildırım and S. J. Wright. Warm-start strategies in interior-point methods for linear programming. Technical Report 1258, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY 14853–3801, 2000.
- [73] S. Zhang. On the strictly complementary slackness relation in linear programming. In D. Z. Du and J. Sun, editors, *Advances in Optimization and Approximation*, volume 1 of *Nonconvex optimization and its applications*, pages 347–361. Kluwer Academic Publishers, Dordrecht ; Boston, 1994.
- [74] Y. Zhang. On extending some primal–dual interior–point algorithms from linear programming to semidefinite programming. *SIAM Journal on Optimization*, 8:365–386, 1998.