**Mathematical Programming manuscript No.**
(will be inserted by the editor)

E. Alper Yıldırım · Michael J. Todd

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

April 12, 2001

**Abstract.** We analyze perturbations of the right-hand side and the cost parameters in linear programming (LP) and semidefinite programming (SDP). We obtain tight bounds on the perturbations that allow interior-point methods to recover feasible and near-optimal solutions in a single interior-point iteration. For the unique, nondegenerate solution case in LP, we show that the bounds obtained using interior-point methods compare nicely with the bounds arising from using the optimal basis. We also present explicit bounds for SDP using the Monteiro-Zhang family of search directions and specialize them to the AHO, H..K..M, and NT directions.

**Key words.** sensitivity analysis – interior-point methods – linear programming – semidefinite programming

## 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 optimization problem. It is often as important



Both authors: School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY 14853–3801, e-mail: {yildirim|miketodd}@orie.cornell.edu

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 [8], Jansen, de Jong, Roos and Terlaky [12] and S. Zhang [26] 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 [21] develop a parametric analysis of the optimal value from the central optimal solution perspective. Nunez and Freund [20] and Holder, Sturm and S. Zhang [11] study the behavior of the central path under perturbations of the input data.

For the SDP case, Goldfarb and Scheinberg [6] investigate the properties of the optimal value function under perturbations of the input parameters. Sturm and S. Zhang [23] 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 [3], H..K..M [9, 14, 15] and NT [19, 18] 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 [4] and Goffin and S.-Mokhtarian [5] 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 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 [13], 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. Linear Programming

We consider the LP given in the following standard form:

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

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:

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

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 $LPD(b, c)$ 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.)

### *2.1. 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 $\mu$. Each point on the central path satisfies the following system for some $\mu > 0$:

$$\begin{aligned} A^T y + s &= c, \\ Ax \quad &= b, \\ XSe \quad &= \mu e, \end{aligned} \tag{1}$$

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

An interior-point iteration is usually a Newton step for this nonlinear system of equations for some $\mu$, 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 $\mu$ (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:

$$\begin{aligned} A^T \Delta y + \quad \Delta s &= r_d, \\ A\Delta x \qquad\qquad &= r_p, \\ S\Delta x \quad + \quad X\Delta s &= r_{xs}, \end{aligned} \tag{2}$$

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 $\mu e$, 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) 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) has a unique solution given by:

$$\begin{aligned} \Delta y &= (AD^2A^T)^{-1}(r_p + AD^2 r_d - AS^{-1} r_{xs}), \\ \Delta s &= r_d - A^T \Delta y, \\ \Delta x &= S^{-1}(r_{xs} - X\Delta s), \end{aligned} \tag{3}$$

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 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,*

$$\|S^{-1}A^T(AD^2A^T)^{-1}\Delta b\|_\infty \leq 1, \tag{4}$$

*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 $x^Ts$.*

*Proof.* Using the above notation in (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):

$$S\Delta x + X\Delta s = 0. \tag{5}$$

Rewriting this equality component-wise, we have:

$$s_i\Delta x_i + x_i\Delta s_i = 0 \quad \text{so} \quad \frac{\Delta x_i}{x_i} + \frac{\Delta s_i}{s_i} = 0, \qquad i = 1, \ldots, n, \tag{6}$$

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 $s_i + \Delta s_i \geq 0$, $i = 1, \ldots, n$, since the equality constraints will automatically be satisfied if a full Newton step is taken. Combining these inequalities with (6), we have that the new iterate will satisfy nonnegativity for both $x$ and $s$ if and only if $|\frac{\Delta s_i}{s_i}| \leq 1$, $i = 1, \ldots n$. Thus, it is necessary and sufficient to have $\left\|S^{-1}\Delta s\right\|_\infty \leq 1$. But using (3), we have:

$$\Delta y = (AD^2A^T)^{-1}\Delta b \quad \text{and} \quad \Delta s = -A^T\Delta y. \tag{7}$$

Hence

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

and this proves the first part of the proposition.

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

$$(x+\Delta x)^T(s+\Delta s) = x^Ts + x^T\Delta s + s^T\Delta x + \Delta x^T\Delta s. \tag{8}$$

Multiplying (5) by $e^T$ from the left, we have $x^T\Delta s + s^T\Delta x = 0$. From (6), $\Delta x_i$ and $\Delta s_i$ have opposite signs, and so $\Delta x^T\Delta s \leq 0$. Thus, we have:

$$(x+\Delta x)^T(s+\Delta s) \leq x^Ts \tag{9}$$

as claimed.

□

We note that the simple bound

$$\|\Delta b\|_\infty \leq \frac{1}{\|S^{-1}A^T(AD^2A^T)^{-1}\|_\infty} \tag{10}$$

implies that $\Delta b$ satisfies the condition (4); moreover (10) defines the largest $L_\infty$-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.** *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,*

$$\|S^{-1}(I - A^T(AD^2A^T)^{-1}AD^2)\Delta c\|_\infty \leq 1, \tag{11}$$

*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 $x^Ts$.*

*Proof.* Once again using (2), we have $r_p = 0$, $r_d = \Delta c$ and $r_{xs} = 0$. By the argument used in the proof of Proposition 1, it is necessary and sufficient that $\|S^{-1}\Delta s\|_\infty \leq 1$. Note that (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. □

The proposed Newton system to regain feasibility for the new problems uses $r_{xs} = 0$ in (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 $L_\infty$ bounds on a single vector.

Goffin and Sharifi-Mokhtarian [5] 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 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*

$$|\beta| \leq \frac{1}{\|S^{-1}(I - A^T(AD^2A^T)^{-1}AD^2)d_c - S^{-1}A^T(AD^2A^T)^{-1}d_b\|_\infty}, \tag{12}$$

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

*Proof.* Using (3), we have $r_p = \beta d_b$, $r_d = \beta d_c$, and $r_{xs} = 0$ by the hypothesis. Therefore,

$$\|S^{-1}\Delta s\|_\infty = |\beta| \|S^{-1}(I - A^T(AD^2A^T)^{-1}AD^2)d_c - S^{-1}A^T(AD^2A^T)^{-1}d_b\|_\infty, \tag{13}$$

from which the result follows immediately. □

## *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 $\Delta b$ 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:

$$B^{-1}(b + \Delta b) \geq 0 \quad \text{or} \quad B^{-1}\Delta b \geq -B^{-1}b = -x^*_B. \tag{14}$$

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 $\Delta c$ as $\Delta c_B$ and $\Delta c_N$, 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):

$$c_N^T + \Delta c_N^T - c_B^T B^{-1} N - \Delta c_B^T B^{-1} N \geq 0 \quad \text{or} \quad \Delta c_N - N^T B^{-T} \Delta c_B \geq -s_N^*, \qquad (15)$$

where $s_N^*$ and $s_B^*$ partition the dual optimal slack $s^*$. Hence, as long as $\Delta c$ 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 (4) and (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 [12] 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 (4) at various primal strictly feasible points as a function of $\beta$. 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 $\beta$ 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, we will show that we can still say something about the quality of the interior-point bounds even under degeneracy.

### *2.3. Comparison of the Simplex and Interior-Point Approaches*

Recall that Propositions 1 through 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 (14) on $\Delta b$ 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

$$\|(X_B^*)^{-1}B^{-1}\Delta b\|_\infty \leq 1. \tag{16}$$

Similarly, the basis condition (15) on $\Delta c$ 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

$$\|(S_N^*)^{-1}(\Delta c_N - N^T B^{-T}\Delta c_B)\|_\infty \leq 1. \tag{17}$$

Recall that the central path is the set of solutions for positive $\mu$ of the system

$$\begin{aligned} A^T y + s &= c, \\ Ax \qquad &= b, \\ XSe \qquad &= \mu e, \end{aligned} \tag{18}$$

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

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 1, we have:

$$\lim_{\mu \to 0} \dot{x}_B(\mu) = -B^{-1}N(s_N^*)^{-1}. \tag{19}$$

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 $\mu$, taking the limit as $\mu$ tends to 0 and using Theorem 1, we have:

$$\lim_{\mu \to 0} \dot{s}_N(\mu) = N^T B^{-T}(x_B^*)^{-1}. \tag{20}$$

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 $\mu$, the point will be a good approximation to $(x(\mu), y(\mu), s(\mu))$. Consequently, we have the following strictly feasible point:

$$\begin{aligned} x_B &= x_B^* - \mu B^{-1}N(s_N^*)^{-1} && \approx x_B(\mu), \\ x_N &= \mu(s_N^*)^{-1} && \approx x_N(\mu), \\ s_B &= \mu(x_B^*)^{-1} && \approx s_B(\mu), \\ s_N &= s_N^* + \mu N^T B^{-T}(x_B^*)^{-1} && \approx s_N(\mu). \end{aligned} \tag{21}$$

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 $\mu$; moreover, it is easy to check that the duality gap of $(x, y, s)$ is $\mu n$, the same as that of the corresponding point on the central path. Therefore, in the case of a unique nondegenerate solution to $LPP(b, c)$, we have a strictly feasible point to use in our analysis for the interior-point approach.

From Proposition 1, we need to compute the following matrix:

$$S^{-1}A^T(AD^2A^T)^{-1}. \tag{22}$$

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

$$\begin{bmatrix} S_B^{-1} & \\ & S_N^{-1} \end{bmatrix} \begin{bmatrix} B^T \\ N^T \end{bmatrix} \left( \begin{bmatrix} B & N \end{bmatrix} \begin{bmatrix} S_B^{-1} & \\ & S_N^{-1} \end{bmatrix} \begin{bmatrix} X_B & \\ & X_N \end{bmatrix} \begin{bmatrix} B^T \\ N^T \end{bmatrix} \right)^{-1} =$$

$$\begin{bmatrix} S_B^{-1} B^T \\ S_N^{-1} N^T \end{bmatrix} \left( B S_B^{-1} X_B B^T + N S_N^{-1} X_N N^T \right)^{-1}. \tag{23}$$

Next, we substitute the values from (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 [7] 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

$$(U+V)^{-1} = U^{-1} - U^{-1}V(I + U^{-1}V)^{-1}U^{-1}. \tag{24}$$

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 (23). We find that

$$S_B^{-1}B^T U^{-1} = (X_B)^{-1}B^{-1} = (X_B^*)^{-1}B^{-1} + O(\mu), \tag{25}$$

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 (23) is

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

This generates the necessary and sufficient condition

$$\left\| \begin{bmatrix} (X_B^*)^{-1}B^{-1} + O(\mu) \\ O(\mu) \end{bmatrix} \Delta b \right\|_\infty \leq 1, \tag{26}$$

which is asymptotically the same as the symmetrized basis condition (16).

Next, we consider a change in the cost vector $c$. From Proposition 2, we need to evaluate the following:

$$S^{-1}(I - A^T(AD^2A^T)^{-1}AD^2). \tag{27}$$

Permuting both the rows and the columns yields the following:

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

Let us examine each block of this $2 \times 2$ 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 $(S_B)^{-1}B^TU^{-1}$ in (24) and (25), we find that this equals

$$(S_B)^{-1} - (S_B)^{-1} + (X_B)^{-1}B^{-1}V(I + U^{-1}V)^{-1}B^{-T}, \tag{28}$$

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 (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

$$\left\| \begin{bmatrix} O(\mu) & O(\mu) \\ -(S_N^*)^{-1}N^TB^{-T} + O(\mu) & (S_N^*)^{-1} + O(\mu) \end{bmatrix} \begin{bmatrix} \Delta c_B \\ \Delta c_N \end{bmatrix} \right\|_\infty \leq 1, \tag{29}$$

and again this is asymptotically identical to the basis condition (17).

We conclude this section by generalizing our results (26) and (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 (26) and (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 $\mu$; 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.** *Under the assumption of a unique, nondegenerate solution, the expressions (4) and (11) yield the asymptotic results (26) and (29), respectively for all strictly feasible points* $(x, y, s)$ *uniformly in* $\mu$ *where* $\mu := x^T s/n$. *These bounds converge to the symmetrized simplex bounds (16) and (17) as* $\mu$ *approaches zero.*

To prove Theorem 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 1.** *Under the assumption of a unique, nondegenerate solution, let* $(x, y, s)$ *be any strictly feasible solution with duality gap* $\mu n$, *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:*

$$\begin{aligned} &x_B = x_B^* + O(\mu), \quad x_N = O(\mu),\ x_N > 0, \\ &s_B = O(\mu),\ s_B > 0,\ s_N = s_N^* + O(\mu), \end{aligned} \tag{30}$$

*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 $\mu n$, the result for $x$ follows from Hoffman's lemma [10]. A similar argument applies to the dual problem. □

Now we are ready to prove Theorem 2.

*Proof of Theorem 2:* Let $(x, y, s)$ be any strictly feasible solution with duality gap $\mu n$. By Lemma 1, $x$ and $s$ have the form (30). Let us re-examine how we obtained (26) from (22). The only use we made of the form of $(x, y, s)$ was that it satisfied (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 (25) which shows how the $S_B^{-1}$ terms disappear.

Next we reconsider the derivation of (29) from (27). Once again, all we required is (30), and the $S_B^{-1}$ terms vanish; see, e.g., (28), where two such terms cancel.

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

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 (26) and (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.

## 3. Semidefinite Programming

We consider the SDP given in the following standard form:

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

where all $A_i \in S\mathbb{R}^{n\times n}$, $b \in \mathbb{R}^m$, $C \in S\mathbb{R}^{n\times n}$ are given, and $X \in S\mathbb{R}^{n\times n}$. Here $S\mathbb{R}^{n\times n}$ 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 $\mathrm{Trace}\,(P^TQ) = \sum_{ij} P_{ij}Q_{ij}$ 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:

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

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

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

A crucial observation is that Newton's method cannot be directly applied to (1). The reason is that the residual map takes an iterate $(X, y, S) \in S\mathbb{R}^{n\times n} \times \mathbb{R}^m \times S\mathbb{R}^{n\times n}$ to a point in $\mathbb{R}^m \times S\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$ (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 (1) so that the residual lies in $S\mathbb{R}^{n\times n}$. Todd [24] 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, $\mathcal{A} : S\mathbb{R}^{n\times n} \to \mathbb{R}^m$ is defined by

$$
\mathcal{A}U := (A_i \bullet U)_{i=1}^m, \tag{2}
$$

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

$$
\mathcal{A}^* y = \sum_{i=1}^m y_i A_i. \tag{3}
$$

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

$$\begin{aligned} \mathcal{A}^*\tilde{y} \quad + \tilde{S} &= C, \\ \mathcal{A}\tilde{X} \qquad\qquad &= b, \\ \Theta(\tilde{X},\tilde{S}) \qquad &= \Theta(X',S'), \end{aligned} \tag{4}$$

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 $\mu$ 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:

$$\begin{aligned} \mathcal{A}^*\Delta y + \quad \Delta S &= R_d, \\ \mathcal{A}\Delta X \qquad\qquad &= r_p, \\ \mathcal{E}\Delta X \quad + \quad \mathcal{F}\Delta S &= R_{EF}, \end{aligned} \tag{5}$$

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 $\Theta$ 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 [3]:

$$(P \odot Q)K := \frac{1}{2}(PKQ^T + QKP^T), \tag{6}$$

where $P, Q \in I\!R^{n\times n}$ and $K \in SI\!R^{n\times n}$, and we will regard it as an operator from $SI\!R^{n\times n}$ to $SI\!R^{n\times n}$. 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

$$Q \odot P = P \odot Q, \quad (P \odot Q)^* = P^T \odot Q^T, \tag{7}$$

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

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

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 $I\!R^{n\times n}$;

for possibly nonsymmetric matrices $K$, we define it by

$$(P \odot Q)K := \frac{1}{2}(PKQ^T + QK^TP^T). \tag{8}$$

Assume that $\mathcal{E}$ is nonsingular. Then the operator $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$ takes $I\!R^m$ 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 (5) has a unique solution iff the Schur complement matrix is nonsingular, and in this case the solution can be found from

$$\begin{aligned}(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)\Delta y &= r_p - \mathcal{A}\mathcal{E}^{-1}(R_{EF} - \mathcal{F}R_d),\\ \Delta S &= R_d - \mathcal{A}^*\Delta y,\\ \Delta X &= \mathcal{E}^{-1}(R_{EF} - \mathcal{F}\Delta S).\end{aligned} \tag{9}$$

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 [15,27,17]. The AHO direction was suggested by Alizadeh, Haeberly and Overton [3]. The H..K..M direction was independently introduced by Helmberg, Rendl, Vanderbei and Wolkowicz [9]; Kojima, Shindoh and Hara [14]; and Monteiro [15]. Finally, Nesterov and Todd [19,18] 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 [22,25] or if $(X, y, S)$ lies in a suitable neighborhood of the central path [16]. 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 $SI\!R^{n\times n}$ to $SI\!R^{n\times n}$ is positive definite if $U \bullet \mathcal{G}U > 0$ for every nonzero $U \in SI\!R^{n\times n}$; it is self-adjoint if $U \bullet \mathcal{G}V = V \bullet \mathcal{G}U$ for every $U, V \in SI\!R^{n\times n}$.) 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 3.2–3.4 on a first reading.

### *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:

$$\mathcal{E} = S \odot M, \qquad \mathcal{F} = MX \odot I, \tag{10}$$

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 (7), the adjoint operators are given by

$$\mathcal{E}^* = S \odot M, \qquad \mathcal{F}^* = XM \odot I. \tag{11}$$

Under the assumption (10), $\mathcal{E}$ 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 4.** *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 (5) be given by (10). 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* $\Delta b \in I\!R^m$*, and suppose a Newton step for the system (4) 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,* $\Delta b$ *satisfies the following inequalities:*

$$\lambda_{\min}\left(X^{-\frac{1}{2}}\left(\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*\left[(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\Delta b\right]\right)X^{-\frac{1}{2}}\right) \geq -1, \tag{12}$$

$$\lambda_{\max}\left(S^{-\frac{1}{2}}\left(\mathcal{A}^*\left[(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\Delta b\right]\right)S^{-\frac{1}{2}}\right) \leq 1. \tag{13}$$

*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 (9), the Newton step $(\Delta X, \Delta y, \Delta S)$ is given by:

$$\begin{aligned} \Delta y &= (\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\Delta b, \\ \Delta S &= -\mathcal{A}^*\left[(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\Delta b\right], \\ \Delta X &= \mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*\left[(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\Delta b\right]. \end{aligned} \tag{14}$$

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,

$$X + \Delta X \succeq 0 \quad \text{holds iff} \quad I + X^{-\frac{1}{2}}\Delta X\, X^{-\frac{1}{2}} \succeq 0. \tag{15}$$

(15) 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 (14) with (15), we have exactly (12) as a necessary and sufficient condition for the new $X$ iterate to be feasible. Similarly,

$$S + \Delta S \succeq 0 \quad \text{holds iff} \quad I + S^{-\frac{1}{2}}\Delta S\, S^{-\frac{1}{2}} \succeq 0; \tag{16}$$

combining (14) with (16), and using the same argument, we have exactly (13) 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

$$(X + \Delta X) \bullet (S + \Delta S) = X \bullet S + X \bullet \Delta S + S \bullet \Delta X + \Delta X \bullet \Delta S. \tag{17}$$

By our hypothesis, $\mathcal{E}$ and $\mathcal{F}$ are given by (10). Therefore, if we use (11), it is easy to verify that

$$\mathcal{E}^* M^{-1} = S, \qquad \mathcal{F}^* M^{-1} = X. \tag{18}$$

Then, using (18), we have:

$$\Delta X \bullet S = \Delta X \bullet \mathcal{E}^* M^{-1} = \mathcal{E}\Delta X \bullet M^{-1}. \tag{19}$$

Similarly,

$$\Delta S \bullet X = \Delta S \bullet \mathcal{F}^* M^{-1} = \mathcal{F}\Delta S \bullet M^{-1}. \tag{20}$$

However, (5) and our hypothesis imply $\mathcal{E}\Delta X + \mathcal{F}\Delta S = 0$. Combining this with (19) and (20), we obtain

$$\Delta X \bullet S + X \bullet \Delta S = 0. \tag{21}$$

Finally,

$$\Delta X \bullet \Delta S = -\Delta S \bullet \mathcal{E}^{-1}\mathcal{F}\Delta S \leq 0, \tag{22}$$

since $\mathcal{E}^{-1}\mathcal{F}$ is positive definite. Hence, (17), (21), and (22) imply:

$$(X + \Delta X) \bullet (S + \Delta S) \leq X \bullet S. \tag{23}$$

This completes the proof. □

Next, we consider perturbations of the cost matrix $C$.

**Proposition 5.** *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 (5) be given by (10). 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 $\Delta C \in S\mathbb{R}^{n\times n}$, and suppose a Newton step for the system (4) 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, $\Delta C$ satisfies the following inequalities:*

$$\lambda_{\max}\left(X^{-\frac{1}{2}}\left[\mathcal{E}^{-1}\mathcal{F}\Delta C - \mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\Delta C\right]X^{-\frac{1}{2}}\right) \leq 1, \tag{24}$$

$$\lambda_{\min}\left(S^{-\frac{1}{2}}\left[\Delta C - \mathcal{A}^*(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\Delta C\right]S^{-\frac{1}{2}}\right) \geq -1. \tag{25}$$

*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 (9), 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:

$$\begin{aligned}
\Delta y &= (\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\Delta C,\\
\Delta S &= \Delta C - \mathcal{A}^*(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\Delta C,\\
\Delta X &= -\mathcal{E}^{-1}\mathcal{F}\Delta C + \mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\Delta C.
\end{aligned} \tag{26}$$

Then, proceeding as in the proof of Proposition 4, 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 (26), we obtain exactly the bounds (24) and (25) we seek. Essentially the same argument as in Proposition 4 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. □

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

**Proposition 6.** *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 (5) be given by (10). 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 $D_C \in S\mathbb{R}^{n \times n}$. Suppose a Newton step for the system (4) 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, $\beta$ satisfies the following:*

$$|\beta| \le \min\{a, b\}, \tag{27}$$

*where $a$ is the reciprocal of*

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

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

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

*(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 4 and 5, the result follows simply by observing that $r_p = \beta d_b$, $R_D = \beta D_C$ and $R_{XS} = 0$ in (9), and imposing the conditions $\lambda_{\min}(S^{-\frac{1}{2}} \Delta S \, S^{-\frac{1}{2}}) \ge -1$ and $\lambda_{\min}(X^{-\frac{1}{2}} \Delta X \, X^{-\frac{1}{2}}) \ge -1$. □

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

$$(\widehat{SDP}) \qquad \min_{\hat{X}} \ \hat{C} \bullet \hat{X}$$
$$\hat{\mathcal{A}}\hat{X} = b,$$
$$\hat{X} \succeq 0,$$

where $\hat{C} := P^{-T}CP^{-1}$, and $\hat{\mathcal{A}}$ and $\hat{\mathcal{A}}^*$ are defined from $\{\hat{A}_i := P^{-T}A_iP^{-1}\}$ as in (2) and (3). The dual of this problem is

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

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 [24]. 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 [24]).

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:

$$u = (\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)e_i = \mathcal{A}\mathcal{E}^{-1}\mathcal{F}A_i, \tag{28}$$

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 (10), we have:

$$\frac{1}{2}(SKM + MKS) = \frac{1}{2}(MXA_i + A_iXM). \tag{29}$$

Then $u$ in (28) is given by

$$u = \begin{bmatrix} A_1 \bullet K \\ \vdots \\ A_m \bullet K \end{bmatrix}. \tag{30}$$

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:

$$u' = (\hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{\mathcal{A}}^*)e_i = \hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{A}_i. \tag{31}$$

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:

$$\begin{gathered}
\frac{1}{2}(\hat{S}\hat{K}\hat{M} + \hat{M}\hat{K}\hat{S}) = \frac{1}{2}(\hat{M}\hat{X}\hat{A}_i + \hat{A}_i\hat{X}\hat{M}) \quad \text{or} \\
\frac{1}{2}(P^{-T}SP^{-1}\hat{K}P^{-T}MP^{-1} + P^{-T}MP^{-1}\hat{K}P^{-T}SP^{-1}) = \\
\frac{1}{2}(P^{-T}MP^{-1}PXP^{T}P^{-T}A_iP^{-1} + P^{-T}A_iP^{-1}PXP^{T}P^{-T}MP^{-1}).
\end{gathered} \tag{32}$$

Multiplying (32) by $P^T$ from the left and $P$ from the right, we get:

$$\frac{1}{2}(SP^{-1}\hat{K}P^{-T}M + MP^{-1}\hat{K}P^{-T}S) = \frac{1}{2}(MXA_i + A_iXM). \tag{33}$$

Comparing (29) with (33), 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,

$$u' = \begin{bmatrix} \hat{A}_1 \bullet \hat{K} \\ \vdots \\ \hat{A}_m \bullet \hat{K} \end{bmatrix} = \begin{bmatrix} \operatorname{Trace}(P^{-T}A_1P^{-1}PKP^T) \\ \vdots \\ \operatorname{Trace}(P^{-T}A_1P^{-1}PKP^T) \end{bmatrix} = \begin{bmatrix} A_1 \bullet K \\ \vdots \\ A_m \bullet K \end{bmatrix}. \tag{34}$$

From (30) and (34), 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.

### *3.2. The AHO Direction*

The AHO direction [3] is the Newton step for the following symmetrization of the third equation in (4):

$$\Theta(\tilde{X}, \tilde{S}) := \frac{1}{2}(\tilde{X}\tilde{S} + \tilde{S}\tilde{X}) = \frac{1}{2}(X'S' + S'X'). \tag{35}$$

It corresponds to taking

$$\mathcal{E} = S \odot I, \quad \mathcal{F} = X \odot I, \quad R_{EF} = \frac{1}{2}(X'S' + S'X') - \frac{1}{2}(XS + SX). \tag{36}$$

Therefore, $M = I$ for the AHO direction. Recall from Section 3.1 that we need the operator $\mathcal{E}^{-1}$ for our analysis. For the AHO direction, $\mathcal{E}$ is given by (36), 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:

$$\begin{aligned} \hat{X} &= Q^{-1}XQ^{-T} = Q^T XQ, \\ \hat{S} &= Q^T SQ \qquad\;\; = D, \\ \hat{A}_i &= Q^T A_i Q. \end{aligned} \tag{37}$$

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

$$\hat{\mathcal{E}}U = R \quad \text{iff} \quad DU + UD = 2R \quad \text{iff} \quad U_{ij} = \frac{2R_{ij}}{d_i + d_j}, \tag{38}$$

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:

$$Ne_i = (\hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{\mathcal{A}}^*)e_i = \hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{A}_i = \hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\left[\tfrac{1}{2}(\hat{A}_i\hat{X} + \hat{X}\hat{A}_i)\right] = \hat{\mathcal{A}}\hat{K}_i, \tag{39}$$

where

$$(\hat{K}_i)_{kl} = \frac{(\hat{A}_i\hat{X} + \hat{X}\hat{A}_i)_{kl}}{d_k + d_l} \tag{40}$$

from (38). Let us also write $K_i$ for $Q\hat{K}_iQ^T$. Therefore,

$$Ne_i = \begin{bmatrix} \hat{A}_1 \bullet \hat{K}_i \\ \vdots \\ \hat{A}_m \bullet \hat{K}_i \end{bmatrix} = \begin{bmatrix} A_1 \bullet K_i \\ \vdots \\ A_m \bullet K_i \end{bmatrix}. \tag{41}$$

(41) implies that the Schur complement matrix $N$ is not symmetric in general (see also Proposition 6.4 in [24]).

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 4 implies that the following bounds on $\Delta b$ are necessary and sufficient:

$$\begin{aligned} \lambda_{\min}\left(\hat{X}^{-\frac{1}{2}}\left[\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{\mathcal{A}}^*(N^{-1}\Delta b)\right]\hat{X}^{-\frac{1}{2}}\right) &\geq -1, \\ \lambda_{\max}\left(\hat{S}^{-\frac{1}{2}}\left[\hat{\mathcal{A}}^*(N^{-1}\Delta b)\right]\hat{S}^{-\frac{1}{2}}\right) &\leq 1. \end{aligned} \tag{42}$$

The first inequality in (42) yields the following:

$$\lambda_{\min}\left(\hat{X}^{-\frac{1}{2}}\left(\hat{\mathcal{E}}^{-1}\left[\frac{1}{2}\sum_{i=1}^{m}(N^{-1}\Delta b)_i\left(\hat{A}_i\hat{X} + \hat{X}\hat{A}_i\right)\right]\right)\hat{X}^{-\frac{1}{2}}\right) =$$

$$\lambda_{\min}\left(\sum_{i=1}^{m}(N^{-1}\Delta b)_i\left(\hat{X}^{-\frac{1}{2}}\hat{K}_i\hat{X}^{-\frac{1}{2}}\right)\right) \geq -1, \tag{43}$$

where $\hat{K}_i$ is as in (40) 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 $X^{-\frac{1}{2}}K_iX^{-\frac{1}{2}}$, since the two are related by an orthogonal similarity. Similarly, the second inequality in (42) yields the following:

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

$$\lambda_{\max}\left(\sum_{i=1}^{m}(N^{-1}\Delta b)_i\left(D^{-\frac{1}{2}}\hat{A}_iD^{-\frac{1}{2}}\right)\right) \leq 1. \tag{44}$$

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

**Proposition 7.** *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 $\Delta b$ satisfies (43) and (44). 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 5 implies that the following bounds on $\widehat{\Delta C}$ are necessary and sufficient:

$$\begin{aligned} \lambda_{\max}\left(\hat{X}^{-\frac{1}{2}}\left[\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\widehat{\Delta C} - \hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{\mathcal{A}}^* N^{-1}\hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\widehat{\Delta C}\right]\hat{X}^{-\frac{1}{2}}\right) &\le 1, \\ \lambda_{\min}\left(\hat{S}^{-\frac{1}{2}}\left[\widehat{\Delta C} - \hat{\mathcal{A}}^* N^{-1}\hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\widehat{\Delta C}\right]\hat{S}^{-\frac{1}{2}}\right) &\ge -1. \end{aligned} \tag{45}$$

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 (38), 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 (45) simplifies to

$$\begin{aligned} &\lambda_{\max}\left(\hat{X}^{-\frac{1}{2}}\left(\hat{L} - \hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{\mathcal{A}}^*(N^{-1}v)\right)\hat{X}^{-\frac{1}{2}}\right) = \\ &\lambda_{\max}\left(\hat{X}^{-\frac{1}{2}}\left(\hat{L} - \hat{\mathcal{E}}^{-1}\left[\frac{1}{2}\sum_{i=1}^m (N^{-1}v)_i\left(\hat{A}_i\hat{X} + \hat{X}\hat{A}_i\right)\right]\right)\hat{X}^{-\frac{1}{2}}\right) = \\ &\lambda_{\max}\left(\hat{X}^{-\frac{1}{2}}\left(\hat{L} - \sum_{i=1}^m (N^{-1}v)_i\hat{K}_i\right)\hat{X}^{-\frac{1}{2}}\right) \le 1, \end{aligned} \tag{46}$$

where $\hat{K}_i$ is again given by (40) and

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

Similarly, the second inequality in (45) yields:

$$\begin{aligned} &\lambda_{\min}\left(D^{-\frac{1}{2}}\left[\widehat{\Delta C} - \hat{\mathcal{A}}^*\left((N^{-1})\hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\widehat{\Delta C}\right)\right]D^{-\frac{1}{2}}\right) = \\ &\lambda_{\min}\left(D^{-\frac{1}{2}}\left[\widehat{\Delta C} - \sum_{i=1}^m (N^{-1}v)_i\hat{A}_i\right]D^{-\frac{1}{2}}\right) \ge -1. \end{aligned} \tag{47}$$

(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 8.** *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 $\Delta C \in S\mathbb{R}^{n\times n}$. 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 (46) and (47). Moreover, the duality gap of the new iterate will be at most $X \bullet S$ if $\mathcal{E}^{-1}\mathcal{F}$ is positive definite.*

□

### *3.3. The H..K..M Direction*

The H..K..M direction [9,14,15] is the Newton step for the following symmetrization of the third equation in (4):

$$\Theta(\tilde{X}, \tilde{S}) := \frac{1}{2}(S\tilde{X}\tilde{S} + \tilde{S}\tilde{X}S) = \frac{1}{2}(SX'S' + S'X'S). \tag{48}$$

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

$$\mathcal{E} = S \odot S, \quad \mathcal{F} = SX \odot I, \quad R_{EF} = \frac{1}{2}(SX'S' + S'X'S) - SXS. \tag{49}$$

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

$$\mathcal{E} = I \odot I, \quad \mathcal{F} = X \odot S^{-1}, \quad R_{EF} = \frac{1}{2}(X'S'S^{-1} + S^{-1}S'X') - X. \tag{50}$$

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:

$$\begin{aligned} \hat{X} &= S^{\frac{1}{2}} X S^{\frac{1}{2}}, \\ \hat{S} &= I, \\ \hat{A}_i &= S^{-\frac{1}{2}} A_i S^{-\frac{1}{2}}, \\ \hat{C} &= S^{-\frac{1}{2}} C S^{-\frac{1}{2}}. \end{aligned} \tag{51}$$

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:

$$Ne_i = (\hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{\mathcal{A}}^*)e_i = \hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{A}_i = \hat{\mathcal{A}}\left[\frac{1}{2}(\hat{A}_i\hat{X} + \hat{X}\hat{A}_i)\right]. \tag{52}$$

Therefore,

$$Ne_i = \frac{1}{2}\begin{bmatrix} \hat{A}_1 \bullet (\hat{A}_i\hat{X} + \hat{X}\hat{A}_i) \\ \vdots \\ \hat{A}_m \bullet (\hat{A}_i\hat{X} + \hat{X}\hat{A}_i) \end{bmatrix} = \begin{bmatrix} \mathrm{Trace}\,(\hat{A}_1\hat{X}\hat{A}_i) \\ \vdots \\ \mathrm{Trace}\,(\hat{A}_m\hat{X}\hat{A}_i) \end{bmatrix}. \tag{53}$$

In the above derivation, we used the obvious facts that $\mathrm{Trace}\,(A) = \mathrm{Trace}\,(A^T)$ and $\mathrm{Trace}\,(PK) = \mathrm{Trace}\,(KP)$, for any square matrices $A$, $P$, and $K$. (53) implies that the Schur complement matrix $N$ is always symmetric (see also Proposition 6.4 in [24]).

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 [24]).

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 4 implies that the following bounds on $\Delta b$ are necessary and sufficient:

$$\begin{aligned} \lambda_{\min}\left(\hat{X}^{-\frac{1}{2}}\left[\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{\mathcal{A}}^*(N^{-1}\Delta b)\right]\hat{X}^{-\frac{1}{2}}\right) &\geq -1, \\ \lambda_{\max}\left(\hat{S}^{-\frac{1}{2}}\left[\hat{\mathcal{A}}^*(N^{-1}\Delta b)\right]\hat{S}^{-\frac{1}{2}}\right) &\leq 1. \end{aligned} \tag{54}$$

The first inequality in (54) yields the following:

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

$$\lambda_{\min}\left(\frac{1}{2}\sum_{i=1}^m (N^{-1}\Delta b)_i\left[\hat{X}^{-\frac{1}{2}}\hat{A}_i\hat{X}^{\frac{1}{2}} + \hat{X}^{\frac{1}{2}}\hat{A}_i\hat{X}^{-\frac{1}{2}}\right]\right) \geq -1. \tag{55}$$

Similarly, the second inequality in (54) yields the following:

$$\lambda_{\max}\left(\sum_{i=1}^m (N^{-1}\Delta b)_i\hat{A}_i\right) \leq 1. \tag{56}$$

Note that (56) bounds the maximum eigenvalue of $E := \sum_{i=1}^m (N^{-1}\Delta b)_i\hat{A}_i$, while (55) bounds the minimum eigenvalue of $(U^{-1}EU + UEU^{-1})/2$ for $U := \hat{X}^{\frac{1}{2}}$. By Lemma 3.3 of Monteiro

[15], the former eigenvalue is always bounded by the maximum eigenvalue of $(U^{-1}EU + UEU^{-1})/2$. Thus, we have

**Proposition 9.** *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 $\Delta b \in I\!R^m$ 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 $\Delta b$ satisfies (55) and (56). 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 $\Delta b$ 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 5 implies that the following bounds on $\widehat{\Delta C}$ are necessary and sufficient:

$$\begin{aligned} \lambda_{\max}\left( \hat{X}^{-\frac{1}{2}} \left[ \hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\widehat{\Delta C} - \hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{\mathcal{A}}^* N^{-1}\hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\widehat{\Delta C} \right] \hat{X}^{-\frac{1}{2}} \right) &\leq 1, \\ \lambda_{\min}\left( \hat{S}^{-\frac{1}{2}} \left[ \widehat{\Delta C} - \hat{\mathcal{A}}^* N^{-1}\hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\widehat{\Delta C} \right] \hat{S}^{-\frac{1}{2}} \right) &\geq -1. \end{aligned} \tag{57}$$

The first inequality in (57) yields the following:

$$\begin{aligned} &\lambda_{\max}\left( \hat{X}^{-\frac{1}{2}} \left[ \frac{1}{2}(\widehat{\Delta C}\hat{X} + \hat{X}\widehat{\Delta C}) - \hat{\mathcal{F}}\hat{\mathcal{A}}^* \left( N^{-1}\hat{\mathcal{A}} \left[ \frac{1}{2}(\widehat{\Delta C}\hat{X} + \hat{X}\widehat{\Delta C}) \right] \right) \right] \hat{X}^{-\frac{1}{2}} \right) = \\ &\lambda_{\max}\left( \hat{X}^{-\frac{1}{2}} \left[ \frac{1}{2}(\widehat{\Delta C}\hat{X} + \hat{X}\widehat{\Delta C}) - \frac{1}{2}\sum_{i=1}^{m}(N^{-1}v)_i(\hat{A}_i\hat{X} + \hat{X}\hat{A}_i) \right] \hat{X}^{-\frac{1}{2}} \right) = \\ &\lambda_{\max}\left( \frac{1}{2}\left( \hat{X}^{-\frac{1}{2}}\widehat{\Delta C}\hat{X}^{\frac{1}{2}} + \hat{X}^{\frac{1}{2}}\widehat{\Delta C}\hat{X}^{-\frac{1}{2}} \right) - \frac{1}{2}\sum_{i=1}^{m}(N^{-1}v)_i \left( \hat{X}^{-\frac{1}{2}}\hat{A}_i\hat{X}^{\frac{1}{2}} + \hat{X}^{\frac{1}{2}}\hat{A}_i\hat{X}^{-\frac{1}{2}} \right) \right) \leq 1, \end{aligned} \tag{58}$$

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

Similarly, the second inequality in (57) yields:

$$\lambda_{\min}\left( \widehat{\Delta C} - \hat{\mathcal{A}}^* \left( N^{-1}\hat{\mathcal{A}}\hat{\mathcal{F}}\widehat{\Delta C} \right) \right) = \lambda_{\min}\left( \widehat{\Delta C} - \sum_{i=1}^{m}(N^{-1}v)_i\hat{A}_i \right) \geq -1, \tag{59}$$

where $v$ is the same as above. Note that (59) bounds the minimum eigenvalue of $E := \widehat{\Delta C} - \sum_{i=1}^m (N^{-1}v)_i \hat{A}_i$, while (58) bounds the maximum eigenvalue of $(U^{-1}EU + UEU^{-1})/2$ for $U := \hat{X}^{\frac{1}{2}}$. Again, the results of Monteiro [15] show that the former is at most the minimum eigenvalue of $(U^{-1}EU + UEU^{-1})/2$. Thus, we have shown

**Proposition 10.** *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 $\Delta C \in S\mathbb{R}^{n\times n}$ 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 $X'S' = XS$. 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 $\widehat{\Delta C} = S^{-\frac{1}{2}} \Delta C S^{-\frac{1}{2}}$ satisfies (58) and (59). 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 $X \bullet S$.*

□

### 3.4. The NT Direction

The NT direction [19,18] is the Newton step for the following symmetrization of the third equation in (4):

$$\Theta(\tilde{X}, \tilde{S}) := \frac{1}{2}(W^{-1}\tilde{X}\tilde{S} + \tilde{S}\tilde{X}W^{-1}) = \frac{1}{2}(W^{-1}X'S' + S'X'W^{-1}), \tag{60}$$

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

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

$$R_{EF} = \frac{1}{2}\left((W^{-1}X'S' + S'X'W^{-1}) - (W^{-1}XS + SXW^{-1})\right). \tag{61}$$

Therefore, $M = W^{-1}$ for the NT direction. It has been shown [25] 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:

$$\mathcal{E} = I \odot I, \quad \mathcal{F} = W \odot W, \quad R_{EF} = \nu S^{-1} - X. \tag{62}$$

We claim that the representation (62) can be generalized to the case when $X'$ and $S'$ are arbitrary matrices if $R_{EF}$ is appropriately chosen. First of all, note that (61) implies that the third equation of (5) is given by

$$(S\Delta X + \Delta S X)W^{-1} + W^{-1}(\Delta X S + X\Delta S) = W^{-1}(X'S' - XS) + (S'X' - SX)W^{-1}. \quad (63)$$

Postmultiplying (63) by $W$, we obtain

$$S\Delta X + \Delta S X + W^{-1}(\Delta X S + X\Delta S)W = W^{-1}(X'S' - XS)W + S'X' - SX. \quad (64)$$

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

$$\Delta X + W\Delta S W = R_{EF}, \quad (65)$$

which is related to (62) is equivalent to (64) for arbitrary $X'$ and $S'$. Proceeding as in [25], we see that (65) is equivalent to each of the following two equations:

$$W^{-1}\Delta X S W + \Delta S X = W^{-1}R_{EF}SW, \quad (66)$$

$$S\Delta X + SW\Delta S W = SR_{EF}. \quad (67)$$

The first equality follows from premultiplying (65) by $W^{-1}$ and postmultiplying by $W^{-1}X$, together with $SW = W^{-1}X$, and the second equality is a consequence of premultiplying (65) by $S$. Adding up (66) and (67), we obtain the same expression on the left hand side of (64). Hence, it follows that $R_{EF}$ should satisfy

$$W^{-1}R_{EF}SW + SR_{EF} = W^{-1}(X'S' - XS)W + S'X' - SX. \quad (68)$$

Postmultiplying (68) by $W^{-1}$, we obtain

$$W^{-1}R_{EF}S + SR_{EF}W^{-1} = W^{-1}(X'S' - XS) + (S'X' - SX)W^{-1}. \quad (69)$$

Using the notations in (6) and (8), (69) is equivalent to

$$(W^{-1} \odot S)R_{EF} = (W^{-1} \odot I)[X'S' - XS]. \quad (70)$$

On the left, $W^{-1} \odot S$ is viewed as an operator from $S\mathbb{R}^{n\times n}$ to itself, while on the right $W^{-1} \odot I$ takes $\mathbb{R}^{n\times n}$ to $S\mathbb{R}^{n\times n}$. However, note that $W^{-1} \odot S$ is positive definite, and therefore (70) has a unique solution $R_{EF}$. This shows that the solution to (65) will also satisfy (63) if $R_{EF}$ is given by (70). However, the solutions to the two Newton systems (5) where the third equations are given by (63) and (65) respectively exist and are unique, and hence must agree. Therefore, we conclude that an alternative representation of the NT direction is given by

$$\mathcal{E} = I \odot I, \quad \mathcal{F} = W \odot W, \quad R_{EF}, \tag{71}$$

where $R_{EF}$ is defined as the solution to (70). Observe that computation of $R_{EF}$ involves solving a Lyapunov system. However, for our purposes, the right hand side of (70) 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:

$$\begin{aligned} \hat{X} &= W^{-\frac{1}{2}} X W^{-\frac{1}{2}}, \\ \hat{S} &= W^{\frac{1}{2}} S W^{\frac{1}{2}}, \\ \hat{A}_i &= W^{\frac{1}{2}} A_i W^{\frac{1}{2}}, \\ \hat{C} &= W^{\frac{1}{2}} C W^{\frac{1}{2}}. \end{aligned} \tag{72}$$

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:

$$\begin{aligned} N e_i &= (\hat{\mathcal{A}} \hat{\mathcal{E}}^{-1} \hat{\mathcal{F}} \hat{\mathcal{A}}^*) e_i = \hat{\mathcal{A}} \hat{A}_i \\ &= \begin{bmatrix} \hat{A}_1 \bullet \hat{A}_i \\ \vdots \\ \hat{A}_m \bullet \hat{A}_i \end{bmatrix}. \end{aligned} \tag{73}$$

Thus the Schur complement matrix $N$ is always symmetric (see also Proposition 6.4 in [24]).

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 [24]).

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 4 implies that the following bounds on $\Delta b$ are necessary and sufficient:

$$\begin{aligned} \lambda_{\min}\left(\hat{X}^{-\frac{1}{2}} \left[\hat{\mathcal{E}}^{-1} \hat{\mathcal{F}} \hat{\mathcal{A}}^*(N^{-1} \Delta b)\right] \hat{X}^{-\frac{1}{2}}\right) &\geq -1, \\ \lambda_{\max}\left(\hat{S}^{-\frac{1}{2}} \left[\hat{\mathcal{A}}^*(N^{-1} \Delta b)\right] \hat{S}^{-\frac{1}{2}}\right) &\leq 1. \end{aligned} \tag{74}$$

The first inequality in (74) yields the following:

$$\lambda_{\min}\left(\sum_{i=1}^{m}(N^{-1}\Delta b)_i\hat{X}^{-\frac{1}{2}}\hat{A}_i\hat{X}^{-\frac{1}{2}}\right)\geq -1. \tag{75}$$

Similarly, since $\hat{X}=\hat{S}$, the second inequality in (74) 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 $L_\infty$-norm of a vector for LP. We state this nice result about the NT direction as:

**Proposition 11.** *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 $\Delta b$ 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 5 implies that the following bounds on $\widehat{\Delta C}$ are necessary and sufficient:

$$\begin{aligned}\lambda_{\max}\left(\hat{X}^{-\frac{1}{2}}\left[\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\widehat{\Delta C}-\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\hat{\mathcal{A}}^*N^{-1}\hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\widehat{\Delta C}\right]\hat{X}^{-\frac{1}{2}}\right)\leq 1,\\ \lambda_{\min}\left(\hat{S}^{-\frac{1}{2}}\left[\widehat{\Delta C}-\hat{\mathcal{A}}^*N^{-1}\hat{\mathcal{A}}\hat{\mathcal{E}}^{-1}\hat{\mathcal{F}}\widehat{\Delta C}\right]\hat{S}^{-\frac{1}{2}}\right)\geq -1.\end{aligned} \tag{76}$$

The first inequality in (76) yields the following:

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

$$\lambda_{\max}\left(\hat{X}^{-\frac{1}{2}}\widehat{\Delta C}\hat{X}^{-\frac{1}{2}}-\sum_{i=1}^{m}(N^{-1}v)_i\hat{X}^{-\frac{1}{2}}\hat{A}_i\hat{X}^{-\frac{1}{2}}\right)\leq 1, \tag{77}$$

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

Similarly, the second inequality in (76) 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 12.** *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 $\Delta C \in S\mathbb{R}^{n \times n}$. 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 $X \bullet S$.*

□

## 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 (4) and (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 (10) 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 [14,15], 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 $\Sigma$ 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

$$\begin{aligned} Ne_i &= \mathcal{A}\mathcal{E}^{-1}\mathcal{F}A_i = \mathcal{A}S^{-1}XA_i \Longrightarrow \\ Ne_i &= \begin{bmatrix} \operatorname{Trace}(A_1 S^{-1} X A_i) \\ \vdots \\ \operatorname{Trace}(A_m S^{-1} X A_i) \end{bmatrix}. \end{aligned} \tag{78}$$

However, (78) 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 (12) and (13) arising from Proposition 4 can be simplified using the fact that all operations yield diagonal matrices and that diagonal matrices commute. In particular, (12) is equivalent to

$$\begin{aligned} &\lambda_{\min}\left(X^{-\frac{1}{2}}\left(\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*\left[(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\Delta b\right]\right)X^{-\frac{1}{2}}\right) = \\ &\qquad \lambda_{\min}\left(S^{-1}\mathcal{A}^*\left[(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\Delta b\right]\right) \geq -1, \end{aligned} \tag{79}$$

which can be combined with (13) to yield the following norm bound:

$$\begin{aligned} &\left\|S^{-1}\mathcal{A}^*\left[(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\Delta b\right]\right\|_2 = \\ &\left\|\sum_{i=1}^m \left((AD^2A^T)^{-1}\Delta b\right)_i A_i S^{-1}\right\|_2 = \\ &\left\|S^{-1}A^T(AD^2A^T)^{-1}\Delta b\right\|_\infty \leq 1, \end{aligned} \tag{80}$$

where the last equality follows from the fact that the $L_2$ operator norm of a diagonal matrix is the same as the $L_\infty$ norm of the vector of its diagonal entries. Therefore, the SDP bounds for $\Delta b$ reduce exactly to the bound given in Proposition 1.

We now consider perturbations of $C$. In a similar manner, the bounds (24) and (25) arising from Proposition 5 can be combined into a single norm bound given by

$$\left\|S^{-1}\left(\Delta C - \mathcal{A}^*(\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)^{-1}\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\Delta C\right)\right\|_2 \leq 1. \tag{81}$$

Using our previous observations, $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\Delta C = AD^2\Delta c$, where $\Delta c$ is the vector obtained from the diagonal entries of $\Delta C$. Therefore, (81) can be rewritten as

$$\begin{aligned} &\left\|S^{-1}\Delta C - \sum_{i=1}^m \left((AD^2A^T)^{-1}AD^2\Delta c\right)_i A_i S^{-1}\right\|_2 = \\ &\qquad \left\|S^{-1}(I - A^T(AD^2A^T)^{-1}AD^2)\Delta c\right\|_\infty \leq 1, \end{aligned} \tag{82}$$

which again is the same as the bound given in Proposition 2.

Therefore, we have proved the following proposition:

**Proposition 13.** *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 $c$.*

□

## 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, 12] 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. 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 [6], 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.

*Acknowledgements.* The research of the first author was partially performed at and supported by Argonne National Laboratory while the author was visiting the Mathematics and Computer Science Division, June – July 1999. The second author was supported in part by NSF through grant DMS-9805602 and ONR through grant N00014-96-1-0050

## References

1. Adler, I., Monteiro, R.D.C. (1991): Limiting behavior of the affine scaling continuous trajectories for linear programming problems. Math. Program. **50**, 29–51
2. Adler, I., Monteiro, R.D.C. (1992): A geometric view of parametric linear programming. Algorithmica **8**, 161–176
3. Alizadeh, F., Haeberly, J.-P. A., Overton, M. L. (1998): Primal-dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results. SIAM J. Optim. **8**, 746–768
4. Goffin, J.-L., Haurie, A., Vial, J.-P. (1992): Decomposition and nondifferentiable optimization with the projective algorithm. Manage. Sci. **38**, 284–302
5. Goffin, J.-L., S.-Mokhtarian, F. (1999): Primal-dual-infeasible Newton approach for the analytic center deep-cutting plane method. J. Optimization Theory Appl. **101**, 35–58
6. Goldfarb, D., Scheinberg, K. (1999): On parametric semidefinite programming. Appl. Numer. Math. **29**, 361–377
7. Golub, G. H., Van Loan, C. F. (1996): Matrix Computations, Third Edition. The Johns Hopkins University Press

8. Greenberg, H. J. (1994): The use of the optimal partition in linear programming solution for post-optimal analysis. Oper. Res. Lett. **15**, 179–185
9. Helmberg, C., Rendl, F., Vanderbei, R., Wolkowicz, H. (1996): An interior-point method for semidefinite programming. SIAM J. Optim. **6**, 342–361
10. Hoffman, A. (1952): On approximate solutions of systems of linear inequalities. J. Res. Natl. Bur. Stand. **49**, 263–265
11. Holder, A. G., Sturm, J. F., Zhang, S. (1998): Analytic central path, sensitivity analysis and parametric linear programming. Technical Report No. 9801/A, Econometric Institute, Erasmus University Rotterdam, The Netherlands
12. Jansen, B., de Jong, J. J., Roos, C., Terlaky, T. (1997): Sensitivity analysis in linear programming: just be careful! Eur. J. Oper. Res. **101**, 15–28
13. Kim, W.-J., Park, C.-K., Park, S. (1999): An $\epsilon$-sensitivity analysis in the primal-dual interior point method. Eur. J. Oper. Res. **116**, 629–639
14. Kojima, M., Shindoh, S., Hara, S. (1997): Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices. SIAM J. Optim. **7**, 86–125
15. Monteiro, R. D. C. (1997): Primal-dual path following algorithms for semidefinite programming. SIAM J. Optim. **7**, 663–678
16. Monteiro, R. D. C., Zanjacomo, P. (1997): A note on the existence of the Alizadeh-Haeberly-Overton direction for semidefinite programming. Math. Program. **78**, 393–396
17. Monteiro, R. D. C., Zhang, Y. (1998): A unified analysis for a class of path-following primal-dual interior-point algorithms for semidefinite programming. Math. Program. **81**, 281–299
18. Nesterov, Yu. E., Todd, M. J. (1997): Self-scaled barriers and interior-point methods in convex programming. Math. Oper. Res. **22**, 1–42
19. Nesterov, Yu. E., Todd, M. J. (1997): Primal-dual interior-point methods for self-scaled cones. SIAM J. Optim. **8**, 324–364
20. Nunez, M. A., Freund, R. M. (1998): Condition measures and properties of the central trajectory of a linear program. Math. Program. **83**, 1–28
21. Roos, C., Terlaky, T., Vial, J.-P. (1997): Theory and Algorithms for Linear Optimization: An Interior Point Approach. John Wiley & Sons, Chichester

22. Shida, M., Shindoh, S., Kojima, M. (1998): Existence and uniqueness of search directions in interior-point algorithms for the SDP and the monotone SDLCP. SIAM J. Optim. **8**, 387–396.
23. Sturm, J. F., Zhang, S. (1998): On sensitivity of central solutions in semidefinite programming. Technical Report No. 9813/A, Econometric Institute, Erasmus University Rotterdam, The Netherlands
24. Todd, M. J. (1999): A study of search directions in primal-dual interior-point methods for semidefinite programming. Optim. Methods Softw. **11**, 1–46
25. Todd, M. J., Toh, K. C., Tütüncü, R. H. (1998): On the Nesterov-Todd direction in semidefinite programming. SIAM J. Optim. **8**, 769–796
26. Zhang, S. (1994): On the strictly complementary slackness relation in linear programming. In: Du, D. Z., Sun, J., eds., Advances in Optimization and Approximation, pp. 347–361, Kluwer
27. Zhang, Y. (1998): On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming. SIAM J. Optim. **8**, 365–386