

A study of search directions in primal-dual interior-point methods for semidefinite programming

M. J. Todd *

February 23, 1999

*School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York 14853, USA (miketodd@cs.cornell.edu). Research supported in part by NSF through grant DMS-9505155 and ONR through grant N00014-96-1-0050.

Abstract

We discuss several different search directions which can be used in primal-dual interior-point methods for semidefinite programming problems and investigate their theoretical properties, including scale invariance, primal-dual symmetry, and whether they always generate well-defined directions. Among the directions satisfying all but at most two of these desirable properties are the Alizadeh-Haeberly-Overton, Helmberg-Rendl-Vanderbei-Wolkowicz/Kojima-Shindoh-Hara/Monteiro, Nesterov-Todd, Gu, and Toh directions, as well as directions we will call the MTW and Half directions. The first five of these appear to be the best in our limited computational testing also.

Key words: semidefinite programming, search direction, invariance properties.

AMS Subject classification: 90C05.

Abbreviated title: Search directions in SDP

1 Introduction

This paper is concerned with interior-point methods for semidefinite programming (SDP) problems and in particular the various search directions they use and their properties. We consider the SDP given in the following standard form:

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

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

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

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

We have stated our semidefinite programming problems with the A_i 's and C arbitrary members of $\mathcal{S}\mathcal{R}^{n \times n}$, and then the variables X and S also range over this space. For many applications, the A_i 's and C lie in a subspace consisting of block-diagonal symmetric matrices with a certain fixed block structure, and then X (without loss of generality) and S (from the constraints) can be restricted to the same subspace. All that follows is also valid when we consider such a subspace (in particular, the directions we propose will automatically lie in this subspace), but we treat the ‘‘dense’’ case for simplicity. If the subspace consists of diagonal matrices, then (SDP) and (SDD) reduce to a linear programming (LP) problem in standard form and its dual (with the vectors x and s embedded as the diagonal entries of diagonal matrices X and S).

SDPs have a wide range of applications in both continuous and combinatorial optimization (we refer the reader to [37] and [6] for an extensive list of applications). Interior-point methods for their solution were pioneered by Alizadeh [1] and Nesterov and Nemirovskii [25, 26] independently. These methods were primal; more efficient algorithms use a primal-dual approach. For a comprehensive list of publications concerned with both algorithms and applications and related software, see the semidefinite programming home pages maintained by Alizadeh [2] and Helmberg [8].

Several different primal-dual interior-point algorithms for SDP have been proposed, differing in their overall strategy (path-following or potential-reduction, for example), whether they restrict the iterates to lie in a neighborhood of the central path, whether they employ a predictor-corrector technique, whether they require all iterates to be feasible, etc. All these are concerns of similar algorithms for LP. But there is another aspect of such methods for SDP that does not arise in LP: how the search directions are defined. In LP, although there are variants that use different right-hand sides

(target-following versus path-following, for example), the left-hand side of the linear system used by (almost) all primal-dual methods is the same, and arises from finding a Newton step for a nonlinear system representing primal and dual feasibility plus perturbed complementarity. This is no longer the case for SDP, as we shall see in the next section. The primary reason for the occurrence of different search directions is that symmetric matrices do not commute! As we will see, if all the symmetric matrices arising in a problem commute, almost all directions proposed yield the same solution (this is the case for the SDP problem corresponding to an LP problem, when all the matrices are diagonal and thus commute).

Our concern in this paper is with the various search directions that have been proposed in the literature, as well as some others that are suggested here. We are concerned with specific directions, rather than the families that have been introduced by Kojima, Shindoh, and Hara [13], Monteiro and Zhang [19, 38, 22, 20], Monteiro and Tsuchiya [21], and Tseng [34] (and the very general family, including all of these, introduced by Kojima, Shida, and Shindoh [15]). However, many of the specific directions we consider do lie in these families; in particular, out of the twenty directions we address, six lie in the Monteiro-Zhang and nine in the Monteiro-Tsuchiya families (with some overlap). The other directions are mostly defined by a linear system that arises from considering Newton's method for some nonlinear system. For this reason, we do not consider other members of the Kojima-Shindoh-Hara or Tseng primal-dual families than the basic ones, since these families do not have this motivation. We also do not consider the Gauss-Newton direction recently proposed by Kruk et al. [16], which is based on a least-squares solution to an over-determined linear system rather than the exact solution to a square system.

We provide motivation for the directions we discuss and investigate their basic properties. We suggest that these desirable properties be considered in designing new directions for use in interior-point methods for SDP. (See also the related paper of Tunçel [36], who studies two particular properties, primal-dual symmetry and scale invariance, in the general context of conic programming.)

We are not concerned here with convergence questions of methods using these directions, which depend heavily on other aspects of the algorithms such as neighborhoods and step sizes. Convergence results for short-step path-following methods were given in [13, 19, 20, 21, 34, 38]. For a very general global convergence result for short-step algorithms and references to the literature, see Kojima, Shida, and Shindoh [15]; they give a bound of $O(\sqrt{n} \ln(1/\epsilon))$ iterations to obtain ϵ -optimal solutions given a suitable starting point, which applies to all the families of search directions described above. Kojima et al. also provide a figure showing the relationship between these families of search directions. Convergence results for long-step algorithms can be found in Monteiro [19] (who proves a bound of $O(n^{3/2} \ln(1/\epsilon))$ iterations for two particular search directions), in Monteiro and Zhang [22] (who extend the results of [19] and establish a bound of $O(n \ln(1/\epsilon))$ iterations for another search direction) and Monteiro and Tsuchiya [21] (who give a bound of $O(n^{3/2} \ln(1/\epsilon))$ iterations for all directions in a subclass of the Monteiro-Tsuchiya family). Ji, Potra, and Sheng [12] describe the literature on local convergence and prove convergence of Q-order 1.5 or 2 for predictor-corrector algorithms using certain search directions from the Monteiro-Zhang family. Nor do we

discuss the amount of computational work involved in computing our search directions in much detail; see Monteiro and Zanjacomo [24] and Toh [32] for flop counts for some of these directions.

In Section 2 we discuss the central path and how directions might be defined to approximate a sequence of points on the path. We also introduce some notation and various notions of geometric mean. The following section describes the Monteiro-Zhang and Monteiro-Tsuchiya families and, in order to develop them, the idea of scaling. Section 4 then gives twenty different primal-dual search directions. Each is defined as the solution of a linear system of equations, and Section 5 discusses how these equations can be solved efficiently. We then turn to an investigation in Section 6 of a number of desirable properties for a search direction for SDP to possess, including that it should extend the well-known primal-dual direction for LP; it should give a predictable decrease to the duality gap if the current iterate is feasible; it should yield a well-defined direction at any iterate; it should satisfy certain scale invariance properties; and it should be symmetric between the primal and dual. We also discuss several other attractive properties. We list the directions that satisfy all but a very small number of these properties: they include the Alizadeh-Haeberly-Overton [3], Helmberg-Rendl-Vanderbei-Wolkowicz/Kojima-Shindoh-Hara/Monteiro [9, 13, 19], and Nesterov-Todd [27, 28, 31] directions, as well as two recently proposed by Gu [7] and Toh [32], and directions we call the MTW and Half directions in addition. The first five of these appear to be the best in our limited computational testing also; this, and our recommendations, are discussed in the concluding Section 7.

2 The central path and preliminaries

Most interior-point methods for semidefinite programming are based (if only loosely) on approximating a sequence of points on the central path. Assuming that both (*SDP*) and (*SDD*) have strictly feasible solutions, we define this path as the set of solutions (X_ν, y_ν, S_ν) for $\nu > 0$ to the central path equations

$$\begin{array}{rcl} & A_i \bullet X & = b_i, \text{ for } i = 1, \dots, m, \\ \sum_{i=1}^m y_i A_i & + S & = C, \\ & XS & = \nu I \end{array} \quad (1)$$

(together with the requirement that X and S be symmetric positive definite). Existence is guaranteed by general results of Nesterov and Nemirovskii [25, 26]; see also Kojima, Shindoh, and Hara [13]. Note that the first two block equations above are linear, while the third is mildly nonlinear. Hence a Newton step seems a natural idea for an iterative algorithm. Unfortunately, the residual map, for which a zero is sought, takes an iterate $(X, y, S) \in \mathcal{S}\mathbb{R}^{n \times n} \times \mathbb{R}^m \times \mathcal{S}\mathbb{R}^{n \times n}$ to a point in $\mathbb{R}^m \times \mathcal{S}\mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n}$ (since $XS - \nu I$ is in general not symmetric), which is a space of higher dimension, and so Newton's method cannot be applied directly.

A natural remedy is to rewrite the last equation so that the resulting residual lies in $\mathcal{S}\mathbb{R}^{n \times n}$. For example, it could be written as $S - \nu X^{-1} = 0$ or $X - \nu S^{-1} = 0$ (corresponding to $S + \nu F'(X) = 0$ or $X + \nu F'(S) = 0$, where $F(U) := -\ln \det U$ is

the standard n -self-concordant barrier for the symmetric positive semidefinite cone). However, the inverse map is not as smooth as the bilinear map $(X, S) \rightarrow XS$, particularly as the solution is approached. Hence other symmetrizations are desirable. We will discuss two general families in the next section, and several specific methods in Section 4.

We devote the rest of this section to introducing some useful notation and definitions. Henceforth, we will usually denote by (X, y, S) the current iterate, not necessarily feasible, but with X and S symmetric positive definite. We therefore use $(\tilde{X}, \tilde{y}, \tilde{S})$ for generic solutions. As is customary, lower-case Roman letters denote vectors and upper-case letters denote $n \times n$ matrices; we reserve K, L, P , and Q (Q will usually be orthogonal) for not necessarily symmetric matrices, with all other letters denoting members of $S\mathbb{R}^{n \times n}$. We write $U^{1/2}$ to denote the symmetric positive semidefinite square root of the symmetric positive semidefinite matrix U . We use lower-case Greek letters for scalars, and script letters for linear operators on (usually symmetric) matrices. In particular, $\mathcal{A} : S\mathbb{R}^{n \times n} \rightarrow \mathbb{R}^m$ is defined by

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

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

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

We will use r and R for residuals. We define

$$r_p := b - \mathcal{A}X, \quad (4)$$

$$R_d := C - \mathcal{A}^*y - S, \quad (5)$$

which are zero if the current iterate is feasible. Each direction will be denoted by $(\Delta X, \Delta y, \Delta S) \in S\mathbb{R}^{n \times n} \times \mathbb{R}^m \times S\mathbb{R}^{n \times n}$.

Several of our directions will be Newton steps for nonlinear systems of the form

$$\begin{aligned} \mathcal{A}\tilde{X} &= b, \\ \mathcal{A}^*\tilde{y} + \tilde{S} &= C, \\ \Theta(\tilde{X}, \tilde{S}) &= 0, \end{aligned} \quad (6)$$

where the last equation is some symmetrization of $\tilde{X}\tilde{S} - \nu I = 0$. (We will refer to such a direction as the Newton step for $\Theta(\tilde{X}, \tilde{S}) = 0$, with the other equations of (6) implicit.) Hence the direction will satisfy the system

$$\begin{aligned} \mathcal{A}\Delta X &= r_p, \\ \mathcal{A}^*\Delta y + \Delta S &= R_d, \\ \mathcal{E}\Delta X + \mathcal{F}\Delta S &= R_{EF}, \end{aligned} \quad (7)$$

where the operators $\mathcal{E} = \mathcal{E}(X, S)$ and $\mathcal{F} = \mathcal{F}(X, S)$ are the derivatives of Θ with respect to \tilde{X} and \tilde{S} respectively, evaluated at (X, S) , and $R_{EF} = R_{EF}(X, S) = -\Theta(X, S)$.

Such a general derivation of search directions was considered also in Monteiro and Zanjacomo [24].

For some directions it will be more convenient to introduce another variable $Z \in S\mathbb{R}^{n \times n}$ and to consider the system

$$\begin{aligned} \mathcal{A}^* \Delta y & \quad \mathcal{A} \Delta X & & = r_p, \\ & + \Delta S & & = R_d, \\ & \mathcal{E} \Delta X & - 2\mathcal{G}Z & = R_E, \\ & & \mathcal{F} \Delta S + 2\mathcal{H}Z & = R_F, \end{aligned} \tag{8}$$

for certain operators \mathcal{E} , \mathcal{F} , \mathcal{G} , and \mathcal{H} , again depending on X and S . (The factor 2 and the minus sign will make the notation later simpler.)

All such operators \mathcal{E} through \mathcal{H} will take $S\mathbb{R}^{n \times n}$ into itself. Almost always they are of the form $\mathcal{E}U = (PUQ^T + QUP^T)/2$, so we will find it convenient to have a special notation for this operator (it corresponds to the symmetric Kronecker product introduced by Alizadeh, Haerberly, and Overton [3], and we will use a similar notation). In fact, we will extend its definition to all of $\mathbb{R}^{n \times n}$:

$$(P \odot Q)K := \frac{1}{2}(PKQ^T + QK^T P^T), \tag{9}$$

but most of the time we regard it as an operator from $S\mathbb{R}^{n \times n}$ to itself. In this case, 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

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

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

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

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

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

but there is no simple expression for $(P \odot Q)^{-1}$ in general. Note that $I \odot I$ is the identity operator. We write \mathcal{E}^{-*} for the inverse of the adjoint of \mathcal{E} , or equivalently the adjoint of its inverse, assuming it exists, and similarly for \mathcal{F} , \mathcal{G} , and \mathcal{H} .

Assuming that \mathcal{E} is nonsingular, we find that (7) has a unique solution iff the $m \times m$ Schur complement matrix $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$ is nonsingular, and in this case the solution can be found from

$$\begin{aligned} (\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*)\Delta y & = r_p - \mathcal{A}\mathcal{E}^{-1}(R_{EF} - \mathcal{F}R_d), \\ \Delta S & = R_d - \mathcal{A}^*\Delta y, \\ \Delta X & = \mathcal{E}^{-1}(R_{EF} - \mathcal{F}\Delta S). \end{aligned} \tag{10}$$

Similarly, assuming \mathcal{E} and \mathcal{H} are nonsingular, we see that (8) has a unique solution iff $\mathcal{A}\mathcal{E}^{-1}\mathcal{G}\mathcal{H}^{-1}\mathcal{F}\mathcal{A}^*$ is nonsingular, and in this case the solution can be found from

$$\begin{aligned} (\mathcal{A}\mathcal{E}^{-1}\mathcal{G}\mathcal{H}^{-1}\mathcal{F}\mathcal{A}^*)\Delta y & = r_p - \mathcal{A}\mathcal{E}^{-1}(R_E + \mathcal{G}\mathcal{H}^{-1}(R_F - \mathcal{F}R_d)), \\ \Delta S & = R_d - \mathcal{A}^*\Delta y, \\ \Delta X & = \mathcal{E}^{-1}(R_E + \mathcal{G}\mathcal{H}^{-1}(R_F - \mathcal{F}\Delta S)). \end{aligned} \tag{11}$$

We will discuss solving systems defined by the linear operators \mathcal{E} and \mathcal{H} (in order to compute the Schur complement matrices and the right-hand sides) in Section 5, after introducing our directions.

For every such method for obtaining a search direction, there is a corresponding dual method defined as follows — roughly by interchanging the roles of X and S . First, the problem (SDD) can be rewritten in the form of (SDP) in terms of the variable \tilde{S} , say as (\overline{SDP}) . An iterate (X, y, S) for (SDP) and its dual can be viewed as an iterate (S, \tilde{y}, X) for (\overline{SDP}) and its dual. (If X is not feasible in (SDP) , then \tilde{y} cannot be chosen so that (\tilde{y}, X) is feasible in the dual of (\overline{SDP}) , so that it is not well-defined. But it can be seen from (7) and (8) that ΔX and ΔS are independent of y , which only affects Δy , so the choice is not important.) We then apply our method to (\overline{SDP}) and its dual at the iterate (S, \tilde{y}, X) to get the direction $(\Delta S, \Delta \tilde{y}, \Delta X)$. The dual direction is then $(\Delta X, \Delta y, \Delta S)$, where Δy is chosen so that the second equation of (7) is satisfied. If a method is defined by (7) with a certain $\mathcal{E}(X, S)$, $\mathcal{F}(X, S)$, and $R_{EF}(X, S)$, then the dual method is again determined by (7), but using $\mathcal{E}'(X, S) := \mathcal{F}(S, X)$, $\mathcal{F}'(X, S) := \mathcal{E}(S, X)$, and $R'_{EF}(X, S) := R_{EF}(S, X)$, i.e., X and S are interchanged. A similar statement holds for methods determined by (8).

We call a method *primal-dual symmetric* if the dual direction always coincides with the original direction. Actually, we need to be a little more careful. Several methods leave the direction undefined at certain iterates, since the corresponding equations have no or multiple solutions due to singularity. We say the method is primal-dual symmetric if any solution to the equations “defining” the direction also solves the equations “defining” the dual direction.

A method given by (7) is clearly primal-dual symmetric if $\mathcal{E}(X, S) = \mathcal{F}(S, X)$ and $R_{EF}(X, S) = R_{EF}(S, X)$. These conditions are sufficient but not necessary, since the last equation of (7) can be written in many equivalent ways, and sometimes the particular form may not exhibit the desired symmetry. However, they often suffice to show primal-dual symmetry in our examples. Similarly, a method given by (8) is certainly primal-dual symmetric if $\mathcal{E}(X, S) = \mathcal{F}(S, X)$, $\mathcal{G}(X, S) = -\mathcal{H}(S, X)$, and $R_E(X, S) = R_F(S, X)$.

To conclude this section we define two geometric means and hence several functions of X and S that will be useful in what follows. Given symmetric positive definite U and V , the unique matrix J satisfying

$$JU^{-1}J = V$$

is called the metric geometric mean of U and V , denoted $U\#V$ [4, 11]. It is easy to see that $U\#V = V\#U = (U^{-1}\#V^{-1})^{-1}$, that $U\#V = U^{1/2}V^{1/2} = V^{1/2}U^{1/2}$ if U and V commute, and that

$$\begin{aligned} U\#V &= V^{1/2}(V^{-1/2}UV^{-1/2})^{1/2}V^{1/2} \\ &= U^{1/2}(U^{-1/2}VU^{-1/2})^{1/2}U^{1/2}. \end{aligned}$$

We define

$$W := X\#S^{-1} = X^{1/2}(X^{1/2}SX^{1/2})^{-1/2}X^{1/2} = S^{-1/2}(S^{1/2}XS^{1/2})^{1/2}S^{-1/2}, \quad (12)$$

so that $WSW = X$. W is the *scaling matrix* defined by X and S ; see Nesterov and Todd [27, 28] and Todd, Toh, and Tütüncü [31]. Let

$$T := W^{-1} = X^{-1} \# S, \quad (13)$$

so that $TXT = S$. Next we define

$$V := W^{1/2} S W^{1/2} = W^{-1/2} X W^{-1/2}, \quad (14)$$

V is the spectral geometric mean of X and S . We denote

$$J := X \# S, \quad (15)$$

so that $JS^{-1}J = X$, and let

$$K := JS^{-1} = XJ^{-1} \quad (16)$$

(note that K is not necessarily symmetric).

3 The Monteiro-Zhang and Monteiro-Tsuchiya families

Here we describe two general symmetrization schemes and use them to define the property of scale invariance.

The first technique we discuss is due to Monteiro [19] and Zhang [38] (see also Monteiro and Zhang [22] and Monteiro [20]): choose a nonsingular matrix $P \in \mathbb{R}^{n \times n}$ and replace $\tilde{X}\tilde{S} = \nu I$ by

$$H_P(\tilde{X}\tilde{S}) = H_P(\nu I) = \nu I, \quad (17)$$

where $H_P := P \odot P^{-T}$ so that

$$H_P(Q) = \frac{1}{2}(PQP^{-1} + P^{-T}Q^T P^T). \quad (18)$$

Hence, if we regard P as fixed (although it may be a function of X and S), the Newton step for (17) is the solution of

$$\begin{aligned} \mathcal{A}^* \Delta y & \quad \mathcal{A} \Delta X & & = r_p, \\ & + \Delta S & & = R_d, \\ \frac{1}{2}(P \Delta X S P^{-1} + P^{-T} S \Delta X P^T) & + \frac{1}{2}(P X \Delta S P^{-1} + P^{-T} \Delta S X P^T) & = R_P, \end{aligned} \quad (19)$$

with $R_P := \nu I - (P X S P^{-1} + P^{-T} S X P^T)/2$. This is of the form (7), with

$$\mathcal{E} := P \odot P^{-T} S, \quad \mathcal{F} := P X \odot P^{-T}, \quad (20)$$

and

$$R_{EF} := R_P = \nu I - \frac{1}{2}(P X S P^{-1} + P^{-T} S X P^T). \quad (21)$$

Directions arising in this way are called members of the Monteiro-Zhang (MZ) family.

We can alternatively pre- and post-multiply the last equation of (19) by P^T and P , to get an equivalent system of the form (7), where now

$$\mathcal{E} := S \odot M, \quad \mathcal{F} := MX \odot I, \quad (22)$$

and

$$R_{EF} := \nu M - \frac{1}{2}(MXS + SXM), \quad (23)$$

with

$$M := P^T P.$$

These equations also arise directly as the Newton system for

$$\frac{1}{2}(M\tilde{X}\tilde{S} + \tilde{S}\tilde{X}M) = \nu M. \quad (24)$$

The second formulation shows that the same direction results if we replace P by QP for any orthogonal Q , since M is thus unchanged. Hence we can remove the ambiguity by requiring P to be symmetric positive definite, $P = M^{1/2}$. The reformulation also shows an alternative parametrization of the Monteiro-Zhang family in terms of M .

The choice of P or M often depends on the current iterates X and S ; we sometimes write $P(X, S)$ or $M(X, S)$ to highlight this dependence. Four particular choices for P and M are: $P = M = I$, giving the Alizadeh-Haerberly-Overton (AHO) direction [3]; $P = S^{1/2}$, $M = S$, giving the Helmberg-Rendl-Vanderbei-Wolkowicz [9]/ Kojima-Shindoh-Hara [13]/ Monteiro [19] (H..K..M) direction; $P = X^{-1/2}$, $M = X^{-1}$, giving the dual H..K..M direction [13, 19]; and $P = W^{-1/2}$, $M = W^{-1}$, where W is the scaling matrix of (12), giving the Nesterov-Todd (NT) direction [27, 28, 31].

The H..K..M direction was independently introduced by Helmberg et al., whose motivation was symmetrizing the direction ΔX resulting from applying Newton's method to the asymmetric equation $\tilde{X}\tilde{S} = \nu I$, and by Kojima et al., whose motivation came from a more general scheme for generating symmetric search directions based on subspaces of skew-symmetric matrices. Monteiro rediscovered the direction based on the motivation above. The dual H..K..M direction is also a member of the Kojima-Shindoh-Hara family and, as we have seen, of the MZ family. It was also discussed by [13, 19]. Alizadeh et al. based their direction directly on applying Newton's method to $\tilde{X}\tilde{S} + \tilde{S}\tilde{X} = 2\nu I$. Finally, Nesterov and Todd's direction is a specialization to SDP of a general primal-dual symmetric scheme for generating search directions for certain conic problems, to be used in path-following or potential-reduction methods. It was shown to be in the MZ family by Todd-Toh-Tütüncü [31], and to be in the KSH family by Kojima-Shida-Shindoh [14].

It is easy to see that the AHO and NT directions are primal-dual symmetric, while small examples show that the H..K..M (and hence the dual H..K..M) direction is not. However, the dual of any direction in the MZ family is also in the family: merely replace $P = P(X, S)$ by $[P(S, X)]^{-T}$ or $M = M(X, S)$ by $[M(S, X)]^{-1}$.

There is another way to view the MZ family, which will be very important in our future discussions. We consider a *change of variables* in problem (SDP), where \tilde{X} is

replaced by $\hat{X} := P\tilde{X}P^T$. In terms of these variables, (SDP) becomes

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

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

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

which can easily be seen to result from (SDD) by the change of variables where (\tilde{y}, \tilde{S}) is replaced by $(\hat{y} := \tilde{y}, \hat{S} := P^{-T}\tilde{S}P^{-1})$.

Note that, if $(\Delta X, \Delta y, \Delta S)$ denotes the direction that solves (19) at (X, y, S) , then the scaled direction $(\Delta \hat{X}, \Delta \hat{y}, \Delta \hat{S}) := (P\Delta X P^T, \Delta y, P^{-T}\Delta S P^{-1})$ at $(\hat{X}, \hat{y}, \hat{S}) = (PXP^T, y, P^{-T}SP^{-1})$ solves

$$\begin{aligned} \hat{A}\Delta\hat{X} &= \hat{r}_p, \\ \hat{A}^*\Delta\hat{y} + \Delta\hat{S} &= \hat{R}_D, \\ \frac{1}{2}(\Delta\hat{X}\hat{S} + \hat{S}\Delta\hat{X}) + \frac{1}{2}(\hat{X}\Delta\hat{S} + \Delta\hat{S}\hat{X}) &= \nu I - \frac{1}{2}(\hat{X}\hat{S} + \hat{S}\hat{X}), \end{aligned}$$

where $\hat{r}_p := b - \hat{A}\hat{X}$, $\hat{R}_D := \hat{C} - \hat{A}^*\hat{y} - \hat{S}$, so it is the AHO direction for the scaled problem. Hence each MZ direction can be viewed as the result of scaling by P , taking the AHO direction, and then unscaling.

A subclass of the MZ family, called the MZ* or TTT family, consists of those members of the MZ family for which $\mathcal{E}^{-1}\mathcal{F}$ is self-adjoint, or equivalently [31], for which PXP^T and $P^{-T}SP^{-1}$ commute. In terms of $M = P^T P$, this is the class for which MXS is symmetric. This includes the H..K..M, dual H..K..M, and NT directions, but not in general the AHO direction.

It is now simple to define the notion of scale invariance. 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 the iterate by an arbitrary nonsingular 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 we restrict ourselves to scalings defined by orthogonal matrices P . Once again we need to be a little careful, in case the method does not define the direction uniquely. We say the method is scale-invariant if, whenever we scale any solution to the equations “defining” the direction for the original problem, the result is a solution to the equations “defining” the direction for the scaled problem.

(A helpful way to view this is for the case where all A_i 's, C , and hence X and S are diagonal matrices, so that (SDP) is a linear programming problem. In order that P define a transformation preserving this structure, it must have the structure of a permutation matrix, but with arbitrary nonzero entries. Then a method is *P-scale-invariant* if it is invariant under permutation of the diagonal components of X and

under arbitrary positive scalings of these components. Since such a P is orthogonal iff all its nonzero entries are ± 1 , a method is Q -scale-invariant if it is invariant under permutations of the diagonal components of X . Thus the simplex method using Dantzig's most negative reduced cost rule is Q -scale-invariant, while using a steepest-edge rule it is P -scale-invariant.)

Now we describe another general symmetrization scheme, due to Monteiro and Tsuchiya [21]. Recall that the dual H..K..M direction is the Newton direction for the first two equations of (6) and

$$\frac{1}{2}(X^{-1/2}\tilde{X}\tilde{S}X^{1/2} + X^{1/2}\tilde{S}\tilde{X}X^{-1/2}) = \nu I.$$

At the current iterate, the left-hand side becomes $X^{1/2}SX^{1/2}$. We can therefore consider the Newton direction for the first two equations of (6) and

$$\tilde{X}^{1/2}\tilde{S}\tilde{X}^{1/2} = \nu I. \quad (25)$$

Note the difference from the equation above; here we are explicitly taking into account the change in $P = P(\tilde{X}, \tilde{S})$, rather than holding it fixed.

To derive the corresponding Newton step, we need to see how $\tilde{X}^{1/2}$ varies as a function of \tilde{X} . Letting \tilde{U} denote (temporarily) the function $\tilde{X}^{1/2}$ of \tilde{X} , and letting $\tilde{U}'[H]$ denote its derivative in the direction H , we find from $\tilde{U}\tilde{U} = \tilde{X}$ that

$$\tilde{U}\tilde{U}'[H] + \tilde{U}'[H]\tilde{U} = H.$$

Hence the Newton system consists of the first two equations of (8) together with

$$\begin{aligned} \Delta X & - X^{1/2}Z - ZX^{1/2} &= 0, \\ X^{1/2}\Delta SX^{1/2} & + X^{1/2}SZ + ZSX^{1/2} &= \nu I - X^{1/2}SX^{1/2}. \end{aligned}$$

We can also equivalently write these equations as

$$\begin{aligned} \Delta X & - X^{1/2}Z - ZX^{1/2} &= 0, \\ \Delta S & + SZX^{-1/2} + X^{-1/2}ZS &= \nu X^{-1} - S, \end{aligned} \quad (26)$$

which is an instance of (8) with

$$\mathcal{E} := \mathcal{F} := I \odot I, \quad \mathcal{G} := X^{1/2} \odot I, \quad \mathcal{H} := S \odot X^{-1/2},$$

and

$$R_E := 0, \quad R_F := \nu X^{-1} - S.$$

Following Monteiro and Zanjacomo [24], we call this the X-MT direction; a similar S-MT direction (the dual direction) arises from interchanging the roles of X and S above.

Monteiro and Tsuchiya obtain a family of directions by first scaling, then applying the X-MT method, and then unscaling. We call this the MT (or X-MT) family. If

instead we use the S-MT method, we obtain the S-MT family. If we scale by P as in the MZ family, equations (26) are replaced by

$$\begin{aligned} P\Delta XP^T &- (PXP^T)^{1/2}Z &- Z(PXP^T)^{1/2} &= 0, \\ P^{-T}\Delta SP^{-1} &+ P^{-T}SP^{-1}Z(PXP^T)^{-1/2} &+ (PXP^T)^{-1/2}ZP^{-T}SP^{-1} &= R'_P. \end{aligned} \quad (27)$$

with $R'_P := \nu(PXP^T)^{-1} - P^{-T}SP^{-1}$.

Let us pre- and post-multiply the first equation by P^{-1} and P^{-T} respectively, and the second by P^T and P respectively. Let us also write Z for $P^{-1}ZP^{-T}$. Finally, we define

$$N := P^T(PXP^T)^{-1/2}P$$

and note that $XN = P^{-1}PXP^T(PXP^T)^{-1/2}P = N^{-1}P^T P$. Then it is not hard to see that (27) is equivalent to

$$\begin{aligned} \Delta X &- XNZ &- ZNX &= 0, \\ \Delta S &+ SZN &+ NZS &= \nu X^{-1} - S. \end{aligned} \quad (28)$$

Now we have

$$\mathcal{E} := \mathcal{F} := I \odot I, \quad \mathcal{G} := XN \odot I, \quad \mathcal{H} := S \odot N, \quad (29)$$

with again

$$R_E := 0, \quad R_F := \nu X^{-1} - S. \quad (30)$$

We note that the MT family can be parametrized as well by N as by P , since any symmetric positive definite N arises in this way from $P = (NXN)^{1/2}$. Again this means that P and QP , for any orthogonal Q , define the same direction (since N is unchanged), so that we may remove the ambiguity by choosing P to be symmetric positive definite.

Particular members of the MT family are: $P = I$, $N = X^{-1/2}$, giving the X-MT direction; $P = X^{-1/2}$, $N = X^{-1}$, giving the dual H..K..M direction; and $P = S^{1/2}$, $N = W^{-1}$, giving the NT direction. We can also choose $P = (SXS)^{1/2}$, $N = S$, which allows us to solve the second equation in (28) for Z and substitute the result in the first equation; this gives the equation for the H..K..M direction (not previously known to belong to the (X-)MT family). Monteiro and Tsuchiya also mention the choice $P = W^{-1/2}$, $N = W^{-1/2}V^{-1/2}W^{-1/2}$, leading to a direction we shall call the MTW direction. We will also discuss two other members of this family in the next section.

There is also a subfamily of these directions, called the MT*-family, for which Monteiro and Tsuchiya prove stronger convergence results. These are those members of the MT family for which $(PXP^T)^{1/2}P^{-T}SP^{-1} + P^{-T}SP^{-1}(PXP^T)^{1/2}$ is positive definite, or equivalently for which $NXS + SXN$ is positive definite. This holds if $(PXP^T)^{1/2}$ and $P^{-T}SP^{-1}$ commute, or if $N^{1/2}XN^{1/2}$ and $N^{-1/2}SN^{-1/2}$ commute; hence the H..K..M, dual H..K..M, NT, and MTW directions are included.

While the dual of any direction in the MZ family is also in this family, it is not known whether the same is true for the MT family. For instance, it is unclear whether the S-MT direction lies in the MT (or X-MT) family, while it obviously lies in the S-MT family.

4 Twenty primal-dual search directions

Here we describe a great variety of different search directions. The first six are in the MZ family while the next six are in or motivated by the MT family; then there are three “simple” directions and finally five more motivated by other Newton systems. Simple examples show that these twenty directions are all different in general.

All are given by linear systems of the form of (7) or (8), so we usually just give some motivation and then the specific choices of the operators and residuals. Along the way we will discover another useful property for search directions in SDP.

The first four directions are by now well known. They belong to the MZ family and satisfy equations (7). We describe the nonlinear Newton system that generates them and the resulting operators and residual.

AHO [3]. This is the Newton step for

$$\frac{1}{2}(\tilde{X}\tilde{S} + \tilde{S}\tilde{X}) = \nu I.$$

It corresponds to taking

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

H..K..M [9, 13, 19]. This is the Newton step for

$$\frac{1}{2}(S\tilde{X}\tilde{S} + \tilde{S}\tilde{X}S) = \nu S$$

(we have chosen the system corresponding to (24)). Here we have

$$\mathcal{E} = S \odot S, \quad \mathcal{F} = SX \odot I, \quad R_{EF} = \nu S - SXS.$$

Alternatively, so that \mathcal{E} doesn't need to be inverted, we can use

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

Dual H..K..M [13, 19]. This is the Newton step for

$$\frac{1}{2}(X^{-1}\tilde{X}\tilde{S} + \tilde{S}\tilde{X}X^{-1}) = \nu X^{-1},$$

again corresponding to (24). (This can be written more symmetrically with the above as $(X\tilde{S}\tilde{X} + \tilde{X}\tilde{S}X)/2 = \nu X$, and the system defining the direction will be equivalent.) Here we have

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

We will not give dual versions of all the methods we will introduce, but this and the S-MT direction will be given because they have been previously defined.

NT [27, 28, 31]. This is the Newton step for

$$\frac{1}{2}(T\tilde{X}\tilde{S} + \tilde{S}\tilde{X}T) = \nu T$$

(recall that $T = W^{-1}$, the inverse of the scaling matrix), again corresponding to (24). (This can be written in the alternative form $(W\tilde{S}\tilde{X} + \tilde{X}\tilde{S}W)/2 = \nu W$, and the system defining the direction will be equivalent.) Hence we have

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

However, it turns out (see [31]) that an equivalent formulation uses

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

(Indeed, the latter form of the equations was the way the direction was introduced in [27, 28]; it was not clear that the resulting direction lay in the MZ family.) We can also pre- and post-multiply by T , to get another form with $\mathcal{F} = I$, or by $T^{1/2} = W^{-1/2}$, to get a symmetric primal-dual form, but the formulation above requires no inversion of \mathcal{E} . However, for future reference, we do want to point out another symmetric primal-dual form, where we pre- and post-multiply the last-mentioned formulation by $V^{1/2}$. Then we find

$$\mathcal{E} = V^{1/2}W^{-1/2} \odot V^{1/2}W^{-1/2}, \quad \mathcal{F} = V^{1/2}W^{1/2} \odot V^{1/2}W^{1/2}, \quad R_{EF} = \nu I - V^2.$$

The next two directions were recently introduced by Gu [7] and Toh [32], based on a characterization by Gu of the TTT (or MZ*) family, itself a restatement of a characterization by Monteiro and Zhang [22]. Suppose $X^{1/2}S^{1/2} = Q_x \Sigma Q_s^T$, with Q_x and Q_s orthogonal and Σ diagonal. Then the TTT family (in the case that all diagonal entries of Σ are distinct) corresponds to those P that can be written as

$$P = \hat{B}\Sigma^{-1/2}Q_s^T S^{1/2} = \hat{B}\Sigma^{1/2}Q_x^T X^{-1/2}$$

for some diagonal positive definite \hat{B} . The H..K..M, NT, and dual H..K..M directions correspond to \hat{B} equal to $\Sigma^{1/2}$, the identity, and $\Sigma^{-1/2}$ respectively. The choices recommended by Gu and Toh are designed to make $P^T P$ or the 3×3 block matrix in (7) better conditioned. Let Φ (Ψ) be the diagonal matrix so that the rows of $\Phi^{-1}Q_s^T S^{1/2}$ ($\Psi^{-1}Q_x^T X^{1/2}$) have unit length.

Gu [7]. Choose $\hat{B} = \Phi^{-1}\Sigma^{1/2}$ so that $P = \hat{B}\Sigma^{-1/2}Q_s^T S^{1/2}$ has rows of unit length. Set

$$M = P^T P = S^{1/2}Q_s \Phi^{-2}Q_s^T S^{1/2}$$

and then

$$\mathcal{E} = S \odot M, \quad \mathcal{F} = MX \odot I, \quad R_{EF} = \nu M - \frac{1}{2}(MXS + SXM).$$

Toh [32]. Choose $\hat{B} = \Phi^{-1/2}\Psi^{1/2}$ and $P = \hat{B}\Sigma^{-1/2}Q_s^T S^{1/2}$. (Then, according to [32], P and P^{-T} have rows of the same lengths.) Set

$$M = P^T P = S^{1/2}Q_s \Sigma^{-1} \Phi^{-1} \Psi Q_s^T S^{1/2}$$

and then \mathcal{E} , \mathcal{F} , and R_{EF} as above.

Next we proceed to the members of the Monteiro-Tsuchiya family. First we have the X-MT direction:

X-MT [21]. This is the Newton step for

$$\tilde{X}^{1/2}\tilde{S}\tilde{X}^{1/2} = \nu I.$$

We have a system of the form (8) with

$$\mathcal{E} = \mathcal{F} = I \odot I, \quad \mathcal{G} = X^{1/2} \odot I, \quad \mathcal{H} = S \odot X^{-1/2}$$

and

$$R_E = 0, \quad R_F = \nu X^{-1} - S.$$

This corresponds to $P = I$ and $N = X^{-1/2}$.

MTI. Our next direction is also in the MT family, corresponding to $P = X^{1/2}$ so that $N = I$. However, there is another motivation which we will see later, so we discuss it now. The X-MT direction involves some square roots. Let us write the last two equations defining the direction as

$$\begin{aligned} \Delta X & - X[X^{-1/2}Z] - [ZX^{-1/2}]X = 0, \\ \Delta S & + S[ZX^{-1/2}] + [X^{-1/2}Z]S = \nu X^{-1} - S. \end{aligned}$$

Now we replace $[X^{-1/2}Z]$ by Z , and pretend that it is symmetric, so that we also replace $[ZX^{-1/2}]$ by Z . The result is of the form of (8), with

$$\mathcal{E} = \mathcal{F} = I \odot I, \quad \mathcal{G} = X \odot I, \quad \mathcal{H} = S \odot I$$

and

$$R_E = 0, \quad R_F = \nu X^{-1} - S.$$

These are also the formulae resulting from setting P and N as above in the MT family.

We have already considered several members of this family. $P = I$, $N = X^{-1/2}$ gives the X-MT direction, while $P = X^{1/2}$, $N = I$ yields the MTI direction. Also, $P = X^{-1/2}$, $N = X^{-1}$ gives the dual H.K.M direction, $P = S^{1/2}$, $N = W^{-1}$ gives the NT direction, and, as mentioned in the last section, $P = (SXS)^{1/2}$, $N = S$ leads to the H.K.M direction. There are two other choices we will discuss. First, given the range of P 's above, it seems worthwhile to consider also $P = S^{-1/2}$, which corresponds to $N = J^{-1}$ (recall that J is the metric geometric mean of X and S as defined in (15)). Second, as mentioned by Monteiro and Tsuchiya, another reasonable choice is $P = W^{-1/2}$ (so that X and S are both scaled to V), which corresponds to $N = X^{-1}\#W^{-1}$.

MTJ. Here we choose $P = S^{-1/2}$ and $N = J^{-1}$ so that

$$\mathcal{E} = \mathcal{F} = I \odot I, \quad \mathcal{G} = XJ^{-1} \odot I, \quad \mathcal{H} = S \odot J^{-1}$$

and

$$R_E = 0, \quad R_F = \nu X^{-1} - S.$$

MTW [21]. Here we choose $P = W^{-1/2}$ and $N = W^{-1/2}V^{-1/2}W^{-1/2} = X^{-1}\#W^{-1}$ so that

$$\mathcal{E} = \mathcal{F} = I \odot I, \quad \mathcal{G} = X(X^{-1}\#W^{-1}) \odot I, \quad \mathcal{H} = S \odot (X^{-1}\#W^{-1})$$

and

$$R_E = 0, \quad R_F = \nu X^{-1} - S.$$

The other two directions motivated by the MT family are the dual X-MT direction, the S-MT direction, and another direction discovered by Monteiro and Zanjacomo [24] using a Cholesky factorization of S .

S-MT [21, 24]. This is the Newton step for

$$\tilde{S}^{1/2} \tilde{X} \tilde{S}^{1/2} = \nu I,$$

which gives a system of the form (8) with

$$\mathcal{E} = \mathcal{F} = I \odot I, \quad \mathcal{G} = X \odot S^{-1/2}, \quad \mathcal{H} = S^{1/2} \odot I$$

and

$$R_E = \nu S^{-1} - X, \quad R_F = 0.$$

S-Chol-MT [24]. This is the Newton step for

$$\tilde{L}_s^T \tilde{X} \tilde{L}_s = \nu I,$$

where $\tilde{L}_s \tilde{L}_s^T = \tilde{S}$ is the Cholesky factorization of \tilde{S} . Let us use P temporarily for \tilde{L}_s and $Q[H]$ for its derivative with respect to \tilde{S} in the direction H . Then we find

$$PQ[H]^T + Q[H]P^T = H.$$

Thus the Newton direction satisfies the first two equations of (8) together with

$$\begin{aligned} L_s^T \Delta X L_s &+ L_s^T X Q + Q^T X L_s = \nu I - L_s^T X L_s, \\ \Delta S &- L_s Q^T - Q L_s^T = 0, \end{aligned} \quad (31)$$

where L_s is the Cholesky factor of S and Q is an auxiliary lower triangular matrix variable. Note that the first equation can equivalently be written

$$\Delta X + X Q L_s^{-1} + L_s^{-T} Q^T X = \nu S^{-1} - X.$$

We now introduce three “simple” directions. The first is just the primal-dual version of the primal direction that results from taking a Newton step for the primal barrier problem. We can view it as taking a Newton step for another equation describing the central path.

Primal [25, 26, 1]. This is the Newton step for

$$\tilde{S} - \nu\tilde{X}^{-1} = 0.$$

It satisfies the linear system (7) with

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

Alternatively, to avoid the inversion of \mathcal{E} , we can equivalently choose

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

Finally, a formulation that looks a little more symmetric between X and S uses

$$\mathcal{E} = \nu X^{-1/2} \odot X^{-1/2}, \quad \mathcal{F} = X^{1/2} \odot X^{1/2}, \quad R_{EF} = \nu I - X^{1/2} S X^{1/2}. \quad (32)$$

Benson et al. [5] have pointed out that for certain applications, where the A_i 's and C are very sparse, S and even S^{-1} are much sparser than X . In this case it might be very worthwhile to consider the **Dual** direction, the Newton step for

$$\tilde{X} - \nu\tilde{S}^{-1} = 0.$$

It satisfies the linear system (7) with

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

It turns out that Δy and ΔS are independent of X , so there is very little computational penalty for having a dense X matrix. This direction shares theoretical properties with the Primal direction (of which it is the dual) and will not be considered further.

Half. This direction is not a Newton step to our knowledge. It is motivated by trying to symmetrize (32), or a similar version of the equation defining the Dual direction, or alternatively symmetrizing the Newton system (for not necessarily symmetric matrices) coming from $\tilde{X}\tilde{S} - \nu I = 0$. This direction was also discussed by Tseng [34]. It has the form of (7) with

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

Again there is an equivalent form with \mathcal{E} the identity:

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

Half. This direction is also not a Newton step, but is again a simple symmetrization of the Primal system. It has the form of (7) with

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

Again there is an equivalent form with \mathcal{E} the identity:

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

The last group of directions we consider is motivated, directly or indirectly, by taking a Newton step. We first note that, since the metric geometric mean J of (15) is symmetric positive definite, $J^2 = \nu I$ iff $J = \sqrt{\nu} I$, and this holds iff $XS = \nu I$. Hence we consider

J. This is the Newton step for

$$\tilde{J}^2 = \nu I,$$

where $\tilde{J} := \tilde{X} \# \tilde{S}$. Recall the defining equation for \tilde{J} : $\tilde{X} - \tilde{J} \tilde{S}^{-1} \tilde{J} = 0$. Hence, if $U[H_x, H_s]$ denotes the derivative of \tilde{J} at (X, S) in the directions H_x for \tilde{X} and H_s for \tilde{S} , we have

$$H_x - JS^{-1}U[H_x, H_s] - U[H_x, H_s]S^{-1}J + JS^{-1}H_s S^{-1}J = 0.$$

But to get the Newton step, we want to choose $H_x = \Delta X$ and $H_s = \Delta S$ so that the resulting $U := U[H_x, H_s]$ satisfies

$$JU + UJ = \nu I - J^2.$$

This Lyapunov equation can be solved directly, to give $U = (\nu J^{-1} - J)/2$. Substituting this into the linking equation, we find that the corresponding direction satisfies (7) with

$$\mathcal{E} = I \odot I, \quad \mathcal{F} = K \odot K, \quad R_{EF} = \frac{\nu}{2} (KJ^{-1} + J^{-1}K^T) - X.$$

where we recall that $K := JS^{-1} = XJ^{-1}$. Note the similarities of this system to that defining the NT direction. Indeed, \mathcal{F} is defined by K rather than W , but we still have $KSK^T = JS^{-1}J = X = WSW$, and $(KJ^{-1} + J^{-1}K^T)/2 = (JS^{-1}J^{-1} + J^{-1}S^{-1}J)/2$ is the mean of two matrices similar to S^{-1} . However, there is a crucial property of the NT direction not possessed by the J direction.

Recall that almost all of our directions are motivated by taking a Newton step to satisfy the feasibility equations for the primal and dual together with some nonlinear equation. There is no *a priori* reason why we should expect that

$$(X + \Delta X) \bullet (S + \Delta S) = \nu n,$$

which certainly follows if the nonlinear equation $\tilde{X} \tilde{S} - \nu I$ is satisfied exactly by the next iterate. However, this does indeed hold for several directions, as long as the current iterate is feasible. In this case, r_p and R_d are zero, so the feasibility equations imply that $\Delta X \bullet \Delta S$ is zero. Then the equation above is equivalent to

$$S \bullet \Delta X + X \bullet \Delta S = \nu n - X \bullet S. \tag{33}$$

We shall say that a direction *predicts the duality gap* if it satisfies (33). For example, if we take the last equation defining the NT direction,

$$\Delta X + W \Delta S W = \nu S^{-1} - X,$$

and take its inner product with S , we obtain exactly (33). As we shall see, this same happy coincidence occurs for almost all our directions. However, taking the inner product of the last equation defining the J direction with S gives

$$S \bullet \Delta X + (K^T SK) \bullet \Delta S = \frac{\nu}{2} S \bullet (JS^{-1}J^{-1} + J^{-1}S^{-1}J) - X \bullet S,$$

which is not (33). A slight modification of the J direction restores this property:

J2. We merely interchange K and K^T in the equations defining the J direction. Hence we obtain the solution to (7) with

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

Our next direction replaces the metric geometric mean J by the spectral geometric mean:

V. This is the Newton step for

$$\tilde{V}^2 = \nu I, \tag{34}$$

where $\tilde{V} = \tilde{W}^{-1/2} \tilde{X} \tilde{W}^{-1/2} = \tilde{W}^{1/2} \tilde{S} \tilde{W}^{1/2}$. Let us write $V'[H_x, H_s]$ ($W'_{1/2}[H_x, H_s]$) for the derivative of \tilde{V} ($\tilde{W}^{1/2}$) in the direction (H_x, H_s) at (X, S) . Then we find

$$\begin{aligned} V'[H_x, H_s] &= -W^{-1/2} W'_{1/2}[H_x, H_s] W^{-1/2} X W^{-1/2} \\ &\quad + W^{-1/2} H_x W^{-1/2} - W^{-1/2} X W^{-1/2} W'_{1/2}[H_x, H_s] W^{-1/2}, \end{aligned}$$

and

$$\begin{aligned} V'[H_x, H_s] &= W'_{1/2}[H_x, H_s] S W^{1/2} + W^{1/2} H_s W^{1/2} \\ &\quad + W^{1/2} S W'_{1/2}[H_x, H_s]. \end{aligned}$$

We want to choose $(\Delta X, \Delta S) = (H_x, H_s)$ so that the resulting $V'[H_x, H_s]$ satisfies

$$V V'[H_x, H_s] + V'[H_x, H_s] V = \nu I - V^2,$$

to get the Newton step for (34). This Lyapunov system can be solved directly to give $V'[H_x, H_s] = (\nu V^{-1} - V)/2$. Hence, letting Z denote $W'_{1/2}[\Delta X, \Delta S]$, the Newton step satisfies the first two equations of (8) as well as

$$\begin{aligned} W^{-1/2} \Delta X W^{-1/2} - W^{-1/2} X W^{-1/2} Z W^{-1/2} - W^{-1/2} Z W^{-1/2} X W^{-1/2} &= \frac{1}{2}(\nu V^{-1} - V), \\ W^{1/2} \Delta S W^{1/2} + W^{1/2} S Z + Z S W^{1/2} &= \frac{1}{2}(\nu V^{-1} - V), \end{aligned}$$

or

$$\begin{aligned} \Delta X - X W^{-1/2} Z - Z W^{-1/2} X &= \frac{1}{2}(\nu S^{-1} - X) \\ \Delta S + S Z W^{-1/2} + W^{-1/2} Z S &= \frac{1}{2}(\nu X^{-1} - S), \end{aligned}$$

which is (8) with

$$\mathcal{E} := \mathcal{F} := I \odot I, \quad \mathcal{G} := X W^{-1/2} \odot I, \quad \mathcal{H} := S \odot W^{-1/2}$$

and

$$R_E := \frac{1}{2}(\nu S^{-1} - X), \quad R_F := \frac{1}{2}(\nu X^{-1} - S).$$

V2. We can obtain a variant of this direction by proceeding as in the first derivation of the MTI direction to eliminate the square roots. We treat $ZW^{-1/2}$ as a new variable, denoted Z in what follows, and ignore the fact that it is not symmetric to obtain the first two equations of (8) and

$$\begin{aligned} \Delta X & - XZ & - ZX & = \frac{1}{2}(\nu S^{-1} - X) \\ \Delta S & + SZ & + ZS & = \frac{1}{2}(\nu X^{-1} - S), \end{aligned}$$

so that

$$\mathcal{E} := \mathcal{F} := I \odot I, \quad \mathcal{G} := X \odot I, \quad \mathcal{H} := S \odot I,$$

with R_E and R_F as above. Note that this is exactly the system defining the MTI direction except that the nonzero right-hand side R_F there is now distributed between R_E and R_F .

SXS. The final direction we consider is the Newton step for

$$\tilde{S}\tilde{X}\tilde{S} = \nu\tilde{S}$$

(compare with the motivation for the H..K..M direction). It follows that the direction solves the first two equations of (7) together with

$$S\Delta X S + S X \Delta S + \Delta S X S - \nu \Delta S = \nu S - S X S,$$

or

$$\Delta X + X \Delta S S^{-1} + S^{-1} \Delta S X - \nu S^{-1} \Delta S S^{-1} = \nu S^{-1} - X,$$

which is (7) with

$$\mathcal{E} := I \odot I, \quad \mathcal{F} := S^{-1} \odot (2X - \nu S^{-1}), \quad R_{EF} := \nu S^{-1} - X.$$

Before ending the section, we consider a modification of the SXS direction. Suppose we take the Newton step for

$$\tilde{S}\tilde{X}\tilde{S} = \nu S.$$

(We can view this as perturbing the right-hand side of the linear system defining the Newton step for $SXS = 0$, rather than first perturbing the nonlinear system and then computing the Newton step.) Then we obtain the direction solving (7) with

$$\mathcal{E} := I \odot I, \quad \mathcal{F} := 2S^{-1} \odot X, \quad R_{EF} := \nu S^{-1} - X. \quad (35)$$

Except for the factor 2, this is exactly the system defining the H..K..M direction. Suppose the current iterate is dual feasible, so that $R_D = 0$. Then the solution to this new system is $(\Delta X, \Delta y/2, \Delta S/2)$, where $(\Delta X, \Delta y, \Delta S)$ is the H..K..M direction. Hence any iterative method that combines this direction with a line search that takes a fraction of the step to the boundary (separately for the primal and the dual) will generate exactly the same iterates as if it used the H..K..M direction. In this way,

we can view the H..K..M method as a perturbed Newton method for $SXS = 0$, with nonstandard step-size control. The situation can be compared to the solution of the scalar equation $s^2 = 0$, which has a singular Jacobian at the solution. The Newton iteration leads to linear convergence, while doubling the Newton step (corresponding to replacing the direction given by (35) with the H..K..M direction) leads to the exact solution in one step.

5 Computation of the search directions

In all cases, we will compute the search direction $(\Delta X, \Delta y, \Delta S)$ through (10) or (11). This involves obtaining the $m \times m$ matrix $B := \mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$ or $B' := \mathcal{A}\mathcal{E}^{-1}\mathcal{G}\mathcal{H}^{-1}\mathcal{F}\mathcal{A}^*$ and evaluating the right-hand sides. Each entry of B or B' is calculated from

$$B_{ij} = A_i \bullet (\mathcal{E}^{-1}\mathcal{F}A_j) \quad \text{or} \quad B'_{ij} = A_i \bullet (\mathcal{E}^{-1}\mathcal{G}\mathcal{H}^{-1}\mathcal{F}A_j), \quad (36)$$

and since application of \mathcal{F} or \mathcal{G} to a matrix is straightforward for all methods, it suffices to show how to solve

$$\mathcal{E}U = R \quad \text{and} \quad \mathcal{H}U = R \quad (37)$$

for arbitrary $R \in S\mathbb{R}^{n \times n}$ and those \mathcal{E} and \mathcal{H} arising in our methods.

The matrices required to define \mathcal{E} , \mathcal{F} , \mathcal{G} , and \mathcal{H} are easily computed. Inverses, such as X^{-1} in the Primal direction, are best evaluated using Cholesky factorization, while square roots, such as $X^{1/2}$ in the Half direction, require an eigenvalue decomposition. Geometric means, like W , V , and J , require Cholesky factors and one or two eigenvalue or singular value decompositions; see [31] for an efficient way to compute W (or similarly J). All of these calculations require $O(n^3)$ arithmetic operations, and in the dense case will take negligible time compared to the computation of B or B' (which requires $O((m+n)mn^2)$ arithmetic operations).

Let us therefore turn to the solution of systems of the form (37) arising in our methods. For eight of them (H..K..M, NT, Primal, Half, IHalf, J, J2, and SXS), \mathcal{H} does not appear and \mathcal{E} can be taken to be the identity, so no such systems need to be solved. For three of them (AHO, MTI, and V2), either \mathcal{H} does not appear and \mathcal{E} is of the form $S \odot I$, or \mathcal{E} is the identity and $\mathcal{H} = S \odot I$. For S-MT, \mathcal{E} is the identity and $\mathcal{H} = S^{1/2} \odot I$. For the remaining directions (dual H..K..M, Gu, Toh, X-MT, MTJ, MTW, and V), either \mathcal{H} does not appear and \mathcal{E} is of the form $S \odot M$ for some symmetric positive definite M , or \mathcal{E} is the identity and $\mathcal{H} = S \odot M$ for some such M . (The attentive reader will have noticed that the S-Chol-MT direction is missing from this list. This requires the solution of a triangular Lyapunov system (see [24]) and will not be addressed further in this section; the order of complexity is the same as the methods here, with slightly smaller constants.) Let us consider these cases in turn.

For the first nontrivial case, we need to solve $(S \odot I)U = R$, or equivalently

$$SU + US = 2R.$$

Let us compute the eigenvalue decomposition QDQ^T of S , where Q is orthogonal and $D = \text{diag}(d_1, \dots, d_n)$ with all d_j 's positive. Then, with $\bar{U} := Q^T U Q$ and $\bar{R} := Q^T R Q$,

we get

$$D\bar{U} + \bar{U}D = 2\bar{R},$$

so that $\bar{U}_{ij} = 2\bar{R}_{ij}/(d_i + d_j)$ for all i, j . Hence we calculate \bar{R} , thence \bar{U} , and lastly U . A similar technique works for the next case, since $S^{1/2} = QD^{1/2}Q^T$.

Finally we need to find the solution of

$$MUS + SUM = 2R,$$

where M is a symmetric positive definite matrix. We first compute an eigenvalue decomposition of S to obtain $S^{1/2}$ and $S^{-1/2}$, and then rewrite the equation above as

$$\hat{M}\hat{U} + \hat{U}\hat{M} = 2\hat{R},$$

where $\hat{M} := S^{-1/2}MS^{-1/2}$, $\hat{U} := S^{1/2}US^{1/2}$, and $\hat{R} := S^{-1/2}RS^{-1/2}$. Now we proceed as in the paragraph above. We find the eigenvalue decomposition $\hat{Q}\hat{D}\hat{Q}^T$ of \hat{M} , where \hat{Q} is orthogonal and $\hat{D} = \text{diag}(\hat{d}_1, \dots, \hat{d}_n)$ with all \hat{d}_j 's positive. Then, with $\bar{U} := \hat{Q}^T\hat{U}\hat{Q}$ and $\bar{R} := \hat{Q}^T\hat{R}\hat{Q}$, we get

$$\hat{D}\bar{U} + \bar{U}\hat{D} = 2\bar{R},$$

so that $\bar{U}_{ij} = 2\bar{R}_{ij}/(\hat{d}_i + \hat{d}_j)$ for all i, j . Hence we obtain \hat{U} and thence U . Note that, with $P := S^{-1/2}Q$, $\bar{R} = P^T R P$ and $U = P\bar{U}P^T$.

6 Properties of SDP search directions

Here we list a number of desirable properties for search directions in semidefinite programming to enjoy, and check whether the directions of Section 3 possess these properties or not. It is somewhat laborious to check each direction for each property, and so our proofs often consider just one or two representative cases.

When we state that a property is true for all our directions except A, B, and C, we mean that all the directions in Section 3 except for these possess the property, and the exceptional directions do not; the counterexamples in such cases are easy to find using MATLAB [18], and indeed we found that one example sufficed for all cases. If the situation for the exceptional directions is unclear (counterexamples have not been found), we write that the property holds for all our directions except possibly A, B, and C. Similarly, if we state that a property holds for just the A, B, and C directions, we mean to imply that it fails for all others in Section 3; if this is not known, we state that the property holds for at least the A, B, and C directions.

Extends linear programming. We mentioned in the introduction that, when the space $SI\mathbb{R}^{n \times n}$ is replaced by the subspace of diagonal matrices, the resulting problems are merely reformulations of linear programming problems. In this case, we would expect that the search directions found for the corresponding semidefinite programming problems would be the diagonal matrices defined by the usual primal-dual directions of linear programming, i.e., would satisfy the first two equations of (7) and

$$S\Delta X + X\Delta S = \nu I - XS, \tag{38}$$

with all matrices diagonal. Hence we say that a direction possesses property (ELP) if, whenever all A_i 's C , X , and S are diagonal, the resulting search direction satisfies (38).

Proposition 6.1 *Property (ELP) holds for all our directions except Primal and SXS.*

Proof:

Let us prove this for all our directions in the MT family. Then, for some symmetric positive definite N , the direction satisfies (28). But all choices for N corresponding to directions in Section 3 are diagonal if X and S are. Hence the auxiliary matrix Z is also diagonal, as the solution of a Lyapunov equation with diagonal matrices on the left- and right-hand sides. Now by multiplying the first equation of (28) by S and the second by X and adding, we get exactly (38), since the terms in Z cancel. The key here, of course, is that diagonal matrices commute. \square

Predicts duality gap. Suppose our current iterate (X, y, S) is primal and dual feasible. If we succeed in taking a full step, the next iterate will be $(X + \Delta X, y + \Delta y, S + \Delta S)$, and this will also be primal and dual feasible, with duality gap equal to

$$(X + \Delta X) \bullet (S + \Delta S) = X \bullet S + S \bullet \Delta X + X \bullet \Delta S,$$

since the directions ΔX and ΔS are orthogonal. We would like this to be equal to the duality gap of the point aimed at, i.e., of the point on the central path with parameter ν . Hence we say that a direction satisfies property (PDG) if we have

$$S \bullet \Delta X + X \bullet \Delta S = \nu n - X \bullet S. \quad (39)$$

Proposition 6.2 *Property (PDG) holds for all our directions except Primal, IHalf, J, and SXS.*

Proof:

Again, we can check this straightforwardly for each case, but instead we will use a more general technique.

In linear programming, the matrix $\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^*$ or $\mathcal{A}\mathcal{E}^{-1}\mathcal{G}\mathcal{H}^{-1}\mathcal{F}\mathcal{A}^*$ is replaced by $\mathcal{A}X\mathcal{S}^{-1}\mathcal{A}^T$, so we might expect $\mathcal{E}^{-1}\mathcal{F}$ or $\mathcal{E}^{-1}\mathcal{G}\mathcal{H}^{-1}\mathcal{F}$ to act like $X\mathcal{S}^{-1}$. It is indeed the case that $\mathcal{E}^{-1}\mathcal{F}\mathcal{S} = X$ or $\mathcal{E}^{-1}\mathcal{G}\mathcal{H}^{-1}\mathcal{F}\mathcal{S} = X$ for most of our directions, but it appears that a more fundamental property is that $\mathcal{F}^*\mathcal{E}^{-*}\mathcal{S} = X$ or $\mathcal{F}^*\mathcal{H}^{-*}\mathcal{G}^*\mathcal{E}^{-*}\mathcal{S} = X$. We say that a direction based on (7) has property (PDG') if

$$\mathcal{F}^*\mathcal{E}^{-*}\mathcal{S} = X, \quad (\mathcal{E}^{-*}\mathcal{S}) \bullet R_{EF} = \nu n - X \bullet \mathcal{S};$$

similarly, a direction based on (8) has property (PDG'') if

$$\mathcal{F} \text{ is nonsingular, } \mathcal{G}^*\mathcal{E}^{-*}\mathcal{S} = \mathcal{H}^*\mathcal{F}^{-*}X, \quad (\mathcal{E}^{-*}\mathcal{S}) \bullet R_E + (\mathcal{F}^{-*}X) \bullet R_F = \nu n - X \bullet \mathcal{S}.$$

By taking the inner product of the last equation of (7) with $\mathcal{E}^{-*}\mathcal{S}$, we find that a direction possessing (PDG') also enjoys (PDG). If we instead take the inner product of the third equation of (8) with $\mathcal{E}^{-*}\mathcal{S}$ and of the fourth equation with $\mathcal{F}^{-*}X$ and add, we find that a direction with (PDG'') also possesses (PDG).

Now we are ready to confirm property (PDG) for the entire MZ family. Here \mathcal{E} has the form $S \odot M$ for some symmetric positive definite M , so that \mathcal{E} is self-adjoint and $\mathcal{E}^{-*}S = M^{-1}$. Also, \mathcal{F} has the form $MX \odot I$, so $\mathcal{F}^* = XM \odot I$ and $\mathcal{F}^*M^{-1} = X$ as desired. Also, $(\mathcal{E}^{-*}S) \bullet R_{EF} = M^{-1} \bullet (\nu M - (MXS + SXM)/2) = \nu n - X \bullet S$, which confirms (PDG') and hence (PDG).

Next we consider the MT family. Here $\mathcal{E} = \mathcal{F} = I \odot I$, $\mathcal{G} = XN \odot I$, and $\mathcal{H} = S \odot N$ for some symmetric positive definite N . We find $\mathcal{G}^*\mathcal{E}^{-*}S = (NX \odot I)S = (NXS + SXN)/2$, while $\mathcal{H}^*\mathcal{F}^{-*}X = (S \odot N)X = (SXN + NXS)/2$, so the two are equal. Also, $(\mathcal{E}^{-*}S) \bullet R_E + (\mathcal{F}^{-*}X) \bullet R_F = X \bullet R_F = X \bullet (\nu X^{-1} - S) = \nu n - X \bullet S$, so we have checked (PDG'') and hence (PDG).

Similar arguments confirm the property for all other directions except Primal, IHalf, J, SXS, and S-Chol-MT. But we can check (PDG) for this last direction by taking the inner product of the first equation of (31) with I and the second with X and adding. (Note that while (PDG) fails for the IHalf and J directions, it holds for the Half and J2 directions.) \square

Well-defined directions. We would like our method for computing directions to give a unique search direction for every symmetric positive definite X and S and surjective operator \mathcal{A} . Given that we define our directions using (7) or (8) with \mathcal{E} or \mathcal{E} and \mathcal{H} nonsingular, this requires that, for all symmetric positive definite X and S ,

$$\mathcal{A}\mathcal{E}^{-1}\mathcal{F}\mathcal{A}^* \text{ or } \mathcal{A}\mathcal{E}^{-1}\mathcal{G}\mathcal{H}^{-1}\mathcal{F}\mathcal{A}^* \text{ is nonsingular for all surjective } \mathcal{A}.$$

It is easy to see that this statement holds iff

$$\mathcal{E}^{-1}\mathcal{F} \text{ or } \mathcal{E}^{-1}\mathcal{G}\mathcal{H}^{-1}\mathcal{F} \text{ is positive (or negative) definite,}$$

but not necessarily self-adjoint (see also Tseng [35]). We say that a direction satisfies (WDD) if the latter holds for all symmetric positive definite X and S . Below we determine for which of our directions (WDD) holds. Shida, Shindoh, and Kojima [30] provide a nice approach to this question for the general KSH family, using the concepts of maximal monotone and antitone affine subspaces of $S\mathbb{R}^{n \times n} \times S\mathbb{R}^{n \times n}$, and also determine via this technique that the H..K..M, NT, and Primal directions satisfy (WDD) as does the AHO direction if $XS + SX$ is symmetric positive definite. Our approach relies on an analysis of the Schur complement matrix as above. Todd, Toh, and Tütüncü [31] show that $\mathcal{E}^{-1}\mathcal{F}$ is self-adjoint and positive definite whenever it is self-adjoint, so that all members of the MZ* or TTT family enjoy the (WDD) property. Monteiro and Tsuchiya [21] show indirectly that all members of the MT* family give well-defined directions, so that (WDD) holds. This family includes the H..K..M, dual H..K..M, and NT directions, which are also in the MZ* family, but also the MTW direction. Since their method was indirect, it was unclear whether the corresponding operator $\mathcal{E}^{-1}\mathcal{G}\mathcal{H}^{-1}\mathcal{F}$ in the last case was self-adjoint as well as positive definite; indeed, from the form of \mathcal{G} and \mathcal{H} it seemed unlikely. The fact that the MTW operator does have this property was surprising, and only discovered after numerical tests showed that the Schur matrix was always symmetric. Thus in all cases where (WDD) is known to hold, the Schur complement matrix is symmetric positive definite.

Proposition 6.3 *Property (WDD) holds for just the H..K..M, dual H..K..M, NT, Gu, Toh, MTW, and Primal directions.*

Proof:

For the H..K..M direction, $\mathcal{E}^{-1}\mathcal{F} = \mathcal{F} = X \odot S^{-1}$, which is self-adjoint and positive definite because X and S^{-1} are symmetric positive definite. For the dual H..K..M direction, $\mathcal{E}^{-1}\mathcal{F} = \mathcal{E}^{-1} = (S \odot X^{-1})^{-1}$ is self-adjoint and positive definite because its inverse is, for reasons identical to those above. Clearly the NT direction has property (WDD), since now $\mathcal{E}^{-1}\mathcal{F} = W \odot W$ which is self-adjoint and positive definite. The same is true for the Primal direction, with $\mathcal{E}^{-1}\mathcal{F} = (1/\nu)X \odot X$. Direct arguments are harder for the other directions.

To take care of the Gu and Toh directions, we note that by results of Gu [7] they belong to the MZ* or TTT family, for which $\mathcal{E}^{-1}\mathcal{F}$ is self-adjoint and positive definite by results of [31]. This class also contains the three directions considered above. Thus we are left with just the MTW direction. This case follows from the lemma below. \square

Lemma 6.1 *For the MTW direction, the operator $\mathcal{E}^{-1}\mathcal{G}\mathcal{H}^{-1}\mathcal{F}$ is self-adjoint and positive definite.*

Proof:

We give two proofs for this result. The first, simple argument is due to Shida [29], based on ideas in [30]. The second, more involved proof also shows how to compute $\mathcal{G}\mathcal{H}^{-1}R$ (or $\mathcal{E}^{-1}\mathcal{G}\mathcal{H}^{-1}\mathcal{F}R$) efficiently for any R , as required to find the MTW direction; indeed, this turns out to be barely more expensive than computing just $\mathcal{H}^{-1}R$.

Recall that for this direction, $\mathcal{E} = \mathcal{F} = I \odot I$, $\mathcal{G} = XN \odot I$, and $\mathcal{H} = S \odot N$, with $N = X^{-1}\#W^{-1}$. We therefore have

$$NXN = W^{-1}, \quad NXNXNXN = S, \quad (40)$$

from which we deduce that

$$XNS^{-1} = N^{-1}W^{-1}S^{-1} = N^{-1}X^{-1}W = WNW \quad (41)$$

is symmetric positive definite.

Now $\mathcal{E}^{-1}\mathcal{G}\mathcal{H}^{-1}\mathcal{F} = \mathcal{G}\mathcal{H}^{-1}$ is self-adjoint and positive definite if and only if $\mathcal{H}\mathcal{G} = \mathcal{H}\mathcal{G}\mathcal{H}^{-1}\mathcal{H}$ is. But

$$\begin{aligned} \mathcal{H}\mathcal{G} &= (S \odot N)(XN \odot I) = \frac{1}{2}(SXN \odot N + S \odot NXN) \\ &= \frac{1}{2}(SXNS^{-1}S \odot N + S \odot NXN) \\ &= \frac{1}{2}(SWNWS \odot N + S \odot W^{-1}) \end{aligned}$$

using (41) and (40). This is clearly self-adjoint and positive definite.

For the second proof we consider $\mathcal{G}\mathcal{H}^{-1}R$ for an arbitrary $R \in S\mathbb{R}^{n \times n}$. From the previous section, to obtain $U_1 := \mathcal{H}^{-1}R$, we find the eigenvalue decomposition

$\hat{Q}\hat{D}\hat{Q}^T$ of $\hat{N} := S^{-1/2}NS^{-1/2}$ (where $\hat{D} = \text{diag}(\hat{d}_1, \dots, \hat{d}_n)$ with all \hat{d}_j 's positive), set $P := S^{-1/2}\hat{Q}$ and $\bar{R} = P^T RP$, and then

$$U_1 = P\bar{U}P^T, \quad \text{where } \bar{U}_{ij} = 2\bar{R}_{ij}/(\hat{d}_i + \hat{d}_j) \quad (42)$$

for all i, j . Next we need

$$U_2 := \mathcal{G}U_1 = \frac{1}{2}(XNU_1 + U_1NX) = \frac{1}{2}(P((P^{-1}XNP)\bar{U} + \bar{U}(P^T NXP^{-T}))P^T). \quad (43)$$

Let us now examine

$$P^{-1}XNP = \hat{Q}^T S^{1/2} X N S^{-1/2} \hat{Q} = \hat{Q}^T S^{1/2} X N S^{-1} S^{1/2} \hat{Q} = \hat{Q}^T S^{1/2} W N W S^{1/2} \hat{Q}.$$

Now, using in turn $WSW = X$, (40), and (41), we obtain

$$\begin{aligned} (S^{1/2}WNWS^{1/2})^3 &= S^{1/2}WNWSWNWSWNWS^{1/2} \\ &= S^{1/2}WNXNXNWS^{1/2} \\ &= S^{1/2}WSN^{-1}X^{-1}WS^{1/2} \\ &= S^{1/2}N^{-1}S^{1/2} \\ &= \hat{N}^{-1} = \hat{Q}\hat{D}^{-1}\hat{Q}^T. \end{aligned}$$

It follows that $P^{-1}XNP = \hat{D}^{-1/3}$, so that

$$\mathcal{G}\mathcal{H}^{-1} = (P \odot P)\mathcal{D}(P^T \odot P^T),$$

with \mathcal{D} the positive definite diagonal operator defined by

$$(\mathcal{D}U)_{ij} = U_{ij}(\hat{d}_i^{-1/3} + \hat{d}_j^{-1/3})/(\hat{d}_i + \hat{d}_j)$$

for all i, j . This is clearly self-adjoint and positive definite, concluding the proof. \square

We should add that for the AHO direction and indeed for all directions in the MT and MZ families, it is known that the search directions are well-defined whenever the iterate (X, y, S) lies in a wide neighborhood of the central path; see [30, 20, 21, 23]. In addition, no problems have been reported in practical computation. Nevertheless, we find it somewhat disturbing that a method may fail to produce a search direction at a strictly feasible iterate.

Let us also record here the status of the Schur complement matrix for all our directions.

Proposition 6.4 *The Schur complement matrix is always symmetric for just the H..K..M, dual H..K..M, NT, Gu, Toh, MTW, Primal, and SXS directions. It is also positive definite for all but the last, which yields a symmetric positive definite Schur complement matrix for all sufficiently small ν (below some tolerance depending on X and S).*

□

Bounded system. Our directions are defined by systems of linear equations of the form (7) or (8) for suitable \mathcal{E} and \mathcal{F} or \mathcal{E} through \mathcal{H} . However, some of these systems are not defined at a solution to the (SDP) and its dual, where X and S are only positive semidefinite. In particular, W and J are not defined at such solutions. However, instead of making a strong condition that the system be defined at such a solution, we ask instead whether the systems (or equivalent restatements of them) remain bounded as X and S converge to positive semidefinite matrices. Clearly this is true for those defined at the solution, such as the AHO and Half directions, but it is also true for the NT, Gu, and other directions.

Proposition 6.5 *The bounded system property holds for at least the AHO, H..K..M, dual H..K..M, NT, Gu, X-MT, MTI, S-MT, and Half directions.*

Proof:

This is clear for the AHO and Half directions; it is also apparent for the H..K..M direction if we employ the first definition of \mathcal{E} and \mathcal{F} , and thus in a similar way for the dual H..K..M direction. For the NT direction, we use the last mentioned choice of \mathcal{E} and \mathcal{F} . Note that $W^{-1/2}VW^{-1/2} = S$, so $V^{1/2}W^{-1/2} = Q'_s S^{1/2}$ for some orthogonal Q'_s , and similarly $V^{1/2}W^{1/2} = Q'_x X^{1/2}$ for some orthogonal Q'_x . Now $\mathcal{E}U = Q'_s S^{1/2} U S^{1/2} (Q'_s)^T$, whose Frobenius norm is at most $\|S\|_F$ times that of U , and similarly for $\mathcal{F}U$. For the Gu direction, recall that P has rows of unit norm, so the entries of M are bounded by n . For the X-MT direction, we eliminate the $X^{-1/2}$ in \mathcal{H} by pre- and post-multiplying by $X^{1/2}$. The proofs for the other directions are similar.

□

Alizadeh et al. [3] show that not only is the linear system for the AHO direction defined at the solution, but that if strict complementarity and primal and dual nondegeneracy hold there, then the coefficient matrix is nonsingular. Numerical tests suggest that this fails for all others among our directions. One might hope that the nonsingular result holds for the Half direction, since its linear system is defined at the solution and is similar. However, this is not the case. The argument below assumes familiarity with [3]. Following the reasoning there, we find that the diagonal matrix with entries $\lambda_i + \lambda_j$ that arises in the AHO direction is replaced by one with entries $\sqrt{\lambda_i \lambda_j}$ for the Half direction, and similarly for the diagonal matrix depending on the ω_i 's. It follows that the matrix $S^T J S$ in [3] will generally be singular for the Half direction at the solution, since the diagonal matrices Φ and Γ do not have enough positive diagonal entries between them.

Scale invariance. We discussed P - and Q -scale invariance in Section 3. Roughly, P -scale invariance (PSI) holds if a direction is invariant under arbitrary congruence transformations on X (and corresponding transformations on the data and S), while Q -scale invariance (QSI) requires only invariance under orthogonal similarity transformations. Here we determine the scale invariance properties of our twenty directions.

Let us first note that it is possible to construct a scale-invariant *method* that uses a *direction* that is *not* scale-invariant, based on a suitable initialization. Indeed, let

us suppose we have some technique for producing an initial iterate (X_0, y_0, S_0) for our problem (SDP) such that $X_0 S_0 = \mu_0 I$ for some $\mu_0 > 0$ (the iterate need not be feasible). Suppose also that the initialization technique is scale-invariant, so that given (\widehat{SDP}) the method would return $(PX_0 P^T, y_0, P^{-T} S_0 P^{-1})$. Then a scale-invariant method can be devised by applying the initialization technique to get (X_0, y_0, S_0) , defining $P := \mu_0^{1/4} X^{-1/2}$, and then applying the basic method (using the direction that may not be scale-invariant) to the problem (\widehat{SDP}) scaled by this P , starting at the scaled point $(PX_0 P^T, y_0, P^{-T} S_0 P^{-1}) = (\mu_0^{1/2} I, y_0, \mu_0^{1/2} I)$. However, we find it more pleasing to use a direction that is itself scale-invariant.

Proposition 6.6 *Property (PSI) holds for just the H..K..M, dual H..K..M, NT, MTW, Primal, and SXS directions.*

Proof:

Let us prove this for the H..K..M and MTW directions. For the first, we need to show that, if $(\Delta X, \Delta y, \Delta S)$ is the H..K..M direction for the original problem (SDP) at (X, y, S) (or satisfies the equations defining this direction, which is the same since this direction is well-defined), then $(\Delta \hat{X}, \Delta \hat{y}, \Delta \hat{S}) := (P \Delta X P^T, \Delta y, P^{-T} \Delta S P^{-1})$ is the H..K..M direction (or equivalently satisfies the defining equations) for the scaled problem (\widehat{SDP}) at $(\hat{X}, \hat{y}, \hat{S}) = (P X P^T, y, P^{-T} S P^{-1})$. Clearly this direction satisfies the first two equations of the scaled (7), and since

$$\Delta X + \frac{1}{2}(X \Delta S S^{-1} + S^{-1} \Delta S X) = \nu S^{-1} - X,$$

we have

$$\begin{aligned} P \Delta X P^T + \frac{1}{2}((P X P^T)(P^{-T} \Delta S P^{-1})(P^{-T} S P^{-1})^{-1} + (P^{-T} S P^{-1})^{-1}(P^{-T} \Delta S P^{-1})(P X P^T)) \\ = \nu (P^{-T} S P^{-1})^{-1} - (P X P^T), \end{aligned}$$

which is the third equation of the scaled (7). (A similar but more general argument for an arbitrary member of the MZ family defined by $M = M(X, S)$ shows that (PSI) holds if M transforms like S , i.e., if $M(\hat{X}, \hat{S}) = P^{-T} M P^{-1}$; this holds for the H..K..M, dual H..K..M, and NT directions.)

Turning now to the MTW direction, we know that this satisfies the first two equations of (8) as well as

$$\begin{aligned} \Delta X - X N Z - Z N X &= 0, \\ \Delta S + S Z N + N Z S &= \nu X^{-1} - S, \end{aligned}$$

where $N := X^{-1} \# W^{-1}$. Hence, by pre- and post-multiplying by P and P^T (first equation) and by P^{-T} and P^{-1} (second equation), and writing $(\Delta \hat{X}, \Delta \hat{y}, \Delta \hat{S}) := (P \Delta X P^T, \Delta y, P^{-T} \Delta S P^{-1})$ and $(\hat{X}, \hat{y}, \hat{S}) = (P X P^T, y, P^{-T} S P^{-1})$, we obtain

$$\begin{aligned} \Delta \hat{X} - \hat{X} \hat{N}' \hat{Z} - \hat{Z} \hat{N}' \hat{X} &= 0, \\ \Delta \hat{S} + \hat{S} \hat{Z} \hat{N}' + \hat{N}' \hat{Z} \hat{S} &= \nu \hat{X}^{-1} - \hat{S}, \end{aligned}$$

where $\hat{N}' := P^{-T}NP^{-1}$ and $\hat{Z} := PZP^T$. These are the appropriate third and fourth equations of (8) for the scaled problem, with auxiliary variable \hat{Z} , as long as \hat{N}' is the appropriate N matrix for the scaled problem, i.e., as long as

$$\hat{N}' = \hat{N} := N(\hat{X}, \hat{S}) := \hat{X}^{-1} \# \hat{W}^{-1}, \quad (44)$$

with \hat{W} the scaling matrix for the scaled iterates \hat{X} and \hat{S} . (That is, again we want N to transform like S .) The defining equation for \hat{W} is $\hat{W}\hat{S}\hat{W} = \hat{X}$, so that $\hat{W} = PWP^T$. Then, since $NWN = X^{-1}$, we find by pre- and post-multiplying by P^{-T} and P^{-1} that

$$\hat{N}'\hat{W}\hat{N}' = \hat{X}^{-1},$$

which shows that (44) is satisfied and completes the proof. \square

Let us illustrate what goes wrong for the AHO direction. The key point is that the third equation of (7) is now

$$\frac{1}{2}(\Delta XS + S\Delta X) + \frac{1}{2}(X\Delta S + \Delta SX) = \nu I - \frac{1}{2}(XS + SX),$$

and under our scaling, XS transforms to $\hat{X}\hat{S} = PXSP^{-1}$ while SX transforms to $\hat{S}\hat{X} = P^{-T}SX P^T$, and similarly for the terms on the left-hand side, and these do not combine gracefully. They do in the case that $P = P^{-T}$, i.e., when P is orthogonal. Indeed we have

Proposition 6.7 *Property (QSI) holds for all our directions except S-Chol-MT.*

Proof:

Let us check this for our directions in the MZ and MT classes. The others follow using similar arguments.

Each of our MZ directions satisfies the first two equations of (7) and

$$S\Delta XM + M\Delta XS + MX\Delta S + \Delta SXM = 2\nu M - MXS - SXM.$$

Hence the scaled direction $(\Delta\hat{X}, \Delta\hat{y}, \Delta\hat{S}) := (Q\Delta X Q^T, \Delta y, Q\Delta S Q^T)$ satisfies the first two equations of the scaled (7) at the scaled point $(\hat{X}, \hat{y}, \hat{S}) = (QXQ^T, y, QSQ^T)$ as well as

$$\hat{S}\Delta\hat{X}\hat{M} + \hat{M}\Delta\hat{X}\hat{S} + \hat{M}\hat{X}\Delta\hat{S} + \Delta\hat{S}\hat{X}\hat{M} = 2\nu\hat{M} - \hat{M}\hat{X}\hat{S} - \hat{S}\hat{X}\hat{M},$$

where $\hat{M} := QMQ^T$. Hence it is only necessary to check that, for our choices of $M = M(X, S)$, \hat{M} equals $M(\hat{X}, \hat{S})$. This certainly holds for $M = I$ (so that $\hat{M} = QQ^T = I$), $M = S$, $M = X^{-1}$, and $M = W^{-1}$, since as we saw in the proof above, $\hat{W} = QWQ^T$ so that $\hat{W}^{-1} = QW^{-1}Q^T$. This takes care of the AHO, H..K..M, dual H..K..M, and NT directions. For the Gu and Toh directions, note that the scaled square roots $\hat{X}^{1/2}$ and $\hat{S}^{1/2}$ are just $QX^{1/2}Q^T$ and $QS^{1/2}Q^T$ respectively. It follows that for the scaled problems, Q_x and Q_s are premultiplied by Q and Σ is unchanged. Hence Φ and Ψ are unchanged, and then it is easy to see that in both cases, \hat{M} is the appropriate M for the scaled problems.

For our MT directions, it is similarly only necessary to check that, for our choices of $N = N(X, S)$, $\hat{N} := QNQ^T$ equals $N(\hat{X}, \hat{S})$. This clearly holds for $N = I$, but it is also easily seen to hold for $N = X^{-1/2}$, $N = J^{-1}$, and $N = X^{-1}\#W^{-1}$. The proof for the S-MT direction is similar. \square

Coincidence of directions. We have already noted that all our directions are different in general, but that for problems arising from linear programming, all but two give the same direction. It is natural to ask if there are other situations where some of the directions coincide, in particular when the iterates lie on the central path.

First we consider a slightly more general case, where our iterates X and S commute. In this case, by making an orthogonal scaling, we can assume that they are both diagonal. Note that we are not assuming that all the A_i 's are diagonal also.

If X and S are diagonal, so are X^{-1} , $S^{1/2}$, etc., W , J , and K . Indeed, $W = K = X^{1/2}S^{-1/2}$. It follows that the NT, Half, IHalf, J, and J2 directions all coincide in this case.

Next, it is easily seen that the orthogonal matrices Q_x and Q_s that arise in the Gu and Toh directions reduce to the identity with $\Sigma = X^{1/2}S^{1/2}$. Hence $\Phi = S^{1/2}$ and $\Psi = X^{1/2}$. Thus, for both the Gu and Toh directions, the matrix M reduces to the identity, so both yield the AHO direction. Let us now consider the MTI direction. One of the equations in its linear system, $\Delta X - XZ - ZX = 0$, can be readily solved, to give $Z_{ij} = \Delta X_{ij}/(X_{ii} + X_{jj})$. Substituting this into the other equation and multiplying the ij th term by $(X_{ii} + X_{jj})/2$ yields the equation for the AHO direction. A similar analysis holds for the V2 direction.

Hence we have

Proposition 6.8 *If X and S commute, the AHO, Gu, Toh, MTI, and V2 directions coincide, as do the NT, Half, IHalf, J, and J2 directions. These two directions are different, and all our remaining directions are also distinct from each other and different from these two in general.*

Proof:

It is only necessary to point out that the directions under discussion are all Q -scale-invariant, so the reduction to the diagonal case above is valid. The other directions can be shown to be different using simple examples. \square

Next we specialize further to the case that we are on the central path, so that $X = \mu S^{-1}$ for some positive μ . In this case it is easily seen that $W = X^{1/2}S^{-1/2} = \mu^{1/2}S^{-1} = \mu^{-1/2}X$. Then the last equation satisfied by the AHO direction, $S\Delta X + \Delta SX + \Delta XS + X\Delta S = 2\nu I - XS - SX$ can be rewritten

$$S(\Delta X + W\Delta SW) + (\Delta X + W\Delta SW)S = 2\nu I - XS - SX,$$

which holds iff $\Delta X + W\Delta SW = \nu S^{-1} - X$, so the direction coincides with the NT direction. This proves that the two classes of five directions each collapse to a single direction on the central path. Similar arguments show

Proposition 6.9 *If the current iterate is on the central path, all our directions except possibly the Primal and SXS directions coincide; if further the iterates are primal and*

dual feasible, then the Primal and SXS directions are scalar multiples (possibly different for the primal and dual parts) of all our other directions. In case the current iterate is on the central path and ν is equal to $X \bullet S$ divided by n so that $X = \nu S^{-1}$, then all our directions coincide, and if further the iterates are primal and dual feasible, then all our directions equal the zero direction.

□

Primal-dual symmetry. We have already defined this property: a direction is primal-dual symmetric (PDS) if it is equal to its dual. We have

Proposition 6.10 *Property (PDS) holds for just the AHO, NT, Toh, Half, IHalf, J, V, and V2 directions.*

Proof:

This is clear from the corresponding linear systems for the AHO, Half, IHalf, and V2 directions. It is easily seen to hold for members of the MZ family if, whenever M results from X and S (in that order), M^{-1} results from S and X . Hence the NT direction is primal-dual symmetric, since it corresponds to $M = W^{-1}$, and the scaling matrix corresponding to S and X is W^{-1} . (We can also argue directly from the simpler linear system with $\mathcal{E} = I \odot I$, $\mathcal{F} = W \odot W$.) From the form for $M = P^T P$ in the Gu and Toh directions, we can see that primal-dual symmetry will hold if \hat{B} for S and X is the inverse of \hat{B} for X and S , and this holds for the Toh direction. (Σ is the same for S and X as for X and S .)

For the J direction, we note that J is unchanged when X and S are interchanged, while K is replaced by its inverse. Primal-dual symmetry then follows by pre- and post-multiplying the equation defining the direction by K^{-1} and K^{-T} respectively.

Finally, for the V direction, notice that if we replace the auxiliary variable Z by $W^{-1/2} Z W^{-1/2}$, then we can change \mathcal{G} to $X \odot W^{1/2}$ and \mathcal{H} to $SW^{1/2} \odot I$. Then our equations become just those for the dual direction. □

From the criteria so far, one direction stands out. Only the NT direction extends linear programming, predicts the duality gap, gives a well-defined direction, is P - as well as Q -scale-invariant, and is primal-dual symmetric. Moreover, the corresponding Schur complement matrix is symmetric and positive definite, and (at least in one form) the direction enjoys the bounded system property.

Four directions satisfy all but one property: the H..K..M, dual H..K..M, and MTW directions satisfy all but primal-dual symmetry, while the Toh direction has all but P -scale invariance. The AHO and Half directions enjoy all except well-defined directions and P -scale invariance, while the Gu direction satisfies all but P -scale invariance and primal-dual symmetry. All other directions fail at least three of the properties we have discussed.

However, there is another property that seems important, although it is not as well-defined. This is that the direction arise in some sense from taking a Newton step for some system.

Newton system. The basic property here (NS) is that the direction be the Newton step for some possibly locally-defined nonlinear system. Here by locally defined we

mean that the nonlinear system can depend on the current iterate. One great advantage of having a direction satisfy this property is that a Mehrotra-type predictor-corrector step [17] can be used in an algorithm for semidefinite programming using this direction [3]. From the derivation of the directions, we have

Proposition 6.11 *Property (NS) holds for all our directions except possibly Half, IHalf, J2, and V2.*

□

(Here and below we say “possibly” or “at least” because there could be another formulation of the defining equations for a direction of which we are unaware showing it to be a Newton step.)

Of course, the full benefit of Newton’s method is its asymptotic convergence rate, for which the nonlinear system to which Newton’s method is applied should be globally defined. We say the semi-strong Newton system (SSNS) property holds if the direction satisfies the Newton system (possibly with the right-hand side perturbed) for some *fixed* nonlinear system, defined for all positive definite \tilde{X} and \tilde{S} .

Proposition 6.12 *Property (SSNS) holds for at least the AHO, X-MT, S-MT, S-Chol-MT, J, V, and SXS directions.*

□

Finally, we say the strong Newton system (SNS) property holds if the fixed nonlinear system is defined in a neighborhood of any pair of symmetric positive semidefinite \tilde{X} and \tilde{S} (and hence also in a neighborhood of the solution).

Proposition 6.13 *Property (SNS) holds for at least the AHO and SXS directions.*

□

These properties show the AHO direction in a much better light. Note also that, according to the discussion at the end of Section 4, if we allow the use of a nonstandard step size (doubling the step for the dual part of the direction), then we can view the H..K..M direction as satisfying (SSNS) and (SNS), and similarly the dual H..K..M direction.

7 Concluding remarks

We have tested methods based on all these directions computationally on a limited set of problems, basing our code on SDPT3 [33]. In order to compare all the directions, we did not use the Mehrotra predictor-corrector variant for any direction, and we used the default step sizes and the default rule for updating the parameter ν . (In particular, we do not impose any requirement that the iterates lie in some neighborhood of the central path. Other authors have had some success with some of the directions which performed poorly in our tests, using step size rules to enforce some such neighborhood requirement.) We tested the methods on five random problems for each of seven different problem classes ((1) to (7) in [33]) of reasonable size.

Of the twenty methods (based on the twenty directions of Section 4), the best were those based on the AHO, Toh, Gu, H..K..M, and NT directions, roughly in that order, as far as robustness and accuracy were concerned. All of these achieved the desired accuracy of 10^{-6} in primal and dual feasibility and duality gap for all but at most two of the 35 problems, and often much higher precision. The next best methods were based on the MTW, dual H..K..M, and Half directions, but these failed to achieve the desired accuracy on between 10 and 15 of the 35 problems.

In terms of generating some iterates that were quite close to the central path, the AHO method was excellent, the Toh method very good, the Gu method good, and the Half and J2 methods fair. The other methods with good performance kept well away from the boundary but did not approach the central path closely.

With respect to the number of iterations to solve the problems, the top five methods were fairly close, with a slight advantage to the AHO method. In terms of time, the other methods were better, because less computation was required to obtain the Schur complement matrix (using its symmetry) than for the AHO method.

The top five methods were all among those satisfying all but at most two of the properties discussed in Section 6. (It is also interesting to note that all are members of the MZ family.) It is hard to argue with the excellent computational results of the AHO method, and the fact that (among the good directions) this is the only one known to have the strong Newton system property might help to explain this. On the other hand, for several problem classes very high accuracy is not required, and then the Toh direction (fast and robust, very accurate, only failing P -scale invariance) and the NT direction (fast and robust, reasonably accurate, with the best theoretical properties) seem attractive. It is also very possible that the best direction to use depends on the problem class; some authors (see, e.g., Helmberg et al. [10] and Benson et al. [5]) have reported excellent behavior for the H..K..M and dual directions for certain relaxations of combinatorial optimization problems.

Acknowledgement. I would like to thank Masayuki Shida for his many insightful comments on an earlier version of this manuscript, and for his simple proof of Lemma 6.1. I would also like to thank Levent Tunçel for very helpful discussions. Finally, I would like to thank the referees for their very careful reading and constructive comments.

References

- [1] F. Alizadeh, *Combinatorial optimization with interior point methods and semidefinite matrices*, Ph.D. thesis, University of Minnesota, Minneapolis, MN, 1991.
- [2] F. Alizadeh, *Semidefinite programming home page*, <http://rutcor.rutgers.edu/~alizadeh/sdp.html>
- [3] F. Alizadeh, J. A. Haeberly, and M. Overton, *Primal-dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results*, SIAM J. Optim., 8 (1998), pp. 746–768.

- [4] T. Ando, *Concavity of certain maps and positive definite matrices and applications to Hadamard products*, Linear Algebra and its Applications, 26 (1979), pp. 203–241.
- [5] S. J. Benson, Y. Ye, and X. Zhang, *Solving large-scale sparse semidefinite programs for combinatorial optimization*, Working Paper, Department of Management Science, University of Iowa, IA, 1997.
- [6] M. X. Goemans, *Semidefinite programming in combinatorial optimization*. Mathematical Programming, 79 (1997), pp. 143–161.
- [7] M. Gu, *On primal-dual interior point methods for semidefinite programming*, CAM report 97-12, Department of Mathematics, University of California, Los Angeles, CA, 1997.
- [8] C. Helmberg, *Semidefinite programming home page*, <http://www.zib.de/helmberg/semidef.html>
- [9] C. Helmberg, F. Rendl, R. Vanderbei, and H. Wolkowicz, *An interior-point method for semidefinite programming*, SIAM Journal on Optimization, 6 (1996), pp. 342–361.
- [10] C. Helmberg, F. Rendl, and R. Weismantel, *A semidefinite programming approach to the quadratic knapsack problem*, Preprint SC 96-10, Konrad-Zuse-Zentrum fuer Informationstechnik Berlin, Berlin, Germany, 1996.
- [11] R. Horn, private communication, February 1996.
- [12] J. Ji, F. Potra, and R. Sheng, *On the local convergence of a predictor-corrector method for semidefinite programming*, Reports on Computational Mathematics 98, Dept. of Mathematics, The University of Iowa, IA, 1997.
- [13] M. Kojima, S. Shindoh, and S. Hara, *Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices*, SIAM Journal on Optimization, 7 (1997), pp. 86–125.
- [14] M. Kojima, M. Shida, and S. Shindoh, *A note on the Nesterov-Todd and the Kojima-Shindoh-Hara search directions in semidefinite programming*, Technical Report B-313, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo 152, Japan, 1996 (to appear in Optimization Methods and Software).
- [15] M. Kojima, M. Shida, and S. Shindoh, *Search directions in the SDP and the monotone SDLCP: Generalization and inexact computation*, Technical Report B-327, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo 152, Japan, 1997 (to appear in Mathematical Programming).
- [16] S. Kruk, M. Muramatsu, F. Rendl, R. J. Vanderbei, and H. Wolkowicz, *The Gauss-Newton direction in linear and semidefinite programming*, in preparation, Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, 1997.
- [17] S. Mehrotra, *On the implementation of a primal-dual interior point method*, SIAM Journal on Optimization, 2 (1992), pp. 575–601.

- [18] The MathWorks, Inc., *MATLAB Reference Guide*, The MathWorks, Inc., Natick, MA, 1992.
- [19] R. D. C. Monteiro, *Primal-dual path following algorithms for semidefinite programming*, SIAM Journal on Optimization, 7 (1997), 663–678.
- [20] R. D. C. Monteiro, *Polynomial convergence of primal-dual algorithms for semidefinite programming based on the Monteiro and Zhang family of directions*, SIAM Journal on Optimization, 8 (1998), 797–812.
- [21] R. D. C. Monteiro and T. T. Tsuchiya, *Polynomial convergence of a new family of primal-dual algorithms for semidefinite programming*, Report Memorandum No. 627, The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106, Japan, 1996.
- [22] R. D. C. Monteiro and Y. Zhang, *A unified analysis for a class of path-following primal-dual interior-point algorithms for semidefinite programming*, Mathematical Programming, 81 (1998), 281–299.
- [23] R. D. C. Monteiro and P. R. Zanjacomo, *A note on the existence of the Alizadeh-Haeberly-Overtton direction for semidefinite programming*, Mathematical Programming, 78 (1997), 393–396.
- [24] R. D. C. Monteiro and P. R. Zanjacomo, *Implementation of primal-dual methods for semidefinite programming based on Monteiro and Tsuchiya Newton directions and their variants*, manuscript, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, 1997.
- [25] Yu. Nesterov and A. S. Nemirovskii, *Conic formulation of a convex programming problem and duality*, Optimization Methods and Software, 1(2) (1992), pp. 95–115.
- [26] Yu. E. Nesterov and A. S. Nemirovskii, *Interior Point Polynomial Methods in Convex Programming : Theory and Algorithms*, SIAM Publications, SIAM, Philadelphia, USA, 1994.
- [27] Yu. Nesterov and M. J. Todd, *Self-scaled barriers and interior-point methods in convex programming*, Mathematics of Operations Research, 22 (1997), pp. 1–42.
- [28] Yu. Nesterov and M. J. Todd, *Primal-dual interior-point methods for self-scaled cones*, SIAM Journal on Optimization, 8 (1998), 324–364.
- [29] M. Shida, private communication, October 1997.
- [30] M. Shida, S. Shindoh, and M. Kojima, *Existence and uniqueness of search directions in interior-point algorithms for the SDP and the monotone SDLCP*, SIAM Journal on Optimization, 8 (1998), 387–396.
- [31] M.J. Todd, K.C. Toh, R.H. Tütüncü, *On the Nesterov-Todd direction in semidefinite programming*, SIAM Journal on Optimization, 8 (1998), 769–796.
- [32] K. C. Toh, *Search directions for primal-dual interior point methods in semidefinite programming*, Technical Report, Department of Mathematics, National University of Singapore, Singapore, 1997.

- [33] K.C. Toh, M.J. Todd, and R.H. Tütüncü, *SDPT3 — a MATLAB software package for semidefinite programming*, Technical Report 1177, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY, 1996.
- [34] P. Tseng, *Search directions and convergence analysis of some infeasible path-following methods for the monotone semi-definite LCP*, *Optimization Methods and Software*, 9 (1998), pp. 245–268.
- [35] P. Tseng, *Further notes on the existence of search directions in interior-point methods for the monotone semi-definite LCP*, Technical Report, Department of Mathematics, University of Washington, Seattle, WA, 1996.
- [36] L. Tunçel, *Primal-dual symmetry and scale invariance of interior-point algorithms for convex optimization*, Research Report CORR 96–18, Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, 1996.
- [37] L. Vandenberghe and S. Boyd, *Semidefinite programming*, *SIAM Review*, 38 (1996), pp. 49–95.
- [38] Y. Zhang, *On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming*, *SIAM Journal on Optimization*, 8 (1998), 365–386.