SEARCH DIRECTIONS AND CONVERGENCE ANALYSIS
OF SOME INFEASIBLE PATH-FOLLOWING METHODS
FOR THE MONOTONE SEMI-DEFINITE LCP[1]

June 26, 1996

PAUL TSENG

Department of Mathematics, University of Washington
Seattle, Washington 98195, U.S.A. (tseng@math.washington.edu)

ABSTRACT

We consider a family of primal/primal-dual/dual search directions for the monotone LCP over the space of $n \times n$ symmetric block-diagonal matrices. We consider two infeasible predictor-corrector path-following methods using these search directions, with the predictor and corrector steps used either in series (similar to the Mizuno-Todd-Ye method) or in parallel (similar to Mizuno et al./McShane's method). The methods attain global linear convergence with a convergence ratio which, depending on the quality of the starting iterate, ranges from $1 - O(\sqrt{n})^{-1}$ to $1 - O(n)^{-1}$. Our analysis is fairly simple and parallels that for the LP and LCP cases.

1 Introduction

Since the original work of Nesterov and Nemirovskii [26], followed by that of Alizadeh [1] and Jarre [11], there has been very active research on interior-point methods for the semi-definite linear programming problem (SDLP) and the semi-definite linear complementarity problem (SDLCP). In particular, various search directions have been proposed [2, 3, 9, 11, 13, 24, 27, 28] and properties of these search directions, such as their existence and their computation, have been analyzed [3, 13, 24, 28, 32]. For SDLP, global linear convergence results have been obtained for feasible primal/primal-dual path-following methods [5, 6, 17, 24, 27, 30, 34, 40] and primal potential-reduction methods [2, 27, 39] and for infeasible primal/primal-dual path-following methods [7, 14, 18, 29, 30, 40].[2] For the more general problem of monotone SDLCP, global linear convergence results have been obtained for feasible and infeasible primal-dual path-following and potential-reduction methods using search directions from the Kojima-Shindo-Hara (KSH) family [13], and for Mizuno-Todd-Ye type infeasible predictor-corrector methods using the KSH/Monteiro (KSH/M) direction [14] or the Alizadeh-Haeberly-Overton (AHO) direction [15]. Under the

[1]This research is supported by National Science Foundation Grant CCR-9311621.

[2]During the writing of this paper, two other works on this topic, [25] and [4], the latter being a substantially expanded version of [3], appeared.

additional assumption that a strictly complementary solution exists and with the enforcement of the central-path neighborhood tending to zero, local superlinear convergence results have been obtained for a feasible predictor-corrector method (using the NT direction) of Nesterov and Todd (NT) [19] and for some infeasible predictor-corrector methods (using the KSH/M direction) of Kojima, Shida and Shindoh [14, 31] and of Potra and Sheng [29, 31]. Alternatively, under the additional assumption that a nondegenerate strictly complementary solution exists, local quadratic convergence result has been obtained for an infeasible predictor-corrector method (using the AHO direction) of Kojima, Shida and Shindoh [15]. Computational experiences with the methods are reported in [3, 9, 11, 18, 36, 39]. Further extension of the results to semi-definite nonlinear complementarity problems is presented in [33].

Motivated by the preceding work (and, in particular, [13, 32]), we study in this paper infeasible path-following methods for the monotone SDLCP defined over the space of $n \times n$ block-diagonal matrices. We address the issues of existence and uniqueness and, to a lesser extent, computation of search directions, and the global (linear) convergence of the methods. More precisely, we give a simple proof of existence and uniqueness for a family of search directions, including the primal-scaling direction [2, 27], the KSH/M direction [9, 13, 24], and the NT direction [28, 36] (see Lem. 2.1). Next, we consider two infeasible predictor-corrector path-following methods that employ these search directions either in series (in the spirit of the Mizuno-Todd-Ye method) or in parallel (in the spirit of Mizuno et al./McShane's method). We show that, assuming a solution exists, the methods attain global linear convergence with a convergence ratio ranging from $1-O(\sqrt{n})^{-1}$ to $1-O(n)^{-1}$, depending on the quality of the starting iterate (see Thms. 5.1 and 5.2). Detection of solution (non)existence is also discussed. While there have been many studies of infeasible path-following methods for SDLP and monotone SDLCP, our analysis is notable for its simplicity and, in particular, its close parallel with that for LP and LCP cases. Also, some of the search directions appear to be new, and the complexity results match the best possible.

We formally describe our problem below. Let $\mathcal{X}$ denote the space of $n \times n$ block-diagonal real matrices with m blocks of sizes $n_1, ..., n_m$, respectively (the blocks are fixed). Thus, $\mathcal{X}$ is closed under matrix addition $x+y$, multiplication xy, transposition x^T, and inversion x^{-1}, where $x, y \in \mathcal{X}$. We endow $\mathcal{X}$ with the inner product

$$\langle x, y \rangle := \mathrm{tr}[x^T y],$$

where $x, y \in \mathcal{X}$ and $\mathrm{tr}[\cdot]$ denotes the matrix trace. Let $\mathcal{S}$ denote the subspace comprising those $x \in \mathcal{X}$ that are symmetric, i.e., $x^T = x$. Let $\mathcal{S}_+$ (respectively, $\mathcal{S}_{++}$) denote the convex cone comprising those elements of $\mathcal{S}$ that are positive semi-definite (respectively, positive definite). Our problem is to find, for a given linear monotone mapping $M : \mathcal{S} \mapsto \mathcal{S}$ (so $\langle x, Mx \rangle \geq 0$ for all $x \in \mathcal{S}$) and a $q \in \mathcal{S}$, an $(x, y) \in \mathcal{S}_+ \times \mathcal{S}_+$ satisfying

$$\langle x, y \rangle = 0, \qquad Mx + q = y. \tag{1.1}$$

This problem contains as special cases the SDLP (for which $n_2 = \cdots = n_m = 1$ and M is skew-symmetric in the sense that $\langle x, Mx \rangle = 0$ for all $x \in \mathcal{S}$) and monotone LCP (for which $n_1 = \cdots = n_m = 1$). Notice that we have written the problem differently from the SDLCP of [13, 14]–here the matrices are block-diagonal and the linear constraints on x and y are expressed explicitly. This makes for a more transparent identification with the LCP case and easy incorporation of weights in the central path. Also, we use lower case, rather than upper case, letters for matrices, to further emphasize the parallel with LP and LCP cases.

Define, for each $x \in \mathcal{S}_{++}$, the linear mapping $x\odot : \mathcal{S} \mapsto \mathcal{S}$ by

$$x \odot y := x^{1/2} y x^{1/2} \quad \forall y \in \mathcal{S},$$

where $x^{1/2} \in \mathcal{S}_{++}$ is the square root of x, i.e., $x^{1/2}x^{1/2} = x$. ["$:=$" means "define".] Notice that $x\odot$ has an inverse, namely $x^{-1}\odot$, and maps $\mathcal{S}_{++}$ to itself. We denote by e the identity matrix in $\Re^{n\times n}$. For any $x \in \mathcal{S}$ with eigenvalues $\xi_1, ..., \xi_n$, we denote by $\|x\|_1$, $\|x\|$, $\|x\|_\infty$, respectively,

$$\|x\|_1 := |\xi_1| + \cdots + |\xi_n|, \|x\| := \sqrt{\langle x, x\rangle} = \sqrt{\xi_1^2 + \cdots + \xi_n^2}, \|x\|_\infty := \max\{|\xi_1|, \cdots, |\xi_n|\}.$$

[$\|x\|$ and $\|x\|_\infty$ are the Frobenius-norm and the 2-norm of x.] Also, denote $\min(x) = \min\{\xi_1, ..., \xi_n\}$. In the LCP case where $n_1 = \cdots = n_m = 1$, the mapping $x\odot$ reduces to componentwise multiplication by x, e reduces to the vector of 1s in $\Re^n$, and $\|\cdot\|_1$, $\|\cdot\|$, $\|\cdot\|_\infty$ reduce to, respectively, the 1-norm, 2-norm, ∞-norm on $\Re^n$. Just as they play key roles in LP and LCP cases, they play analogous roles in the SDLCP case [2]. In particular, we will approximate our problem by that of finding an $(x, y) \in \mathcal{S}_{++} \times \mathcal{S}_{++}$ satisfying the "central-path" equation:

$$x \odot y = \omega e, \qquad Mx + q - y = 0, \tag{1.2}$$

with $\omega \in (0, \infty)$ the "centering" parameter. Then, starting with an ω and an inexact solution to (1.2), we take a few (typically one or two) Newton steps so to improve the accuracy of the solution, after which we decrease ω and repeat this process. [A method is "feasible" if it maintains the second equation of (1.2) satisfied at all iterations, and otherwise is "infeasible" [12].] We will measure accuracy of solution to the first equation of (1.2) using the popular measure of proximity to the central path:

$$\phi_\omega(x, y) \ := \ \|x \odot y - \omega e\|/\omega, \tag{1.3}$$

where $(x, y) \in \mathcal{S}_{++} \times \mathcal{S}$.

2 Choices of Search Directions

A key ingredient of path-following methods is the Newton step used to improve the accuracy of an approximate solution to (1.2). In particular, for a given approximate solution (x, y) to (1.2), we need to find a search direction (u, v), obtained as the solution to a "linearization" of (1.2), along which to move (x, y). More precisely, $(u, v) \in \mathcal{S} \times \mathcal{S}$ should be the unique solution to equations of the form:

$$Au + Bv = r, \qquad Mu - v = -s. \tag{2.1}$$

where $(r, s) \in \mathcal{S} \times \mathcal{S}$, and A and B are suitably chosen linear mappings from $\mathcal{S}$ to $\mathcal{S}$. We will focus on the following choices of A and B:

1. Primal Scaling:

$$Au := \omega x^{-1} \odot u, \quad Bv := x \odot v. \tag{2.2}$$

This choice of A and B was used by Alizadeh [1, 2], Nesterov and Todd [27], and others [5, 6, 39].

2. Primal-Dual Scaling:

$$Au := \frac{1}{2}\left((x \odot y)^{\tau}(x^{-1} \odot u)(x \odot y)^{1-\tau} + (x \odot y)^{1-\tau}(x^{-1} \odot u)(x \odot y)^{\tau}\right), \quad Bv := x \odot v, \tag{2.3}$$

where $\tau \in [0, \frac{1}{2}]$ is any rational number whose denominator is an integer power of 2. We will also consider a variant whereby $x \odot y$ in (2.3) is replaced by $y \odot x$. When $\tau = \frac{1}{2}$, the above choice of A and B coincides with that proposed by Nesterov and Todd [28]. When $\tau = 0$, the above choice coincides with one proposed by Monteiro [24, Eq. (3)] (also see [40]), which in turn is a special case of the following choice proposed by Kojima, Shindo and Hara [13]:

$$Au := (x^{-1} \odot (u + \tilde{u}))(x \odot y), \quad Bv := x \odot (v + \tilde{v}),$$

where the auxiliary matrices $\tilde{u}$ and $\tilde{v}$ are skew-symmetric and depending on u, v. In particular, for any $(u, v) \in \mathcal{S} \times \mathcal{S}$, by setting

$$\tilde{u} := 0, \qquad \tilde{v} := \frac{1}{2}x^{-1} \odot \left((x \odot y)(x^{-1} \odot u) - (x^{-1} \odot u)(x \odot y)\right),$$

the above Au and Bv reduce to that given by (2.3) with $\tau = 0$. And the corresponding search direction (u, v) remains unchanged. It is shown recently that the NT direction also belongs to the KSH family [16]. This raises the open question of whether, in general, the direction corresponding to (2.3) belongs to this family.

3. Dual Scaling:

$$Au := y \odot u, \quad Bv := \omega y^{-1} \odot v. \tag{2.4}$$

This choice of A and B is obtained from (2.2) by interchanging the roles of x and y.

The choice of A and B in (2.3) corresponds to primal-dual scaling in the sense that, when $n_1 = \cdots = n_m = 1$, the corresponding search direction reduces to that used by primal-dual interior point methods for LP and LCP (independent of τ and of whether $x \odot y$ is replaced by $y \odot x$). Similarly, the choices (2.2) and (2.4) correspond to primal and dual scaling, respectively. Below, we show that the search direction (u, v) corresponding to each of the above choices of A and B exists and is unique. We will also investigate other choices of search direction.

Lemma 2.1 *Consider two linear mappings $A : \mathcal{S} \mapsto \mathcal{S}$ and $B : \mathcal{S} \mapsto \mathcal{S}$ such that B^*, the adjoint of B (i.e., $\langle d, Bd\rangle = \langle B^*d, d\rangle$ for all $d \in \mathcal{S}$), is invertible and AB^* is strictly monotone (i.e., $\langle d, AB^*d\rangle > 0$ whenever $d \neq 0$). For any $(r, s) \in \mathcal{S} \times \mathcal{S}$, there exists a unique $(u, v) \in \mathcal{S} \times \mathcal{S}$ satisfying* (2.1). *In particular, the above hypothesis on A, B are satisfied by the choices* (2.2), (2.3) *(with or without replacing $x \odot y$ by $y \odot x$) and* (2.4)*, where $(x, y) \in \mathcal{S}_{++} \times \mathcal{S}_{++}$.*

Proof. Upon eliminating v from the equations in (2.1), we have $Au + BMu = r - Bs$. Upon making the substitution $d = (B^*)^{-1}u$, we obtain equivalently $AB^*d + BMB^*d = r - Bs$. Since AB^* is strictly monotone, by virtue of BMB^* being monotone ($\langle d, BMB^*d\rangle = \langle B^*d, MB^*d\rangle \geq 0$ for all $d \in \mathcal{S}$), we see that $AB^* + BMB^*$ is also strictly monotone and hence invertible. [In general, any strictly monotone, linear mapping $L : \mathcal{S} \mapsto \mathcal{S}$ is invertible. This can be more easily seen by translating the setting from $\mathcal{S}$ to $\Re^\nu$, where $\nu := \frac{n_1(n_1+1)}{2} + \cdots + \frac{n_m(n_m+1)}{2}$: Choose a basis $x_1, ..., x_\nu$, for $\mathcal{S}$ (e.g., the element of $\mathcal{S}$ with a 1 in its (i, j)th and (j, i)th entry and 0 everywhere else, for all $i \leq j$) and define the invertible linear mappings $K : (\alpha_1, ..., \alpha_\nu) \in \Re^\nu \mapsto \alpha_1 x_1 + \cdots + \alpha_\nu x_\nu \in \mathcal{S}$. The linear mapping $K^{-1}LK : \Re^\nu \mapsto \Re^\nu$ has $\{0\}$ as its nullspace and hence is invertible. Then, L must be invertible.]

For A, B given by (2.2), B^* (which equals B) is clearly invertible and $d \mapsto AB^*d = \omega d$ is clearly strictly monotone. For A, B given by (2.3), B^* is clearly invertible and $d \mapsto AB^*d = \frac{1}{2}((x \odot y)^\tau d(x \odot y)^{1-\tau} + (x \odot y)^{1-\tau} d(x \odot y)^\tau)$ is strictly monotone since $\langle d, AB^*d\rangle = \mathrm{tr}\,[d(x \odot y)^\tau d(x \odot y)^{1-\tau}] = \|(x \odot y)^{\tau/2} d(x \odot y)^{(1-\tau)/2}\|^2$ which equals 0 if and only if $d = 0$. If instead $x \odot y$ in (2.3) is replaced by $y \odot x$, the argument is similar. For A, B given by (2.4), B^* (which equals B) is invertible and $d \mapsto AB^*d = \omega d$ is clearly strictly monotone. ■

The proof of Lem. 2.1 suggests a general strategy for solving (2.1), namely, express u and v as linear combinations of some basis elements for $\mathcal{S}$ and solve for the coefficients in the combinations (which can be done using Gaussian elimination in $O(n^6)$ time). Instead of using the basis elements all at once, we can also use a column generation strategy: Choose a subset of basis elements and find the corresponding least square solution to (2.1); if this solution is not sufficiently accurate, add more basis elements and update it (using, say, a low-rank updating scheme). For specific choices of A and B, specialized strategies using Cholesky factorization, singular value decomposition, and/or Schur complement equation [3, 9, 34, 36] or using conjugate-gradient iterations [39] have been developed.

In addition to the choices (2.2), (2.3) and (2.4), there are many other choices of A and B that satisfy the hypothesis of Lem. 2.1 and, in the cases of LP and LCP, yield the *same* search direction. In fact, there are *infinitely many* such search directions, obtained by, for example, replacing $x \odot y$ in (2.3) with $x^{1/2} \odot (y \odot x^{1/2})$ or with $x^2 \odot (y \odot x^{-1})$ or with..., and similarly for (2.2) and (2.4). On the other hand, not all these search directions may be amenable to a convergence analysis, and it remains to determine which ones are amenable. Lastly, we note that the simple choice of

$$Au := y \odot u, \quad Bv := x \odot v,$$

also considered by Todd [35], does not satisfy the hypothesis of Lem. 2.1. The difficulty lies in that $\langle d, AB^* d\rangle = \mathrm{tr}\left[dy^{1/2}x^{1/2}dx^{1/2}y^{1/2}\right]$ which, because $y^{1/2}x^{1/2}$ need not be symmetric positive definite, need not be positive for $d \neq 0$. However, this choice does satisfy the hypothesis of Lem. 2.1 provided that, in addition, (x, y) is well centered in the sense that the proximity measure $\|y^{1/2}x^{1/2} - \sqrt{\omega}e\|/\sqrt{\omega}$, which can be seen to be bounded below by $\phi_\omega(x, y)/(1 + \sqrt{2})$, is less than 1. Whether this or other search directions can be used in interior-point methods remains to be investigated.

3 Infeasible Predictor-Corrector Path-Following Methods

A very effective way to accelerate path-following methods is to alternately set the centering parameter ω to 0 ("predictor") and to 1 ("corrector") when taking a Newton step for (1.2). We will consider two such methods–the first using the predictor and corrector steps in series, in the spirit of the Mizuno-Todd-Ye method [22], and the second using the steps in parallel, in the spirit of Mizuno et al./McShane's method [20, 21]. We describe the first method below. This is an infeasible method based on that studied in [38] for the LCP case, and it differs from other predictor-corrector methods in that ω and $Mx + q - y$ decrease in norm at both the predictor step and the corrector step.

Algorithm 3.1 *Choose* $(\beta_0, \beta_1, \beta_2) \in \Re^3$ *satisfying* $0 < \beta < \beta_0 < \beta_1 < \beta_2 < 1$, *where* β *depends on* $\beta_0, \beta_1, \beta_2$, *and let*

$$\eta_1 = (\beta_0 - \beta)/(\beta_0 + \sqrt{n}). \tag{3.1}$$

Also, for each $(x, y, \omega) \in \mathcal{S}_{++} \times \mathcal{S}_{++} \times (0, \infty)$, *choose two linear mappings* $A : \mathcal{S} \mapsto \mathcal{S}$ *and* $B : \mathcal{S} \mapsto \mathcal{S}$ *with the properties that* B^* *is invertible and* AB^* *is strictly monotone. Choose any* $(x^0, y^0, \omega^0) \in \mathcal{S}_{++} \times \mathcal{S}_{++} \times (0, \infty)$ *satisfying* $\phi_{\omega^0}(x^0, y^0) \le \beta_0$. *For* $t = 0, 1, ...,$ *compute* $(x^{t+1}, y^{t+1}, \omega^{t+1})$ *from* (x^t, y^t, ω^t) *according to the following iteration:*

Predictor Step: *Denote* $(x, y, \omega) := (x^t, y^t, \omega^t)$. *Let* $(u_p, v_p) \in \mathcal{S} \times \mathcal{S}$ *be the solution to*

$$Au + Bv = -x \odot y, \qquad Mu - v = -(Mx + q - y). \tag{3.2}$$

Let θ^t *be the largest* $\theta \in (0, 1]$ *satisfying*

$$(x + \alpha u_p, y + \alpha v_p) \in \mathcal{S}_{++} \times \mathcal{S}_{++}, \quad \phi_{(1-\alpha)\omega}(x + \alpha u_p, y + \alpha v_p) \le \beta_1 \quad \forall \alpha \in (0, \theta], \tag{3.3}$$

and let

$$\hat{x}^t := x + \theta^t u_p, \qquad \hat{y}^t := y + \theta^t v_p, \qquad \hat{\omega}^t := (1-\theta^t)\omega. \tag{3.4}$$

<u>*Corrector Step*</u>*: Denote* $(x, y, \omega) := (\hat{x}^t, \hat{y}^t, \hat{\omega}^t)$*. Let* γ^t *be the largest* $\gamma \in (0, \eta_1]$ *satisfying*

$$\|x \odot y - \omega e + \gamma x \odot (Mx + q - y)\|/\omega \le \beta_2. \tag{3.5}$$

Let $(u_c, v_c) \in \mathcal{S} \times \mathcal{S}$ *be the solution to*

$$Au + Bv = \omega e - x \odot y, \qquad Mu - v = -\gamma^t(Mx + q - y). \tag{3.6}$$

Let $\hat{\theta}^t$ *be any (e.g., largest)* $\theta \in [1, 1/\gamma^t]$ *satisfying*

$$(x + \theta u_c, y + \theta v_c) \in \mathcal{S}_{++} \times \mathcal{S}_{++}, \quad \phi_{(1-\gamma^t\theta)\omega}(x + \theta u_c, y + \theta v_c) \le \beta_0, \tag{3.7}$$

and let

$$x^{t+1} := x + \hat{\theta}^t u_c, \qquad y^{t+1} := y + \hat{\theta}^t v_c, \qquad \omega^{t+1} := (1 - \gamma^t\hat{\theta}^t)\omega. \tag{3.8}$$

The scalars θ^t, γ^t are well defined due to $\beta_0 < \beta_1 < \beta_2$ (see the proof of Thm. 5.1 for a careful argument of this) and can be found by solving, respectively, a quartic and a quadratic equation in 1 variable. We will show in Lem. 4.2 that $\hat{\theta}^t$ is well defined and, in particular can be set to 1. However, this requires β to be chosen judiciously, depending on A and B: For A, B given by (2.2), choose

$$\beta := \beta_2^2. \tag{3.9}$$

For A, B given by (2.3), choose

$$\beta := \frac{\beta_2\beta_1}{(1-\beta_1)} + \frac{\beta_2^2(1+\beta_1)}{(1-\beta_1)^2}. \tag{3.10}$$

For A, B given by (2.4), choose

$$\beta := (\beta_1 + \beta_2)^2, \tag{3.11}$$

and, in addition, replace $x \odot y$ in (3.6) by $y \odot x$ and replace (3.5) by

$$\|y \odot x - \omega e - \gamma\omega y^{-1} \odot (Mx + q - y)\|/\omega \le \beta_2. \tag{3.12}$$

This replacement reflects the change in B (from $x\odot$ to $\omega y^{-1}\odot$) and the minus sign in front of v in (3.6). Each of the three choices of β above has the property that it tends to 0 quadratically with β_1 and β_2, so the requirement $\beta < \beta_0 < \beta_1 < \beta_2$ is met by choosing $\beta_1 < \beta_2$ sufficiently near 0 (so that $\beta < \beta_1$) and then choosing β_0 between β and β_1. In particular, this requirement is met by the choice $\beta_0 = .7, \beta_1 = .75, \beta_2 = .8$

when β is given by (3.9), by the choice $\beta_0 = .19, \beta_1 = .2, \beta_2 = .25$ when β is given by (3.10), by the choice $\beta_0 = .19, \beta_1 = .2, \beta_2 = .22$ when β is given by (3.11).

A disadvantage of Algorithm 3.1 (and of Mizuno-Todd-Ye type predictor-corrector method in general) is that A and B are different in the predictor step and the corrector step, so the computational effort per iteration is double that of ordinary path-following methods in which only corrector steps are used. A way around this, employed in our second method, is to take the two steps in parallel, rather than in series, and then take a convex combination of the steps. This idea has been used by Mizuno et al. [21], McShane [20] and Gonzaga [8] in the context of feasible path-following methods for LP and monotone LCP. Our method is an infeasible method based on an unpublished work for monotone LCP, done by the author (circa 1994) while unaware of the works [21, 8, 20].

Algorithm 3.2 *Choose* $(\beta_0, \beta_1, \beta_2) \in \Re^3$ *satisfying* $0 < \beta < \beta_0 < \beta_1 < \beta_2 < 1$, *where* β *depends on* $\beta_0, \beta_1, \beta_2$. *Also, for each* $(x, y, \omega) \in \mathcal{S}_{++} \times \mathcal{S}_{++} \times (0, \infty)$, *choose two linear mappings* $A : \mathcal{S} \mapsto \mathcal{S}$ *and* $B : \mathcal{S} \mapsto \mathcal{S}$ *with the properties that* B^* *is invertible and* AB^* *is strictly monotone. Choose any* $(x^0, y^0, \omega^0) \in \mathcal{S}_{++} \times \mathcal{S}_{++} \times (0, \infty)$ *satisfying* $\phi_{\omega^0}(x^0, y^0) \le \beta_0$. *For* $t = 0, 1, ...,$ *compute* $(x^{t+1}, y^{t+1}, \omega^{t+1})$ *from* (x^t, y^t, ω^t) *according to the following iteration:*

Predictor & Corrector Steps: Denote $(x, y) := (x^t, y^t)$. *Let* γ^t *be the largest* $\gamma \in (0, 1]$ *satisfying* $\phi_\omega(x, y) \le \beta_1$ *and* (3.5), *with* $\omega = (1 - \gamma)\omega^t$. *Let* $\omega := (1 - \gamma^t)\omega^t$. *Let* $(u_p, v_p) \in \mathcal{S} \times \mathcal{S}$ *be the solution to* (3.2) *and let* $(u_c, v_c) \in \mathcal{S} \times \mathcal{S}$ *be the solution to* (3.6). *Let* θ^t *be the largest* $\theta \in [0, 1]$ *satisfying*

$$\begin{aligned} (x + \theta u_p + (1 - \theta)u_c, y + \theta v_p + (1 - \theta)v_c) &\in \mathcal{S}_{++} \times \mathcal{S}_{++}, \\ \phi_{(1-\theta)\omega}(x + \theta u_p + (1 - \theta)u_c, y + \theta v_p + (1 - \theta)v_c) &\le \beta_0, \end{aligned} \tag{3.13}$$

and let

$$x^{t+1} := x + \theta^t u_p + (1 - \theta^t)u_c, \quad y^{t+1} := y + \theta^t v_p + (1 - \theta^t)v_c, \quad \omega^{t+1} := (1 - \theta^t)\omega. \tag{3.14}$$

The scalar γ^t is well defined since $\beta_0 < \beta_1 < \beta_2$. We will show in Lem. 4.2 that (3.13) is satisfied by $\theta = 0$ so θ^t is well defined. Both γ^t and θ^t are computable by solving, respectively, a quadratic equation and a quartic equation in 1 variable. We remark that, for our global convergence result (Thms. 3.1, 3.2), it suffices that (u_p, v_p) satisfies the second equation in (3.2) only. However, for practical computation and for local convergence analysis, the first equation (with the primal-dual choice (2.3) of A, B) is also needed. Also, analogous to LP and LCP cases, we can use weighted central path, whereby e in the definition of $\phi_\omega(\cdot)$ is replaced by a diagonal matrix whose diagonal entries are positive and are the same in each block. Our analysis extends accordingly as per [37].

4 Centering Properties of Search Directions

In this section, we derive properties of the linear mapping $x\odot$ and use these properties to analyze the search direction (u_c, v_c), obtained by solving (3.6) with A and B given by either (2.2) or (2.3) or (2.4) and with (x, y) satisfying the centering condition $\phi_\omega(x, y) \le \beta_1$ for some ω and γ^t satisfying (3.5). These results are key to the convergence analysis of Algorithms 3.1 and 3.2 in the next section. The proofs are fairly simple and reveal close parallels with the LCP case.

In what follows, we will freely use the following facts about trace [10]: For any $x, y \in \mathcal{X}$, $\mathrm{tr}[x^k] = \xi_1^k + \cdots + \xi_n^k$ $(k \in \mathbb{N})$, $\mathrm{tr}[xy] = \mathrm{tr}[yx]$, and $\mathrm{tr}[x + y] = \mathrm{tr}[x] + \mathrm{tr}[y]$, where $\xi_1, ..., \xi_n$ denote the eigenvalues of x (which are real if $x \in \mathcal{S}$ and are positive if $x \in \mathcal{S}_{++}$). Also, $\|\cdot\|$ is a norm on $\mathcal{X}$ and, in particular, the triangle inequality and the Cauchy-Schwarz inequality hold for $\|\cdot\|$. Moreover, $\|x^T\| = \sqrt{\mathrm{tr}[x^T x]} = \sqrt{\mathrm{tr}[xx^T]} = \|x\|$ so that $\mathrm{tr}[x^2] = \langle x, x^T\rangle \le \|x\|\|x^T\| = \|x\|^2$, and, for any orthogonal $p \in \mathcal{X}$ (i.e., $p^T = p^{-1}$), $\|pxp^T\| = \sqrt{\mathrm{tr}[px^Tp^Tpxp^T]} = \sqrt{\mathrm{tr}[x^Tp^Tpxp^Tp]} = \sqrt{\mathrm{tr}[x^Tx]} = \|x\|$.

Then, for any $d, w \in \mathcal{S}$ and $z \in \mathcal{S}_{++}$, we have

$$\langle d, z \odot w\rangle = \mathrm{tr}\left[dz^{1/2}wz^{1/2}\right] = \mathrm{tr}\left[z^{1/2}dz^{1/2}w\right] = \langle z \odot d, w\rangle, \tag{4.1}$$

$$\mathrm{tr}[d^4] = \delta_1^4 + \cdots + \delta_n^4 \le (\delta_1^2 + \cdots + \delta_n^2)^2 = \mathrm{tr}[d^2]^2 = \|d\|^4, \tag{4.2}$$

where $\delta_1, ..., \delta_n$ denote the eigenvalues of d. For any positive rational numbers τ, σ whose denominators are integer powers of 2, we have

$$\begin{aligned}
\|z^\tau dz^\sigma\|^2 &= \|pz^\tau p^T pdp^T pz^\sigma p^T\|^2 \\
&= \|\mathrm{diag}[\zeta_1, ..., \zeta_n]^\tau pdp^T \mathrm{diag}[\zeta_1, ..., \zeta_n]^\sigma\|^2 \\
&= \sum_{i=1}^n \sum_{j=1}^n \left(\zeta_i^\tau [pdp^T]_{ij} \zeta_j^\sigma\right)^2 \\
&\ge \sum_{i=1}^n \sum_{j=1}^n \left([pdp^T]_{ij}\right)^2 \zeta_1^{2(\tau+\sigma)} \\
&= \mathrm{tr}[(pdp^T)^2]\zeta_1^{2(\tau+\sigma)} \\
&= \mathrm{tr}[d^2]\zeta_1^{2(\tau+\sigma)} = \|d\|^2 \min(z^{2(\tau+\sigma)}),
\end{aligned} \tag{4.3}$$

where $0 < \zeta_1 \le ... \le \zeta_n$ denote the eigenvalues of z and $p \in \mathcal{X}$ is orthogonal and satisfies $pzp^T = \mathrm{diag}[\zeta_1, ..., \zeta_n]$. Similarly (also using (4.2)),

$$\begin{aligned}
\|dz^\tau dz^\sigma\|^2 &= \|pdp^T pz^\tau p^T pdp^T pz^\sigma p^T\|^2 \\
&= \|pdp^T \mathrm{diag}[\zeta_1, ..., \zeta_n]^\tau pdp^T \mathrm{diag}[\zeta_1, ..., \zeta_n]^\sigma\|^2 \\
&= \sum_{i=1}^n \sum_{j=1}^n \left(\sum_{k=1}^n [pdp^T]_{ik}\zeta_k^\tau [pdp^T]_{kj}\zeta_j^\sigma\right)^2 \\
&\le \sum_{i=1}^n \sum_{j=1}^n \left(\sum_{k=1}^n [pdp^T]_{ik}[pdp^T]_{kj}\right)^2 \zeta_n^{2(\tau+\sigma)}
\end{aligned}$$

$$
\begin{aligned}
&= \mathrm{tr}[(pdp^T)^4]\zeta_n^{2(\tau+\sigma)} \\
&= \mathrm{tr}[d^4]\|z^{\tau+\sigma}\|_\infty^2 \le \|d\|^4\|z^{\tau+\sigma}\|_\infty^2. && (4.4)
\end{aligned}
$$

If in addition $w \in \mathcal{S}_+$ (so $pwp^T \in \mathcal{S}_+$ and $\mathrm{tr}[w] = \|w\|_1$), then

$$
\begin{aligned}
\langle z, w\rangle &= \mathrm{tr}[zw] \\
&= \mathrm{tr}[pzp^T pwp^T] \\
&= \mathrm{tr}\left[\mathrm{diag}\left[\zeta_1, ..., \zeta_n\right] pwp^T\right] \\
&= \sum_{i=1}^n \zeta_i [pwp^T]_{ii} \\
&\ge \zeta_1 \sum_{i=1}^n [pwp^T]_{ii} \\
&= \zeta_1 \mathrm{tr}[pwp^T] = \zeta_1 \mathrm{tr}[w] = \min(z)\|w\|_1. && (4.5)
\end{aligned}
$$

Lemma 4.1 *For any $(x, y, \omega) \in \mathcal{S}_{++} \times \mathcal{S} \times (0, \infty)$ and any $\beta \in (0, 1)$ satisfying $\phi_\omega(x, y) \le \beta$, we have upon letting $z := x \odot y$ that*
(a) $(1-\beta)\omega \le \min(z) \le \|z\|_\infty \le (1+\beta)\omega$;
(b) $\langle x, y\rangle \le n(1+\beta)\omega$;
(c) $y \in \mathcal{S}_{++}$;
(d) $\|\omega d - \frac{1}{2}(z^\tau d z^{1-\tau} + z^{1-\tau} d z^\tau)\| \le \|d\|\beta\omega$, for any positive rational number τ whose denominator is an integer power of 2.

Proof. Since $\sqrt{(\zeta_1-\omega)^2 + \cdots + (\zeta_n - \omega)^2} = \|z - \omega e\| = \phi_\omega(x, y)\omega \le \beta\omega$, where $\zeta_1, ..., \zeta_n$ denote the eigenvalues of z, we have $0 < (1-\beta)\omega \le \zeta_i \le (1+\beta)\omega$ for $i = 1, ..., n$, which yields (a) and (b) ($\langle x, y\rangle = \mathrm{tr}[z] = \zeta_1 + \cdots + \zeta_n \le n(1+\beta)\omega$) and $z \in \mathcal{S}_{++}$ (cf. [13, Lem. 7.2]). The latter and $x \in \mathcal{S}_{++}$ imply (c). Lastly,

$$
\begin{aligned}
& \|\omega d - \frac{1}{2}(z^\tau d z^{1-\tau} + z^{1-\tau} d z^\tau)\|^2 \\
=\ & \|\omega pdp^T - \frac{1}{2}(pz^\tau p^T pdp^T pz^{1-\tau}p^T + pz^{1-\tau}p^T pdp^T pz^\tau p^T)\|^2 \\
=\ & \Big\|\omega pdp^T - \frac{1}{2}(\mathrm{diag}[\zeta_1, ..., \zeta_n]^\tau pdp^T \mathrm{diag}[\zeta_1, ..., \zeta_n]^{1-\tau} \\
+\ & \mathrm{diag}[\zeta_1, ..., \zeta_n]^{1-\tau} pdp^T \mathrm{diag}[\zeta_1, ..., \zeta_n]^\tau)\Big\|^2 \\
=\ & \sum_{i=1}^n \sum_{j=1}^n \left(\left(\omega - \frac{1}{2}(\zeta_i^\tau \zeta_j^{1-\tau} + \zeta_i^{1-\tau}\zeta_j^\tau)\right)[pdp^T]_{ij}\right)^2 \\
\le\ & \sum_{i=1}^n \sum_{j=1}^n (\beta\omega)^2 [pdp^T]_{ij}^2 \\
=\ & (\beta\omega)^2 \mathrm{tr}[(pdp^T)^2] = (\beta\omega)^2 \mathrm{tr}[d^2] = (\beta\omega)^2\|d\|^2,
\end{aligned}
$$

where $p \in \mathcal{X}$ is orthogonal and satisfies $pzp^T = \mathrm{diag}[\zeta_1, ..., \zeta_n]$; the inequality follows from $(1-\beta)\omega \le \zeta_i \le (1+\beta)\omega$ for $i = 1, ..., n$. ■

The following key lemma shows that the direction (u_c, v_c) is a centering direction at (x, y). The proof of this lemma is patterned after that of [38, Lem. 1] for the special case of monotone LCP and with A, B, β chosen by (2.2), (3.9).

Lemma 4.2 *Consider any $(\beta_0, \beta_1, \beta_2) \in \Re^3$ satisfying $0 < \beta < \beta_0 < \beta_1 < \beta_2 < 1$, where A, B, β is chosen by either (a)* (2.2), (3.9) *or (b)* (2.3), (3.10) *or (c)* (2.3), (3.10) *with $x \odot y$ replaced by $y \odot x$ or (d)* (2.4), (3.11). *Let η_1 be given by* (3.1). *For any $(x, y, \omega) \in \mathcal{S}_{++} \times \mathcal{S}_{++} \times (0, \infty)$ satisfying $\phi_\omega(x, y) \le \beta_1$, the $(u, v) \in \mathcal{S}_{++} \times \mathcal{S}_{++}$ satisfying* (3.6)*, with γ^t being any $\gamma \in (0, 1]$ satisfying* (3.5) *(with $x \odot y$ in* (3.6) *replaced by $y \odot x$ and* (3.5) *replaced by* (3.12) *for choice (d)), also satisfies*

$$(x+u, y+v) \in \mathcal{S}_{++} \times \mathcal{S}_{++}, \qquad \phi_{(1-\gamma)\omega}(x+u, y+v) \le \beta_0 \ \forall \gamma \in [0, \eta_1].$$

Proof. (a) Suppose A, B, β is given by (2.2), (3.9). Let $r := \omega e - x \odot y$, $s := \gamma^t(Mx + q - y)$ and $d := x^{-1} \odot u$. Then (3.6) may be rewritten as

$$\omega d + x \odot v = r, \qquad M(x \odot d) - v = -s, \tag{4.6}$$

yielding

$$[\omega I + x \odot Mx\odot]d = r - x \odot s.$$

Since the linear mapping $x \odot Mx\odot$ is monotone (in particular, by (4.1) and monotonicity of M, we have $\langle d, x \odot (M(x \odot d))\rangle = \langle x \odot d, M(x \odot d)\rangle \ge 0$), this implies

$$\omega\|d\|^2 \le \langle d, [\omega I + x \odot Mx\odot]d\rangle = \langle d, r - x \odot s\rangle \le \|d\|\|r - x \odot s\|,$$

and it follows from (3.5) that

$$\|d\| \le \|r - x \odot s\|/\omega \le \beta_2. \tag{4.7}$$

Let $x' := x + u$ and $y' := y + v$. By (4.7), we have $\|d\| \le \beta_2 < 1$ and so $e + d \in \mathcal{S}_{++}$. Then $x \in \mathcal{S}_{++}$ implies $x' = x + x \odot d = x \odot (e + d) \in \mathcal{S}_{++}$, so

$$\begin{aligned}
\|x' \odot y' - \omega e\|^2 &= \mathrm{tr}\left[x'y'x'y' - 2\omega x'y' + \omega^2 e\right] \\
&= \mathrm{tr}\left[\left((e+d)x^{1/2}y'x^{1/2} - \omega e\right)^2\right] \\
&= \mathrm{tr}[\omega^2 d^4] \\
&\le \omega^2\|d\|^4 \ \le \ \omega^2(\beta_2)^4 = \omega^2\beta^2,
\end{aligned}$$

where the first and second equalities use the symmetry of d, x, x', y' and properties of trace; the third equality uses $x \odot y' = \omega e - \omega d$ (see (4.6) and definition of r); the first and second inequalities follow from (4.2) and (4.7), respectively. Since $\beta < 1$ and

$x' \in \mathcal{S}_{++}$, the above relation implies that $y' \in \mathcal{S}_{++}$ (see Lem. 4.1(c)). Also, using the triangle inequality and the above relation, we have that, for every $\gamma \in [0, \eta_1]$,

$$\begin{aligned} \phi_{(1-\gamma)\omega}(x', y') &= \|x' \odot y' - (1-\gamma)\omega e\|/((1-\gamma)\omega) \\ &\leq \|x' \odot y' - \omega e\|/((1-\gamma)\omega) + \gamma\|e\|/(1-\gamma) \\ &\leq \beta/(1-\gamma) + \gamma\sqrt{n}/(1-\gamma), \end{aligned}$$

whose right-hand side, by $\gamma \leq \eta_1$ and (3.1), is below β_0.

(b) Suppose A, B, β is given by (2.3), (3.10). Let $r := \omega e - x \odot y$, $s := \gamma^t(Mx + q - y)$, $d := x^{-1} \odot u$, and $z := x \odot y$. Then (3.6) may be rewritten as

$$\frac{1}{2}(z^\tau dz^{1-\tau} + z^{1-\tau}dz^\tau) + x \odot v = r, \qquad M(x \odot d) - v = -s, \tag{4.8}$$

yielding

$$\frac{1}{2}(z^\tau dz^{1-\tau} + z^{1-\tau}dz^\tau) + [x \odot Mx \odot]d = r - x \odot s.$$

Since the linear mapping $x \odot Mx \odot$ is monotone, this implies

$$\begin{aligned} (1-\beta_1)\omega\|d\|^2 &\leq \min(z)\|d\|^2 \\ &\leq \|z^{\tau/2}dz^{(1-\tau)/2}\|^2 \\ &= \operatorname{tr}\left[dz^\tau dz^{1-\tau}\right] \\ &= \langle d, \frac{1}{2}(z^\tau dz^{1-\tau} + z^{1-\tau}dz^\tau)\rangle \\ &\leq \langle d, \frac{1}{2}(z^\tau dz^{1-\tau} + z^{1-\tau}dz^\tau) + [x \odot Mx \odot]d\rangle \\ &= \langle d, r - x \odot s\rangle \leq \|d\|\|r - x \odot s\|, \end{aligned}$$

where the first inequality uses Lem. 4.1(a) and the second inequality uses (4.3). Then, we have using (3.5) that

$$\|d\| \leq \|r - x \odot s\|/(1-\beta_1)\omega \leq \beta_2/(1-\beta_1). \tag{4.9}$$

Let $x' := x + u$ and $y' := y + v$. By (4.9), (3.10) and $\beta < 1$, we have $\|d\| \leq \beta_2/(1-\beta_1) < 1$ and so $e + d \in \mathcal{S}_{++}$. Then $x \in \mathcal{S}_{++}$ implies $x' = x + x \odot d = x \odot (e + d) \in \mathcal{S}_{++}$, so

$$\begin{aligned} \|x' \odot y' - \omega e\|^2 &= \operatorname{tr}\left[x'y'x'y' - 2\omega x'y' + \omega^2 e\right] \\ &= \operatorname{tr}\left[((e+d)x^{1/2}y'x^{1/2} - \omega e)^2\right] \\ &= \operatorname{tr}\left[((e+d)(\omega e - \frac{1}{2}(z^\tau dz^{1-\tau} + z^{1-\tau}dz^\tau) - \omega e)^2\right] \\ &= \operatorname{tr}\left[\left(\omega d - \frac{1}{2}(z^\tau dz^{1-\tau} + z^{1-\tau}dz^\tau) - \frac{1}{2}(dz^\tau dz^{1-\tau} + dz^{1-\tau}dz^\tau)\right)^2\right] \end{aligned}$$

$$
\begin{aligned}
&\le \left(\|\omega d - \frac{1}{2}(z^\tau dz^{1-\tau} + z^{1-\tau} dz^\tau)\| + \frac{1}{2}\|dz^\tau dz^{1-\tau}\| + \frac{1}{2}\|dz^{1-\tau} dz^\tau\|\right)^2 \\
&\le (\|d\|\beta_1\omega + \|d\|^2\|z\|_\infty)^2 \\
&\le (\beta_2\beta_1\omega/(1-\beta_1) + \beta_2^2(1+\beta_1)\omega/(1-\beta_1)^2)^2 \\
&= (\beta_2\beta_1/(1-\beta_1) + \beta_2^2(1+\beta_1)/(1-\beta_1)^2)^2\omega^2 = \beta^2\omega^2,
\end{aligned}
$$

where the first and second equalities use the symmetry of d, x, x', y' and properties of trace; the third equality uses $x \odot y' = \omega e - \frac{1}{2}(z^\tau dz^{1-\tau} + z^{1-\tau} dz^\tau)$ (see (4.8) and definition of r); the first inequality uses a property of trace and the triangle inequality; the second inequality uses (4.4) and Lem. 4.1(d); the third inequality uses Lem. 4.1(a) and (4.9). The remainder of the argument is identical to that for part (a).

(c) Suppose A, B, β is given by (2.3), (3.10) with $x \odot y$ replaced by $y \odot x$. The argument is identical to that for part (b), except with $z := y \odot x$ (and using the observation that $\phi_\omega(y,x) = \phi_\omega(x,y)$).

(d) Suppose A, B, β is given by (2.4), (3.11), with $x \odot y$ in (3.6) replaced by $y \odot x$ and (3.5) replaced by (3.12). Let $r := \omega e - y \odot x$, $s := \gamma^t(Mx + q - y)$ and $h := y^{-1} \odot v$. Then (3.6) may be rewritten as

$$
y \odot u + \omega h = r, \qquad Mu - y \odot h = -s. \tag{4.10}
$$

Since the linear mapping M is monotone, this implies

$$
\begin{aligned}
\omega\|h\|^2 &= \langle h, r - y \odot u\rangle \\
&= \langle h, r\rangle - \langle y \odot h, u\rangle \\
&= \langle h, r\rangle - \langle Mu + s, u\rangle \\
&\le \langle h, r\rangle - \langle s, u\rangle \\
&= \langle h, r\rangle - \langle s, y^{-1} \odot (r - \omega h)\rangle \\
&= \langle h, r\rangle - \langle y^{-1} \odot s, r - \omega h\rangle \\
&= \langle h, r + \omega y^{-1} \odot s\rangle - \langle y^{-1} \odot s, r\rangle \\
&\le \|h\|\|r + \omega y^{-1} \odot s\| + \|y^{-1} \odot s\|\|r\| \le \|h\|\beta_2\omega + (\beta_1 + \beta_2)\beta_1\omega,
\end{aligned}
$$

where the third and fourth equalities use (4.10); the second and fifth equalities use (4.1); the last inequality uses (3.12), $\|y^{-1} \odot s\| \le \|r\|/\omega + \|r + \omega y^{-1} \odot s\|/\omega$, and $\|r\|/\omega = \phi_\omega(y,x) = \phi_\omega(x,y) \le \beta_1$. Thus, $\|h\|^2 - \|h\|\beta_2 - (\beta_1 + \beta_2)\beta_1 \le 0$, which is a quadratic inequality in $\|h\|$ and implies

$$
\|h\| \le \beta_1 + \beta_2. \tag{4.11}
$$

Let $x' := x + u$ and $y' := y + v$. By (4.11), (3.11), and $\beta < 1$, we have $\|h\| \le \beta_1 + \beta_2 < 1$ and so $e + h \in \mathcal{S}_{++}$. Then $y \in \mathcal{S}_{++}$ implies $y' = y + y \odot h = y \odot (e + h) \in \mathcal{S}_{++}$, so

$$
\|x' \odot y' - \omega e\|^2 = \operatorname{tr}\left[x'y'x'y' - 2\omega x'y' + \omega^2 e\right]
$$

$$
\begin{aligned}
&= \operatorname{tr}\left[((e+h)y^{1/2}x'y^{1/2}-\omega e)^2\right] \\
&= \operatorname{tr}\left[\omega^2 h^4\right] \\
&\le \omega^2\|h\|^4 \;\le\; \omega^2(\beta_1+\beta_2)^4 = \omega^2\beta^2,
\end{aligned}
$$

where the first and second equalities also use the symmetry of h, x, x', y' and properties of trace; the third equality uses $y \odot x' = \omega e - \omega h$ (see (4.10) and definition of r); the first and second inequalities use, respectively, (4.2) and (4.11). The remainder of the argument is identical to that for part (a) but with the roles of x and y switched. ■

5 Global Linear Convergence Analysis

Let Σ denote the set of solutions (possibly empty) of the monotone SDLCP, i.e.,

$$\Sigma = \{\ (x,y) \in \mathcal{S}_+ \times \mathcal{S}_+ \ :\ (x,y) \text{ satisfies } (1.1)\ \}.$$

The next lemma, due originally to Mizuno [23] for the LP case (see Lem. 3.3 therein), gives a bound on $\|x\|$ and is widely used in the convergence analysis of infeasible-interior-point methods. For completeness, we include a short proof.

Lemma 5.1 *Assume that $\Sigma \neq \emptyset$. For any $\mu \in [0,1]$, any $(x^*, y^*) \in \Sigma$ and any $(x,y) \in \mathcal{S}_+ \times \mathcal{S}_+$ satisfying $Mx + q - y = \mu(Mx^0 + q - y^0)$, we have*

$$\mu(\langle x, y^0\rangle + \langle x^0, y\rangle) \;\le\; \langle x, y\rangle + \mu(\langle x^0, y^0\rangle + \langle x^*, y^0\rangle + \langle x^0, y^*\rangle). \tag{5.1}$$

Proof. We have from $Mx + q - y = \mu(Mx^0 + q - y^0)$ and $Mx^* + q = y^*$ that

$$M(x - \mu x^0 - (1-\mu)x^*) - (y - \mu y^0 - (1-\mu)y^*) = 0.$$

Taking inner product of both sides with $x - \mu x^0 - (1-\mu)x^*$ and using the monotonicity of M yields

$$\langle x - \mu x^0 - (1-\mu)x^*, y - \mu y^0 - (1-\mu)y^*\rangle \ge 0,$$

which, upon rearranging terms and using $\langle x^*, y^*\rangle = 0$, gives

$$
\begin{aligned}
&\mu(\langle x, y^0\rangle + \langle x^0, y\rangle) + (1-\mu)(\langle x, y^*\rangle + \langle x^*, y\rangle) \\
\le\;& \langle x, y\rangle + \mu^2\langle x^0, y^0\rangle + \mu(1-\mu)(\langle x^*, y^0\rangle + \langle x^0, y^*\rangle),
\end{aligned}
$$

Since (x,y), (x^0,y^0) and (x^*,y^*) are all in $\mathcal{S}_+ \times \mathcal{S}_+$ so that, by (4.5), the above inner products are all nonnegative, this and $0 \le \mu \le 1$ proves (5.1). ■

By using Lems. 4.2 and 5.1, we can now establish the global linear convergence of Algorithms 3.1 and 3.2. The proof is similar to that of [38, Thm. 3.1] and is included for completeness.

Theorem 5.1 *Let β_0, β_1, β_2, η_1, and $\{(x^t, y^t, \omega^t, \hat{x}^t, \hat{y}^t, \hat{\omega}^t, \theta^t, \gamma^t, \hat{\theta}^t)\}_{t=0,1,\ldots}$ be generated by Algorithm 3.1 with A, B, β given by any of the choices cited in Lem. 4.2 (with $x \odot y$ in (3.6) replaced by $y \odot x$ and (3.5) replaced by (3.12) for choice (d) therein). Then,*

$$(x^t, y^t) \in \mathcal{S}_{++} \times \mathcal{S}_{++}, \qquad \phi_{\omega^t}(x^t, y^t) \leq \beta_0, \tag{5.2}$$
$$\omega^t = \mu^t \omega^0, \qquad Mx^t + q - y^t = \mu^t(Mx^0 + q - y^0), \tag{5.3}$$

for all t, where

$$\mu^t := (1 - \gamma^{t-1}\hat{\theta}^{t-1})(1 - \theta^{t-1}) \cdots (1 - \gamma^0 \hat{\theta}^0)(1 - \theta^0). \tag{5.4}$$

If $\{(\hat{x}^t, \hat{y}^t)\}$ is bounded, then $\Sigma \neq \emptyset$. If $\Sigma \neq \emptyset$, then $\{(\hat{x}^t, \hat{y}^t)\}$ and $\{(x^t, y^t)\}$ are bounded and, for any $(x^*, y^*) \in \Sigma$, we have $\gamma^t \geq \min\{\eta_1, \eta_2\}$ for all t, where

$$\eta_2 \;:=\; \frac{(\beta_2 - \beta_1)\omega^0 \min(y^0)/\|Mx^0 + q - y^0\|_\infty}{n(1+\beta_1)\omega^0 + \langle x^0, y^0\rangle + \langle x^*, y^0\rangle + \langle x^0, y^*\rangle}. \tag{5.5}$$

Proof. If (5.2)-(5.3) hold for some t, then by $\beta_0 < \beta_1 < \beta_2$ and Lem. 4.2, the scalars θ^t, γ^t, and $\hat{\theta}^t$ are well defined ($\hat{\theta}^t$ can be set to 1) so, also using (3.4) and (3.8) and the second equation in (3.2) and (3.6), the relations (5.2)-(5.3) still hold when t is replaced by $t+1$. Then, an induction argument shows that (5.2)-(5.3) hold for all $t = 0, 1, \ldots$

For any t, we have from (3.3) with $\alpha = \theta^t$ and (3.4) that $\phi_{\hat{\omega}^t}(\hat{x}^t, \hat{y}^t) \leq \beta_1$, so Lem. 4.1(b) yields

$$\langle \hat{x}^t, \hat{y}^t \rangle \leq n(1+\beta_1)\hat{\omega}^t = n(1+\beta_1)(1-\theta^t)\mu^t\omega^0, \tag{5.6}$$

where the equality follows from (3.4) and (5.3). Similarly, for any $\gamma \in [0, \infty)$, we have

$$\begin{aligned} &\|\hat{x}^t \odot \hat{y}^t - \hat{\omega}^t e + \gamma \hat{x}^t \odot (M\hat{x}^t + q - \hat{y}^t)\|/\hat{\omega}^t \\ \leq\;& \phi_{\hat{\omega}^t}(\hat{x}^t, \hat{y}^t) + \gamma\|\hat{x}^t \odot (M\hat{x}^t + q - \hat{y}^t)\|/\hat{\omega}^t \;\leq\; \beta_1 + \gamma\|\hat{x}^t \odot (M\hat{x}^t + q - \hat{y}^t)\|/\hat{\omega}^t. \end{aligned}$$

Let $\bar{\gamma}^t$ be the largest $\gamma \in [0, \infty)$ for which the left-hand side is below β_2 (so $\gamma^t = \min\{\eta_1, \bar{\gamma}^t\}$). Then $\bar{\gamma}^t$ must exceed the largest γ for which the right-hand side is below β_2, so

$$\bar{\gamma}^t \geq \frac{(\beta_2 - \beta_1)\hat{\omega}^t}{\|\hat{x}^t \odot (M\hat{x}^t + q - \hat{y}^t)\|} \geq \frac{(\beta_2 - \beta_1)\hat{\omega}^t}{\|\hat{x}^t\|\|M\hat{x}^t + q - \hat{y}^t\|_\infty} = \frac{(\beta_2 - \beta_1)\omega^0}{\|\hat{x}^t\|\|Mx^0 + q - y^0\|_\infty}, \tag{5.7}$$

where the second inequality follows from an identity analogous to (4.4); the equality follows from $\hat{\omega}^t = (1-\theta^t)\mu^t\omega^0$ and $M\hat{x}^t + q - \hat{y}^t = (1-\theta^t)\mu^t(Mx^0 + q - y^0)$ (see (3.2), (3.4), (5.3)).

Assume $\{(\hat{x}^t, \hat{y}^t)\}$ is bounded. Then (5.7) yields that $\bar{\gamma}^t \geq \eta$ for some constant $\eta > 0$ and so $\gamma^t = \min\{\eta_1, \bar{\gamma}^t\} \geq \min\{\eta_1, \eta\}$ for all t. By (5.4) and $\hat{\theta}^t \geq 1$ for all t,

$\{\mu^t\} \to 0$. Since $(\hat{x}^t, \hat{y}^t) \in \mathcal{S}_{++} \times \mathcal{S}_{++}$, $M\hat{x}^t + q - \hat{y}^t = (1-\theta^t)\mu^t(Mx^0 + q - y^0)$ and (5.6) holds for all t, this implies any cluster point of $\{(\hat{x}^t, \hat{y}^t)\}$ is in Σ, so $\Sigma \neq \emptyset$.

Assume $\Sigma \neq \emptyset$ and fix any $(x^*, y^*) \in \Sigma$. For each t, we have from (5.2)-(5.3) and Lem. 5.1 that (5.1) holds (with $(x,y) = (\hat{x}^t, \hat{y}^t)$, $\mu = (1-\theta^t)\mu^t$), so

$$\begin{aligned}
\min(y^0)\|\hat{x}^t\|_1 + \min(x^0)\|\hat{y}^t\|_1 &\leq \langle \hat{x}^t, y^0 \rangle + \langle x^0, \hat{y}^t \rangle \\
&\leq \langle \hat{x}^t, \hat{y}^t \rangle / \mu^t + \langle x^0, y^0 \rangle + \langle x^*, y^0 \rangle + \langle x^0, y^* \rangle \\
&\leq n(1+\beta_1)\omega^0 + \langle x^0, y^0 \rangle + \langle x^*, y^0 \rangle + \langle x^0, y^* \rangle,
\end{aligned}$$

where the first inequality follows from (4.5) and the last inequality follows from (5.6). Since $(x^0, y^0) \in \mathcal{S}_{++} \times \mathcal{S}_{++}$ so $\min(x^0) > 0, \min(y^0) > 0$, this shows $\{(\hat{x}^t, \hat{y}^t)\}$ is bounded. A similar argument shows $\{(x^t, y^t)\}$ is bounded. Using the above inequality and $\|\hat{x}^t\| \leq \|\hat{x}^t\|_1$ to bound the right-hand side of (5.7) yields $\bar{\gamma}^t \geq \eta_2$, with η_2 given by (5.5). Thus, $\gamma^t = \min\{\eta_1, \bar{\gamma}^t\} \geq \min\{\eta_1, \eta_2\}$. ∎

Theorem 5.2 *Let β_0, β_1, β_2, and $\{(x^t, y^t, \omega^t, \theta^t, \gamma^t)\}_{t=0,1,\ldots}$ be generated by Algorithm 3.2 with A, B, β given by any of the choices cited in Lem. 4.2 (with $x \odot y$ in* (3.6) *replaced by $y \odot x$ and* (3.5) *replaced by* (3.12) *for choice (d) therein). Then,* (5.2) *and* (5.3) *hold for all t, where*

$$\mu^t := (1-\gamma^{t-1})(1-\theta^{t-1})\cdots(1-\gamma^0)(1-\theta^0).$$

If $\{(x^t, y^t)\}$ is bounded, then $\Sigma \neq \emptyset$. If $\Sigma \neq \emptyset$, then $\{(x^t, y^t)\}$ is bounded and, for any $(x^*, y^*) \in \Sigma$, we have $\gamma^t \geq \min\{\eta_1, \eta_2\}$ for all t, where

$$\eta_1 := (\beta_1 - \beta_0)/(\beta_1 + \sqrt{n}), \tag{5.8}$$

$$\eta_2 := \frac{.5(\beta_2 - \beta_1)\omega^0 \min(y^0)/\|Mx^0 + q - y^0\|_\infty}{n(1+\beta_0)\omega^0 + \langle x^0, y^0 \rangle + \langle x^*, y^0 \rangle + \langle x^0, y^* \rangle}. \tag{5.9}$$

Proof. The proof is nearly identical to that of Thm. 5.1. The only part that may not be evident is $\gamma^t \geq \min\{\eta_1, \eta_2\}$, which we now show. By using an argument similar to that used in the proof of Thm. 5.1, we have that (3.5) holds whenever $\gamma \leq \eta_2$, with $\omega := (1-\gamma)\omega^t$ and $(x,y) := (x^t, y^t)$. We also have from using the triangle inequality that, whenever $\gamma \leq \eta_1$,

$$\begin{aligned}
\phi_{(1-\gamma)\omega^t}(x^t, y^t) &= \|x^t \odot y^t - (1-\gamma)\omega^t e\|/(1-\gamma)\omega^t \\
&\leq (\|x^t \odot y^t - \omega^t e\| + \gamma\omega^t\|e\|)/(1-\gamma)\omega^t \\
&= (\phi_{\omega^t}(x^t, y^t) + \gamma\sqrt{n})/(1-\gamma) \leq \beta_0 + \gamma\sqrt{n}/(1-\gamma) \leq \beta_1,
\end{aligned}$$

where the last inequality follows from (5.2) and the definition of η_1. ∎

Thus, if Σ is nonempty, Algorithms 3.1 and 3.2 have global linear convergence of ratio $1 - \min\{\eta_1, \eta_2\}$ and can find a solution with accuracy ϵ in $O(\min\{\eta_1, \eta_2\}^{-1}|\log \epsilon|)$

iterations. It is easily seen from (3.1) and (5.8) that $\eta_1^{-1} = O(\sqrt{n})$, so it only remains to estimate η_2^{-1}. If (x^0, y^0, ω^0) is chosen so that η_2^{-1} is $O(\sqrt{n})$ (such as when in addition $Mx^0 + q - y^0 = 0$), then the iteration count is $O(\sqrt{n}|\log \epsilon|)$. If (x^0, y^0, ω^0) is chosen by the well-known choice of

$$x^0 = \rho_p e, \quad y^0 = \rho_d e, \quad \omega^0 = \rho_p \rho_d, \quad \rho_p \geq \frac{\|x^*\|_1}{n}, \quad \rho_d \geq \max\left\{\frac{\|y^*\|_1}{n}, \|\rho_p Me + q\|_\infty\right\},$$

where (x^*, y^*) is any element of Σ, then η_2 given by (5.5) satisfies

$$\eta_2^{-1} \leq 3(4 + \beta_1)n/(\beta_2 - \beta_1) = O(n),$$

so the iteration count is $O(n|\log \epsilon|)$, and similarly for η_2 given by (5.9). In the LCP case, it is known that Algorithm 3.1 with A, B, β given by (2.3), (3.10) has local quadratic convergence, assuming a strictly complementary solution exists [38, Thm. 4.1]; similar results can be shown for Algorithm 3.2, assuming that a nondegenerate strictly complementary solution exists. For the general problem however, local convergence analysis appears more difficult (see [14, 15, 19, 31]).

To detect whether Σ is empty, we can use the following known technique [12]: In Algorithm 3.1 and 3.2, for every t let ρ^t be the largest ρ such that

$$\langle x^t, y^0\rangle + \langle y^t, x^0\rangle > \langle x^t, y^t\rangle/\mu^t + \langle x^0, y^0\rangle + (\|x^0\| + \|y^0\|)\rho.$$

Then, by (5.1), there is no $(x^*, y^*) \in \Sigma$ satisfying $\max\{\|x^*\|, \|y^*\|\} \leq \rho^t$. However, as noted in [2], $\|x^*\|$ could be super-exponential in the problem size.

References

[1] F. Alizadeh, Optimization over the positive-definite cone: interior point methods and combinatorial applications, in *Advances in Optimization and Parallel Computing*, P. Pardalos, ed., North-Holland, Amsterdam, 1992, 1-25.

[2] F. Alizadeh, Interior point methods in semidefinite programming with application to combinatorial optimization, *SIAM J. Optim.*, **5** (1995), 13-51.

[3] F. Alizadeh, J.-P. A. Haeberly, and M. L. Overton, Primal-dual interior point methods for semidefinite programming, Manuscript, Presented at the Math Programming Symposium, Ann Arbor, 1994.

[4] F. Alizadeh, J.-P. A. Haeberly, and M. L. Overton, Primal-dual interior point methods for semidefinite programming: convergence rates, stability and numerical results, Working paper, RUTCOR, Rutgers University, New Brunswick, May 1996.

[5] K. M. Anstreicher and M. Frampa, A long-step path following algorithm for semidefinite programming problems, Working paper, Department of Management Sciences, University of Iowa, Iowa City, January 1996.

[6] L. Faybusovich, Semi-definite programming: a path-following algorithm for a linear-quadratic functional, Report, Department of Mathematics, University of Notre Dame, Notre Dame, 1995; to appear in *SIAM J. Optimization.*

[7] R. M. Freund, Complexity of an algorithm for finding an approximate solution of a semidefinite program with no regularity assumption, Working paper OR 302-94, Operations Research Center, Massachusetts Institute of Technology, Cambridge, December 1994.

[8] C. C. Gonzaga, The largest step path following algorithm for monotone linear complementarity problems, Report 94-07, Faculty of Technical Mathematics and Informatics, Delft University, Delft, January 1994; to appear in *Math. Prog.*

[9] C. Helmberg, F. Rendl, R. J. Vanderbei, and H. Wolkowicz, An interior-point method for semidefinite programming, *SIAM J. Optim.*, **6** (1996), 342-361.

[10] R. A. Horn, and C. R. Johnson, *Matrix Analysis*, Cambridge Press, Cambridge, 1985.

[11] F. Jarre, An interior-point method for minimizing the maximum eigenvalue of a linear combination of matrices, *SIAM J. Control Optim.*, **31** (1993), 1360-1377.

[12] M. Kojima, N. Megiddo, and S. Mizuno, A primal-dual infeasible-interior-point algorithm for linear programming, *Math. Prog.*, **61** (1993), 263-280.

[13] M. Kojima, S. Shindoh, and S. Hara, Interior-point methods for the monotone semidefinite linear complementarity problems, Research Report 282, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo, April 1994; revised October 1995; to appear in *SIAM J. Optim.*

[14] M. Kojima, M. Shida, and S. Shindoh, Local convergence of predictor-corrector infeasible interior-point algorithms for SDPs and SDLCPs, Research Report 306, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo, December 1995. This is a revised version of Research Report 305, titled Global and local convergence of predictor-corrector infeasible interior-point algorithms for semidefinite programs.

[15] M. Kojima, M. Shida, and S. Shindoh A predictor-corrector interior-point algorithms for the semidefinite linear complementarity problem using the

Alizadeh-Haeberly-Overton search direction, Research Report 311, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo, January 1996; revised March 1996.

[16] M. Kojima, M. Shida, and S. Shindoh A note on the Nesterov-Todd and the Kojima-Shindo-Hara search directions in semidefinite programming, Research Report 313, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo, April 1996.

[17] C.-J. Lin and R. Saigal, A predictor-corrector method for semi-definite linear programming, Working paper, Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, October 1995; revised December 1995.

[18] C.-J. Lin and R. Saigal, An infeasible start predictor-corrector method for semi-definite linear programming, Working paper, Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, December 1995.

[19] Z.-Q. Luo, J. F. Sturm and S. Zhang, Superlinear convergence of a symmetric primal-dual path following algorithm for semidefinite programming, Report 9607/A, Econometric Institute, Erasmus University, Rotterdam, January 1996.

[20] K. A. McShane, Superlinearly convergent $O(\sqrt{n}L)$-iteration interior point algorithms for linear programming and the monotone linear complementarity problem, *SIAM J. Optim.*, **4** (1994), 247–261.

[21] S. Mizuno, A. Yoshise, and T. Kikuchi, Practical polynomial time algorithms for linear complementarity problems, *J. Oper. Res. Soc. Japan*, **32** (1989), 75-92.

[22] S. Mizuno, M. J. Todd, and Y. Ye, On adaptive step primal-dual interior-point algorithms for linear programming, *Math. Oper. Res.*, **18** (1993), 964-981.

[23] S. Mizuno, Polynomiality of infeasible interior point algorithms for linear programming, *Math. Prog.*, **67** (1994), 109-119.

[24] R. D. C. Monteiro, Primal-dual path following algorithms for semidefinite programming, Working paper, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, September 1995.

[25] R. D. C. Monteiro and Y. Zhang, A unified analysis for a class of path-following primal-dual interior-point algorithms for semidefinite programming, Working paper, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, June 1996.

[26] Y. E. Nesterov and A. S. Nemirovskii, *Interior Point Polynomial Methods in Convex Programming: Theory and Applications*, SIAM, Philadelphia, 1993.

[27] Y. E. Nesterov and M. J. Todd, Self-scaled cones and interior-point methods in nonlinear programming, Working paper, CORE, Catholic University of Louvain, Louvain-la-Neuve, April 1994.

[28] Y. E. Nesterov and M. J. Todd, Primal-dual interior-point methods for self-scaled cones, Technical report 1125, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, 1995; revised February 1996.

[29] F. Potra and R. Sheng, A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming, Reports on Computational Mathematics 78, Department of Mathematics, University of Iowa, Iowa City, October 1995.

[30] F. Potra and R. Sheng, Homogeneous interior-point algorithms for semidefinite programming, Reports on Computational Mathematics 82, Department of Mathematics, University of Iowa, Iowa City, November 1995.

[31] F. Potra and R. Sheng, Superlinear convergence of interior-point algorithms for semidefinite programming, Reports on Computational Mathematics 86, Department of Mathematics, University of Iowa, Iowa City, April 1996.

[32] M. Shida, S. Shindoh, and M. Kojima, Existence of search directions in interior-point-algorithms for the SDP and the monotone SDLCP, Research Report 310, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo, January 1996.

[33] M. Shida and S. Shindoh, Monotone semidefinite complementarity problems, Research Report 312, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo, March 1996; revised April 1996.

[34] J. F. Sturm and S. Zhang, Symmetric primal-dual path following algorithms for semidefinite programming, Report 9554/A, Econometric Institute, Erasmus University, Rotterdam, November 1995; revised February 1996.

[35] M. J. Todd, Private communication, May 1996.

[36] M. J. Todd, K. C. Toh, and R. H. Tütüncü, On the Nesterov-Todd direction in semidefinite programming, Technical report, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, March 1996; revised May 1996.

[37] P. Tseng, Complexity analysis of a linear complementarity algorithm based on a Lyapunov function, *Math. Prog.*, **53** (1992), 297-306.

[38] P. Tseng, Simplified analysis of an $O(nL)$-iteration infeasible predictor-corrector path following method for monotone LCP, in *Recent Trends in Optimization Theory and Applications*, R. P. Agarwal, ed., World Scientific, Singapore, 1995, 423-434.

[39] L. Vandenberghe and S. Boyd, A primal-dual potential reduction method for problems involving matrix inequalities, *Math. Prog.* **69** (1995), 205-236.

[40] Y. Zhang, On extending primal-dual interior-point algorithms from linear programming to semidefinite programming, Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, October 1995.