

A One-Factorization Predictor–Corrector Long-Step Arc-Search Method and a Curvature-Amplified Variant for Semidefinite Programming with a Homogeneous Self-Dual Embedding

Makoto Yamashita*

Yaguang Yang†

Abstract

In a predictor–corrector arc-search method, a point on the predictor arc is first selected and a corrector is then computed at that point. Since the Karush–Kuhn–Tucker (KKT) matrix changes with the selected point, this correction may require a second factorization in each iteration. We propose a one-factorization arc-search method (OFAS) for semidefinite programming (SDP) that avoids this second factorization, retains the corrector after the arc step, and uses only one KKT factorization per iteration. The key is a three-term decomposition of the complementarity mismatch at the selected point on the arc, which allows the correction to be assembled from linear-system solutions using the factorization at the current iterate. We also propose a curvature-amplified variant (CA-OFAS) that amplifies the second-order arc term. The proposed methods are formulated with a homogeneous self-dual embedding. For a wide long-step neighborhood, we prove that both methods reduce the homogeneous complementarity and residuals below ε in $O(n \log(1/\varepsilon))$ iterations for an n -by- n SDP matrix variable. Numerical experiments show that the one-factorization structure and the correction after the arc step are associated with reductions in computation time. Compared with the Mehrotra-type method, CA-OFAS reduces the geometric-mean computation time by 17% on SDPLIB and 13% on neural network verification SDP problems.

Keywords: Semidefinite programming · Interior-point method · Arc-search · One-factorization · Homogeneous self-dual embedding

1 Introduction

Semidefinite programs (SDPs) arise in control, combinatorial optimization, statistical modeling, and convex relaxations for nonconvex problems [1, 2]. Primal–dual interior-point methods remain useful when high-accuracy solutions are required, as illustrated by widely used implementations such as SDPA, SDPT3, and SeDuMi [3–5]. Recent developments also include regularized SDP interior-point algorithms and general conic solvers built around homogeneous embeddings [6, 7]. Their cost is strongly influenced by the linear algebra for the Newton/Karush–Kuhn–Tucker (KKT) equations. For SDP in particular, a Schur-complement or KKT factorization can dominate an iteration. This motivates algorithms that control the number of matrix factorizations as well as the iteration count [8].

Homogeneous self-dual (HSD) embeddings let primal–dual interior-point iterations start without first constructing a feasible primal–dual pair. Self-dual LP formulations go back to Goldman and Tucker [9], and Ye–Todd–Mizuno later gave an HSD interior-point algorithm for LP [10]. The embedding combines the primal and dual systems in an augmented formulation; after normalization, a limiting solution yields either an optimizer of the original problem or an infeasibility certificate.

*Corresponding author. Department of Mathematical and Computing Science, Institute of Science Tokyo, 2-12-1-W8-29 Ookayama, Meguro-ku, Tokyo 152-8550, Japan. E-mail: Makoto.Yamashita@comp.isct.ac.jp. ORCID: 0000-0002-8409-036X

†Independent researcher, Rockville, MD 20850, USA. E-mail: yaguang.yang@verizon.net. ORCID: 0000-0002-2943-9389

Homogeneous embeddings also underlie modern conic solvers such as Clarabel [7], where symmetric cones are handled with Nesterov–Todd (NT) scaling in a primal–dual interior-point framework [11, 12].

Arc-search interior-point methods incorporate local curvature by replacing a straight predictor with an arc when tracking the central path or an analogous path for infeasible iterates. At the current iterate W_k , denote the first two derivatives of a local path by $\dot{W}_k$ and $\ddot{W}_k$. With α as the arc angle, the predictor has the familiar form

$$W_k(\alpha) = W_k - \dot{W}_k \sin \alpha + \ddot{W}_k(1 - \cos \alpha).$$

Arc-search methods were first developed as polynomial-time algorithms for convex quadratic programming and LP [13, 14], and were later extended to infeasible interior-point methods, symmetric-cone optimization, and SDP [15–20]. Other studies include a computationally efficient arc-search method for infeasible LP [21], an extension to general convex optimization [22], a method combined with Nesterov’s restarting strategy [23], an extension to nonlinear constrained problems [24], and a method that solves Newton systems inexactly [25]. See also Yang [26] for further developments of arc-search methods. We are not aware of an earlier arc-search analysis carried out directly in an HSD formulation.

Some predictor–corrector arc-search methods first select a point on the arc and then compute a separate corrector at that point, whereas other methods do not perform this additional correction [15, 21, 23]. In the former case, the KKT coefficient matrix changes with the scaling at the selected point and may require another factorization. In the latter case, the coefficient matrix at the current iterate can be reused. A detailed comparison of these computational structures is given in Section 2.4.

Motivated by this factorization cost, we construct a one-factorization arc-search method (OFAS). It keeps the corrector applied after the arc point is selected, but factors the KKT matrix only at the current iterate. The central device is an exact decomposition of the selected-point complementarity mismatch into three terms in the current NT coordinates. We solve for the associated correction bases with the current KKT factorization and combine them only after the candidate values of the arc angle α and the centering-recovery parameter ρ are fixed. Candidate changes therefore do not trigger a new KKT factorization. In OFAS, the single factorization serves five algorithm-specific right-hand sides. HSD and predictor–corrector methods commonly reuse a factorization across several right-hand sides, as does Clarabel [7, 10]; here the three-term decomposition extends that reuse to a post-arc corrector whose matrix would otherwise vary with the selected point.

We further introduce a curvature-amplified one-factorization arc-search method (CA-OFAS), which selectively amplifies the arc term corresponding to the second-order direction. This modification uses curvature information more aggressively while keeping the same one-factorization structure as OFAS. The amplified candidates are constructed using the directions and correction bases already computed in OFAS and require no additional matrix factorization or linear-system solve. If an amplified candidate is not accepted, the method uses an OFAS candidate. Thus, CA-OFAS tests curvature-amplified candidates while retaining the worst-case iteration bound established for OFAS.

The contributions are threefold.

- (i) We develop OFAS for SDP, which retains a separate corrector after selecting the point on the arc while requiring only one KKT matrix factorization per iteration. We also construct CA-OFAS, a curvature-amplified method with the same one-factorization structure.
- (ii) Using a wide long-step neighborhood, we establish the well-definedness of OFAS, finite termination of the candidate selection, and the existence of a uniform admissible window at each iteration. The resulting iteration bound is $O(\nu \log(1/\varepsilon))$ for reducing the homogeneous complementarity and residuals below ε , where $\nu = n + 1$ is the rank of the HSD cone. Under Assumption 2.2, we also study the projectively normalized HSD sequence and prove that its distance to the primal–dual optimal set tends to zero. Since CA-OFAS has an OFAS fallback candidate, it inherits the same iteration bound.

- (iii) We implement OFAS and CA-OFAS in the HSD framework of Clarabel.jl and evaluate them on SDPLIB and neural network verification SDP (NNV-SDP) problems. First, we compare OFAS with Two-factorization, which refactorizes the KKT coefficient matrix after selecting the point on the arc, and Corrector-free, which omits the separate corrector at that point. These comparisons examine the contributions of factorization reuse and the corrector. Among matched problems reaching the high-accuracy termination criterion, OFAS reduced the computation time relative to Two-factorization by about 44% on SDPLIB and about 38% on NNV-SDP, and relative to Corrector-free by about 37% and 42%, respectively. We then compare CA-OFAS with OFAS and observe that curvature amplification reduced the computation time by about 4% and 19%, respectively. Finally, we compare the methods with a Mehrotra-type method and the Liu-type method based on Liu–Liu (2012) [27, 28]. Among matched problems reaching the high-accuracy termination criterion, CA-OFAS reduced the computation time relative to the Mehrotra-type method by about 17% on SDPLIB and about 13% on NNV-SDP, and relative to the Liu-type method by about 20% and 12%, respectively.

Section 2 develops the HSD-SDP setting, NT scaling, the wide long-step neighborhood, and the arc equations used later. Section 3 then introduces OFAS and CA-OFAS, while Section 4 establishes candidate-search termination, the iteration bound, and projective normalization for the solvable case. Numerical results are reported in Section 5, followed by the conclusions in Section 6.

2 SDP with an HSD Embedding and the Foundations of Long-Step Arc-Search

2.1 SDP and the HSD Formulation

Let $n, m \in \mathbb{N}$. Write $\mathbb{S}^n$ for the space of real symmetric $n \times n$ matrices; its positive semidefinite and positive definite cones are denoted by $\mathbb{S}_+^n$ and $\mathbb{S}_{++}^n$. For $U, V \in \mathbb{S}^n$, the relations $U \succeq V$ and $U \succ V$ mean $U - V \in \mathbb{S}_+^n$ and $U - V \in \mathbb{S}_{++}^n$, respectively. The problem data are $C \in \mathbb{S}^n$, $b \in \mathbb{R}^m$, and a linear map $\mathcal{A} : \mathbb{S}^n \rightarrow \mathbb{R}^m$ with adjoint $\mathcal{A}^*$. We write $U \bullet V := \text{tr}(UV)$ for the matrix inner product and use I and O for the identity and zero matrices of order n . Unless a different norm is indicated, vector norms are Euclidean and symmetric-matrix norms are Frobenius norms.

We consider the following primal SDP in equality standard form and its dual:

$$p^* := \min_{X \in \mathbb{S}^n} \{C \bullet X \mid \mathcal{A}(X) = b, X \succeq O\}, \quad d^* := \max_{y \in \mathbb{R}^m, S \in \mathbb{S}^n} \{b^T y \mid \mathcal{A}^*(y) + S = C, S \succeq O\}.$$

Assumption 2.1 (Surjectivity). The map $\mathcal{A}$ is surjective.

Assumption 2.1 is equivalent to linear independence of the equality constraints. In the setting considered here, linearly dependent equality constraints can be removed in preprocessing, and we assume that this preprocessing has been applied.

We embed the original primal–dual SDP into a homogeneous self-dual (HSD) system. The basic idea of the HSD formulation goes back to Ye et al. [10] for linear programming. Self-dual skew-symmetric embeddings and homogeneous interior-point methods for semidefinite programming were studied by de Klerk et al. [29], Potra and Sheng [30]. When the quadratic objective term is absent, the homogeneous embedding used in Clarabel reduces to the standard homogeneous self-dual embedding [7]. Let the HSD variable be $W = (X, y, S, \tau, \kappa)$. We use the following system:

$$\begin{aligned} \mathcal{A}(X) - b\tau &= 0, & \mathcal{A}^*(y) + S - C\tau &= 0, & C \bullet X - b^T y + \kappa &= 0, \\ X \succeq O, & S \succeq O, & \tau \geq 0, & \kappa \geq 0. \end{aligned} \tag{1}$$

Here, $\tau \in \mathbb{R}$ homogenizes the primal–dual variables, and $\kappa \in \mathbb{R}$ is an auxiliary variable associated with the primal–dual gap. We refer to this embedded system as the HSD-SDP formulation. The pairs (X, τ) and (S, κ) are primal and dual cone variables over $\mathbb{S}_+^n \times \mathbb{R}_+$, and their complementarity is measured by $X \bullet S + \tau\kappa$. If $\tau > 0$, the original SDP variables are recovered by projective

normalization, $(\bar{X}, \bar{y}, \bar{S}) := (X/\tau, y/\tau, S/\tau)$. For an exact HSD solution, the normalized point satisfies the primal and dual equalities of the original SDP. At an exact HSD solution with $\tau = 0$ and $\kappa > 0$, we have $C \bullet X - b^T y = -\kappa < 0$, so at least one of $C \bullet X < 0$ and $b^T y > 0$ holds. The former certifies dual infeasibility, while the latter certifies primal infeasibility. The present complexity analysis focuses on the solvable case with $\tau > 0$; iteration bounds for certificate recovery are outside its scope. An infeasible interior-point method need not satisfy the equality part of (1) exactly at every iteration. We therefore define the residuals $R_p(W) := \mathcal{A}(X) - b\tau$, $R_d(W) := \mathcal{A}^*(y) + S - C\tau$, and $R_g(W) := C \bullet X - b^T y + \kappa$. Let the residual space be $\mathcal{F} := \mathbb{R}^m \times \mathbb{S}^n \times \mathbb{R}$. We also let $0_m \in \mathbb{R}^m$ be the zero vector and define the zero element of $\mathcal{F}$ by $O_{\mathcal{F}} := (0_m, O, 0)$. We collect the three residuals in the HSD residual map

$$\Psi(W) := (R_p(W), R_d(W), R_g(W)) \in \mathcal{F}.$$

and equip $\mathcal{F}$ with the norm $\|(u, V, \chi)\|_{\mathcal{F}} := (\|u\|^2 + \|V\|_F^2 + |\chi|^2)^{1/2}$. The map Ψ is linear, and $\Psi(W) = O_{\mathcal{F}}$ is equivalent to the three linear equalities in the HSD system. In the interior-point iterations below, we call an HSD point $W = (X, y, S, \tau, \kappa)$ *strictly interior* if $X, S \succ O$ and $\tau, \kappa > 0$.

For the iteration-complexity analysis and Algorithms 1–2, we start from the standard HSD initial point $W_0 := (I, 0, I, 1, 1)$. This point lies in the interior of the cone, but in general does not satisfy $\Psi(W_0) = O_{\mathcal{F}}$. We call W_0 the *standard initial point*.

The next assumption is needed only for the analysis of normalized approximate solutions of the original SDP. Define the primal, dual, and joint optimal sets by

$$\begin{aligned} \mathcal{P}^* &:= \{X \succeq O : \mathcal{A}(X) = b, C \bullet X = p^*\}, \\ \mathcal{D}^* &:= \{(y, S) : \mathcal{A}^*(y) + S = C, S \succeq O, b^T y = d^*\}, \\ \mathcal{Z}^* &:= \{(X, y, S) : X \in \mathcal{P}^*, (y, S) \in \mathcal{D}^*\}. \end{aligned}$$

Assumption 2.2 (Solvability and bounded optimal solution sets of the primal–dual problem). Both problems attain finite optimal values and strong duality holds, so $d^* = p^* \in \mathbb{R}$. The sets $\mathcal{P}^*$ and $\mathcal{D}^*$ are bounded.

This assumption is not used in the homogeneous long-step analysis itself. It is needed only when an HSD point is quantitatively converted into a normalized approximate solution of the original SDP. More specifically, Assumption 2.1 is needed for nonsingularity of the KKT operator at the current iterate, whereas Assumption 2.2 is used in Section 4.5 to exclude nonzero directions that would make the primal or dual optimal solution set unbounded and to control the projective scale. Therefore, the homogeneous iteration bounds in Theorem 4.7 and Corollary 4.8 do not require Assumption 2.2. For the projective normalization introduced immediately after (1), the residual identities are $\mathcal{A}(\bar{X}) - b = R_p(W)/\tau$ and $\mathcal{A}^*(\bar{y}) + \bar{S} - C = R_d(W)/\tau$ whenever $\tau > 0$. In Section 4.5, under Assumption 2.2, we evaluate the relative scales of τ_k and the homogeneous progress measure along the sequence generated by Algorithms 1–2. We then show that the normalized primal residual, dual residual, complementarity, and duality gap converge to zero.

2.2 NT Scaling and Complementarity Geometry

Throughout this subsection, let $\mathbb{R}_{++} := \{t \in \mathbb{R} : t > 0\}$, and let $\lambda(A) := (\lambda_1(A), \dots, \lambda_n(A))$ denote the eigenvalues of $A \in \mathbb{S}^n$, ordered as $\lambda_1(A) \leq \dots \leq \lambda_n(A)$. We use the direct-product Euclidean Jordan algebra $\mathcal{E} := \mathbb{S}^n \times \mathbb{R}$ for one positive semidefinite (PSD) block and the scalar HSD block. For $U = (U_X, u)$ and $V = (V_X, v)$ in $\mathcal{E}$, define the Jordan product by

$$U \circ V := \left(\frac{U_X V_X + V_X U_X}{2}, uv \right), \quad U \circ V = V \circ U.$$

The identity and zero elements of $\mathcal{E}$ are $\bar{I} := (I, 1)$ and $\bar{O} := (O, 0)$, respectively, and the inner product is $\langle U, V \rangle_{\mathcal{E}} := U_X \bullet V_X + uv$. The induced norm is written as $\|U\|_{\mathcal{E}} := \sqrt{\langle U, U \rangle_{\mathcal{E}}}$. When the arguments clearly belong to $\mathcal{E}$, we omit the subscript $\mathcal{E}$ and write $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$. We also define

$\text{Tr}_{\mathcal{E}}(U) := \langle \bar{I}, U \rangle_{\mathcal{E}}$ and $\lambda_{\min}(U) := \min\{\lambda_1(U_X), u\}$. For simplicity, we describe the method for a single PSD block. The same construction applies to multiple PSD blocks over a direct-product cone, as briefly discussed in Remark 4.12.

Since $\mathbb{S}_+^n \times \mathbb{R}_+$ is a symmetric cone, NT scaling applies to this product cone [11, 12]. We set $\nu := n + 1$ and define the average complementarity by $\mu(W) := (X \bullet S + \tau\kappa)/\nu$. For an HSD interior point W with $X, S \succ O$, $\tau > 0$, and $\kappa > 0$, define $\lambda(X, S) := \lambda(X^{1/2}SX^{1/2})$ and $h(W) := (\lambda(X, S), \tau\kappa) \in \mathbb{R}_{++}^\nu$, and denote the i th component of $h(W)$ by $h_i(W)$. The componentwise average of $h(W)$ is $\mu(W)$.

For the homogeneous analysis, we use the progress measure

$$\mathcal{Q}(W) := \max\{\mu(W), \|R_p(W)\|, \|R_d(W)\|, |R_g(W)|\}$$

which combines the average complementarity and the HSD residuals. When $\tau > 0$, we instead measure the KKT accuracy for the original SDP by

$$\mathcal{Q}_{\text{SDP}}(W) := \max\left\{\frac{\|R_p(W)\|}{\tau}, \frac{\|R_d(W)\|}{\tau}, \frac{X \bullet S}{\tau^2}, \frac{|C \bullet X - b^T y|}{\tau}\right\}.$$

This quantity measures the primal residual, dual residual, complementarity, and duality gap of the normalized point $(X/\tau, y/\tau, S/\tau)$. To recover a solution of the original SDP, we use the stopping condition $\mathcal{Q}_{\text{SDP}}(W) \leq \delta$ for a tolerance $\delta > 0$. The quantity $\mathcal{Q}(W)$ is used for the homogeneous progress analysis in Section 4. Section 4.5 shows that the normalized stopping condition is satisfied after finitely many iterations.

For the PSD cone, we use the same NT scaling as in Clarabel.jl [7]. For an HSD interior point $\widehat{W} = (\widehat{X}, \widehat{y}, \widehat{S}, \widehat{\tau}, \widehat{\kappa})$, let $M(\widehat{W})$ denote the inverse scaling matrix maintained at that point. For the PSD block, an explicit symmetric NT choice is

$$M(\widehat{W}) = \left[\widehat{X}^{-1/2} (\widehat{X}^{1/2} \widehat{S} \widehat{X}^{1/2})^{1/2} \widehat{X}^{-1/2} \right]^{1/2},$$

where all matrix square roots are the principal positive definite square roots. Indeed, with $G := M(\widehat{W})^T M(\widehat{W})$, this choice gives $G \widehat{X} G = \widehat{S}$. Multiplying this identity by $M(\widehat{W})^{-T}$ from the left and by $M(\widehat{W})^{-1}$ from the right gives the equality of the primal and dual variables in the scaled coordinates in (2). Thus,

$$M(\widehat{W}) \widehat{X} M(\widehat{W})^T = M(\widehat{W})^{-T} \widehat{S} M(\widehat{W})^{-1} =: Q_X^{\text{NT}}(\widehat{W}) \succ O \quad (2)$$

holds. We also define $\xi(\widehat{W}) := \sqrt{\widehat{\kappa}/\widehat{\tau}}$. For a generic HSD vector $V = (V_X, v_y, V_S, v_\tau, v_\kappa)$, define its NT coordinates with respect to the reference interior point $\widehat{W}$ by

$$\begin{aligned} P(V; \widehat{W}) &:= (M(\widehat{W}) V_X M(\widehat{W})^T, \xi(\widehat{W}) v_\tau), \\ R(V; \widehat{W}) &:= (M(\widehat{W})^{-T} V_S M(\widehat{W})^{-1}, \xi(\widehat{W})^{-1} v_\kappa). \end{aligned} \quad (3)$$

In particular,

$$Q^{\text{NT}}(\widehat{W}) := P(\widehat{W}; \widehat{W}) = R(\widehat{W}; \widehat{W}) = (Q_X^{\text{NT}}(\widehat{W}), \sqrt{\widehat{\tau}\widehat{\kappa}}).$$

We represent complementarity in the NT coordinates of the same reference point by the quadratic map

$$\Phi(V; \widehat{W}) := P(V; \widehat{W}) \circ R(V; \widehat{W}).$$

At the current iterate W_k , we use the abbreviations

$$M_k := M(W_k), \quad \xi_k := \xi(W_k), \quad Q_k^{\text{NT}} := Q^{\text{NT}}(W_k).$$

Candidate points are evaluated in the NT coordinates determined by W_k .

The standard initial point W_0 satisfies $\mu(W_0) = 1$ and $\Phi(W_0; W_0) = \bar{I}$. Let $\mathcal{K} := \mathbb{S}_+^n \times \mathbb{R}_+$, and denote the central-path parameter by μ_c . By the standard central-path theory for HSD-SDP

[11, 12, 30], under Assumption 2.1, for every $0 < \mu_c \leq 1$ there exists a unique interior point $W^c(\mu_c)$ satisfying

$$\Psi(W^c(\mu_c)) = \mu_c \Psi(W_0), \quad \Phi(W^c(\mu_c); W^c(\mu_c)) = \mu_c \bar{I}.$$

The second equation is the standard centrality condition based on the Jordan product on $\mathcal{K}$. For the PSD block, it is equivalent to $S = \mu_c X^{-1}$, and for the scalar HSD block, to $\kappa = \mu_c/\tau$. We call $\{W^c(\mu_c) : 0 < \mu_c \leq 1\}$ the HSD central path. Along this path, $\mu(W^c(\mu_c)) = \mu_c$. We introduce the central path only to describe the geometry behind the local tracking curve. Its use here does not impose a solvability assumption for the original SDP on the homogeneous iteration bound in Section 4. Projective normalization for the original SDP is treated in Section 4.5.

Since $P(\cdot; \widehat{W})$ and $R(\cdot; \widehat{W})$ are linear, $\Phi(\cdot; \widehat{W})$ is quadratic. For an HSD interior point $\widehat{W}$ and HSD variables or directions $V^{(1)}, V^{(2)}$, define the associated cross term by

$$\mathcal{B}(V^{(1)}, V^{(2)}; \widehat{W}) := P(V^{(1)}; \widehat{W}) \circ R(V^{(2)}; \widehat{W}) + P(V^{(2)}; \widehat{W}) \circ R(V^{(1)}; \widehat{W}).$$

The map $\mathcal{B}(\cdot, \cdot; \widehat{W})$ is symmetric and bilinear, and

$$\Phi(V^{(1)} + V^{(2)}; \widehat{W}) = \Phi(V^{(1)}; \widehat{W}) + \mathcal{B}(V^{(1)}, V^{(2)}; \widehat{W}) + \Phi(V^{(2)}; \widehat{W}).$$

For an HSD direction D , define the linearized HSD KKT operator at the reference interior point $\widehat{W}$ by

$$\mathcal{L}(\widehat{W})[D] := (\Psi(D), \mathcal{B}(\widehat{W}, D; \widehat{W})). \quad (4)$$

2.3 Current-Iterate KKT System, Predictor Arc, and Wide Long-Step Neighborhood

A general infeasible interior point W_k need not lie on the HSD central path. Following the infeasible arc-search literature, we introduce a local path on which the HSD residual map and the NT-coordinate complementarity are multiplied by the same scalar parameter [15, 23]. Its first two derivatives determine the predictor arc.

Let an HSD direction be $D = (D_X, d_y, D_S, d_\tau, d_\kappa)$. At the current iterate W_k , (3) becomes

$$\begin{aligned} P(D; W_k) &= (M_k D_X M_k^T, \xi_k d_\tau), \\ R(D; W_k) &= (M_k^{-T} D_S M_k^{-1}, \xi_k^{-1} d_\kappa). \end{aligned} \quad (5)$$

Since $P(W_k; W_k) = R(W_k; W_k) = Q_k^{\text{NT}}$ at the current iterate, $\mathcal{B}(W_k, D; W_k) = Q_k^{\text{NT}} \circ \{P(D; W_k) + R(D; W_k)\}$. Therefore, the specialization of (4) to $\widehat{W} = W_k$ is

$$\mathcal{L}(W_k)[D] = \begin{pmatrix} \mathcal{A}(D_X) - b d_\tau \\ \mathcal{A}^*(d_y) + D_S - C d_\tau \\ C \bullet D_X - b^T d_y + d_\kappa \\ Q_k^{\text{NT}} \circ \{P(D; W_k) + R(D; W_k)\} \end{pmatrix} \in \mathcal{F} \times \mathcal{E}. \quad (6)$$

Thus, (6) gives the matrix representation of (4) at the current iterate. By Lemma 3.3, $\mathcal{L}(W_k)$ is nonsingular under Assumption 2.1 and strict interiority of W_k . In this paper, “one factorization” means that this matrix is factorized only once in each interior-point iteration and that the resulting factorization is reused for multiple right-hand sides.

Using the NT scaling determined by W_k , consider a local curve $\widehat{W}_k(t)$ near $t = 1$ satisfying

$$\widehat{W}_k(1) = W_k, \quad \Psi(\widehat{W}_k(t)) = t \Psi(W_k), \quad \Phi(\widehat{W}_k(t); W_k) = t \Phi(W_k; W_k). \quad (7)$$

The Jacobian with respect to W in (7) at $(W, t) = (W_k, 1)$ is $\mathcal{L}(W_k)$. Since this operator is nonsingular, the implicit function theorem gives a unique smooth curve locally around W_k . This curve is distinct from the HSD central path. Along it, the parameter t scales both the residual map and the NT-coordinate complementarity from their values at the current iterate.

Let $\dot{W}_k$ and $\ddot{W}_k$ be the first and second derivatives at $t = 1$, respectively. We call them the first-order and second-order directions. Differentiating (7) once and using the linearity of Ψ and the quadratic structure of Φ gives

$$\mathcal{L}(W_k)[\dot{W}_k] = (\Psi(W_k), \Phi(W_k; W_k)). \quad (8)$$

Differentiating twice, the right-hand side has zero second derivative, and only the quadratic term of Φ remains. Thus,

$$\mathcal{L}(W_k)[\ddot{W}_k] = (O_{\mathcal{F}}, -2\Phi(\dot{W}_k; W_k)). \quad (9)$$

Hence, $\dot{W}_k$ and $\ddot{W}_k$ are obtained from two right-hand sides with the same KKT coefficient matrix.

Using these directions, we define the predictor arc by

$$W_k(\alpha) = W_k - \dot{W}_k \sin \alpha + \ddot{W}_k (1 - \cos \alpha). \quad (10)$$

We have $W_k(0) = W_k$, the tangent direction at $\alpha = 0$ is $-\dot{W}_k$, and the second derivative is $\ddot{W}_k$. The minus sign reflects that the arc follows the local tracking curve as its parameter t decreases from 1. We write $W_k(\alpha_k)$ for the point corresponding to an accepted angle α_k .

For the long-step analysis, we use the wide long-step neighborhood

$$\mathcal{N}_{\text{long}}(\gamma) := \left\{ W = (X, y, S, \tau, \kappa) \mid \begin{array}{l} X \succ O, S \succ O, \tau > 0, \kappa > 0, \\ \min_i h_i(W) \geq \gamma \mu(W) \end{array} \right\}, \quad 0 < \gamma < 1,$$

which is independent of the particular scaling representation. This complementarity neighborhood does not require HSD feasibility and therefore contains infeasible interior points. We test candidate points in the NT coordinates of the current iterate W_k . At the current iterate, $\lambda_{\min}(\Phi(W_k; W_k)) = \min_i h_i(W_k)$. For a general candidate W , Lemma 4.2 shows that $\lambda_{\min}(\Phi(W; W_k)) \geq \gamma \mu(W)$ implies $W \in \mathcal{N}_{\text{long}}(\gamma)$. Thus, the candidate search can be performed in the current iterate's NT coordinates, while the analysis uses a scaling-independent neighborhood.

2.4 Comparison of Corrector Structures in Existing Arc-Search Methods

To compare the computational structure of OFAS with existing methods, we first consider methods that rebuild the NT scaling at the accepted point on the predictor arc and compute the corrector there. In such methods, restoring centrality at the accepted point changes the coefficient matrix. In Algorithm 3.1 of Yang and Yamashita [15] and Algorithm 1 of Iida and Yamashita [23], the first-order and second-order directions are computed at the current iterate, whereas the corrector system uses the selected point on the arc in its coefficient matrix. Below, we give an HSD-SDP reference construction in NT coordinates that captures only this computational structure.

Assume that the accepted point $W_k(\alpha_k)$ is strictly interior. Following the centering target used in these LP methods, set $\mu_k^{\text{ref}} := (1 - \sin \alpha_k) \mu(W_k)$ and $T_k^{\text{ref}} := \mu_k^{\text{ref}} \bar{I}$. Define the centering deviation in the NT coordinates of the accepted point by $\Delta_k^{\text{ref}} := \Phi(W_k(\alpha_k); W_k(\alpha_k)) - T_k^{\text{ref}}$. A corrector $D_k^{\text{cor,ref}}$ that linearly removes this mismatch without changing the HSD residuals is obtained from

$$\Psi(D_k^{\text{cor,ref}}) = O_{\mathcal{F}}, \quad \mathcal{B}(W_k(\alpha_k), D_k^{\text{cor,ref}}; W_k(\alpha_k)) = -\Delta_k^{\text{ref}}.$$

This corrector uses $\mathcal{L}(W_k(\alpha_k))$, namely the KKT operator formed at the accepted point; in general $\mathcal{L}(W_k(\alpha_k)) \neq \mathcal{L}(W_k)$. Thus, with a direct linear solver, constructing the predictor arc requires one factorization of $\mathcal{L}(W_k)$, whereas the accepted-point corrector requires another factorization of $\mathcal{L}(W_k(\alpha_k))$.

Other methods incorporate centering information directly into the arc construction and do not solve a separate corrector system at the accepted point. For example, Yang [21] uses the same current-iterate coefficient matrix for the first-order and second-order directions, adds a centering term to the right-hand side for the second-order direction, and directly updates to a point on the resulting arc. Such a method can reuse the factorization at the current iterate for multiple right-hand sides, but it does not apply a separate post-arc corrector based on the complementarity mismatch at the accepted point. Reusing one coefficient-matrix factorization for several right-hand sides is standard. The key distinction is that, when a method uses an accepted-point corrector, updating the scaling at the accepted point changes the coefficient matrix.

3 Predictor–Corrector Arc-Search with Current-Iterate NT Scaling and Curvature Amplification

In this section, we construct OFAS and CA-OFAS in the current-iterate NT coordinates introduced in Section 2. We use Φ , $\mathcal{B}$, $\mathcal{L}(W_k)$, the first-order direction $\dot{W}_k$, the second-order direction $\ddot{W}_k$, and the predictor arc $W_k(\alpha)$ defined there. For a general candidate angle α , we write $\sigma := \sin \alpha$ and $\psi := 1 - \cos \alpha$ throughout this section.

The main idea of OFAS is to retain an explicit post-arc correction while expressing its candidate-dependent complementarity mismatch as an exact sum of three components that depend only on the current iterate W_k . We precompute the correction response to each component using the same KKT factorization at W_k . The dependence on the candidate parameters then appears only through scalar coefficients, so the correction can be reconstructed without refactorizing the KKT matrix at the accepted point. We first derive this three-term decomposition and the corresponding one-factorization correction basis. We then describe the corrected candidates, candidate selection, OFAS, and CA-OFAS, which uses the same basis.

3.1 Three-Term Decomposition of the Complementarity Mismatch on the Predictor Arc

Fix iteration k and define $U_k(\alpha) := W_k(\alpha) - W_k = \sigma(-\dot{W}_k) + \psi\ddot{W}_k$. The complementarity rows of (8) and (9), together with the predictor arc (10), give

$$\mathcal{B}(W_k, -\dot{W}_k; W_k) = -\Phi(W_k; W_k), \quad \mathcal{B}(W_k, \ddot{W}_k; W_k) = -2\Phi(-\dot{W}_k; W_k). \quad (11)$$

Together with the quadratic structure of $\Phi(\cdot; W_k)$, these identities give the following exact decomposition.

Lemma 3.1 (Exact quadratic decomposition of complementarity on the predictor arc). *For any α ,*

$$\begin{aligned} \Phi(W_k(\alpha); W_k) &= (1 - \sigma)\Phi(W_k; W_k) + \sigma\psi\mathcal{B}(-\dot{W}_k, \ddot{W}_k; W_k) \\ &\quad + \psi^2\{\Phi(\ddot{W}_k; W_k) - \Phi(-\dot{W}_k; W_k)\}. \end{aligned} \quad (12)$$

Proof. A quadratic expansion of $\Phi(W_k + U_k(\alpha); W_k)$, together with (11), gives

$$\Phi(W_k(\alpha); W_k) = (1 - \sigma)\Phi(W_k; W_k) + (\sigma^2 - 2\psi)\Phi(-\dot{W}_k; W_k) + \sigma\psi\mathcal{B}(-\dot{W}_k, \ddot{W}_k; W_k) + \psi^2\Phi(\ddot{W}_k; W_k).$$

Using $\sigma^2 = 2\psi - \psi^2$ gives (12). $\square$

For a centering-recovery parameter $\rho \in [0, 1]$, define the corrector complementarity target by

$$T_k(\alpha, \rho) := (1 - \sigma)\{(1 - \rho)\Phi(W_k; W_k) + \rho\mu(W_k)\bar{I}\} \quad (13)$$

and define the complementarity mismatch by

$$\Delta_k^{\text{comp}}(\alpha, \rho) := \Phi(W_k(\alpha); W_k) - T_k(\alpha, \rho). \quad (14)$$

Here, ρ specifies the fraction of the current centrality deviation to be corrected, while $1 - \sigma$ gives the complementarity reduction produced by the predictor arc.

Proposition 3.2 (Three-term decomposition of the complementarity mismatch). *For any α and $\rho \in [0, 1]$,*

$$\Delta_k^{\text{comp}}(\alpha, \rho) = \rho(1 - \sigma)\Delta_k^{\text{cen}} + \sigma\psi\Delta_k^{\text{mix}} + \psi^2\Delta_k^{\text{rem}}, \quad (15)$$

where

$$\begin{aligned} \Delta_k^{\text{cen}} &:= \Phi(W_k; W_k) - \mu(W_k)\bar{I}, \\ \Delta_k^{\text{mix}} &:= \mathcal{B}(-\dot{W}_k, \ddot{W}_k; W_k), \\ \Delta_k^{\text{rem}} &:= \Phi(\ddot{W}_k; W_k) - \Phi(-\dot{W}_k; W_k). \end{aligned} \quad (16)$$

Proof. Rearranging (13) gives $T_k(\alpha, \rho) = (1 - \sigma)\Phi(W_k; W_k) - \rho(1 - \sigma)\{\Phi(W_k; W_k) - \mu(W_k)\bar{I}\}$. Substituting this expression into (14) and applying Lemma 3.1 gives (15). $\square$

The three components in (15) correspond to the centrality deviation at the current iterate, the cross term between the first- and second-order directions, and the quadratic remainder, respectively. Thus, once W_k , $\dot{W}_k$, and $\ddot{W}_k$ are fixed, all three components are fixed. The candidate (α, ρ) enters only through the scalar coefficients $\rho(1 - \sigma)$, $\sigma\psi$, and ψ^2 .

3.2 One-Factorization Correction Basis and Correction-Response Map

Lemma 3.3 (Unique solvability of the NT KKT operator). *Under Assumption 2.1, the NT KKT operator $\mathcal{L}(\widehat{W})$ is nonsingular at any HSD interior point $\widehat{W}$.*

A proof is given in Appendix A.1.

For an HSD interior point $\widehat{W}$ and $\Delta \in \mathcal{E}$, define the correction-response map $\mathcal{G}(\Delta; \widehat{W})$ by

$$\mathcal{L}(\widehat{W})[\mathcal{G}(\Delta; \widehat{W})] = (O_{\mathcal{F}}, -\Delta). \quad (17)$$

By Lemma 3.3, this map is well defined and $\mathcal{G}(\cdot; \widehat{W})$ is linear. If $D = \mathcal{G}(\Delta; \widehat{W})$, then $\Psi(D) = O_{\mathcal{F}}$ and $\mathcal{B}(\widehat{W}, D; \widehat{W}) = -\Delta$. Below, we use it with $\widehat{W} = W_k$.

Define the unit-complementarity direction by

$$\mathcal{L}(W_k)[D_k^{\text{unit}}] = (O_{\mathcal{F}}, \bar{I}), \quad (18)$$

and define $D_k^{\text{mix}} := \mathcal{G}(\Delta_k^{\text{mix}}; W_k)$, $D_k^{\text{rem}} := \mathcal{G}(\Delta_k^{\text{rem}}; W_k)$, and $D_k^{\text{cen}} := \mathcal{G}(\Delta_k^{\text{cen}}; W_k)$. Each of D_k^{unit} , D_k^{cen} , D_k^{mix} , and D_k^{rem} is an HSD residual-invariant direction; equivalently, $\Psi(D_k^{\text{term}}) = O_{\mathcal{F}}$ for $\text{term} \in \{\text{unit}, \text{cen}, \text{mix}, \text{rem}\}$. For D_k^{mix} and D_k^{rem} , the right-hand sides in (17) can be formed directly from $\dot{W}_k$ and $\ddot{W}_k$ using (16). The centering direction D_k^{cen} can instead be recovered algebraically by the following lemma, so it requires no additional linear-system solve.

Lemma 3.4 (Algebraic reconstruction of the centering correction).

$$D_k^{\text{cen}} = -W_k + \dot{W}_k + \mu(W_k)D_k^{\text{unit}}. \quad (19)$$

Proof. Using (8) and $\mathcal{L}(W_k)[W_k] = (\Psi(W_k), 2\Phi(W_k; W_k))$, we obtain $\mathcal{G}(\Phi(W_k; W_k); W_k) = -W_k + \dot{W}_k$. Equation (18) also gives $D_k^{\text{unit}} = \mathcal{G}(-\bar{I}; W_k)$. The identity then follows from the linearity of $\mathcal{G}$ and $\Delta_k^{\text{cen}} = \Phi(W_k; W_k) - \mu(W_k)\bar{I}$. $\square$

For any (α, ρ) , define

$$D_k^{\text{cor}}(\alpha, \rho) := \rho(1 - \sigma)D_k^{\text{cen}} + \sigma\psi D_k^{\text{mix}} + \psi^2 D_k^{\text{rem}}.$$

By Proposition 3.2 and the linearity of $\mathcal{G}$,

$$\mathcal{L}(W_k)[D_k^{\text{cor}}(\alpha, \rho)] = (O_{\mathcal{F}}, -\Delta_k^{\text{comp}}(\alpha, \rho)). \quad (20)$$

Thus, $D_k^{\text{cor}}(\alpha, \rho)$ is also an HSD residual-invariant direction.

Once the KKT matrix has been factorized, solving systems with additional right-hand sides by reusing the existing factors is much less expensive than performing another factorization. At each iteration, we factorize the matrix representation of $\mathcal{L}(W_k)$ only once and reuse its factors for the five algorithm-specific right-hand sides associated with $\dot{W}_k$, $\ddot{W}_k$, D_k^{unit} , D_k^{mix} , and D_k^{rem} . The direction D_k^{cen} is reconstructed from (19), and the amplification-specific direction used by CA-OFAS below can also be reconstructed algebraically from $\ddot{W}_k$. Therefore, changing the candidate parameters or testing an amplification candidate requires neither an additional KKT factorization nor an additional algorithm-specific linear-system solve. Evaluating a candidate then requires only linear combinations of precomputed directions, computation of auxiliary scalars, and evaluation of the acceptance conditions.

Remark 3.5 (Relation to the reduced KKT system in Clarabel.jl). The direct solver in Clarabel.jl does not directly factorize the full matrix representation of (6). Instead, it eliminates the cone-slack directions and the HSD scalar coupling, forms a reduced symmetric KKT system, and applies a direct LDL factorization to the reduced matrix [7, Section 3]. Once the NT scaling at the current iterate is fixed, the same reduced KKT matrix is used for $\dot{W}_k$, $\ddot{W}_k$, D_k^{unit} , D_k^{mix} , and D_k^{rem} ; only the right-hand side changes. Thus, the abstract one-factorization construction in (6) can be implemented in Clarabel.jl by reusing the factors of one reduced KKT matrix for multiple right-hand sides.

3.3 Corrected Candidates and Candidate Selection with Theoretical Guarantees

We add the correction $D = D_k^{\text{cor}}(\alpha, \rho)$ in (20) to the point $W_k(\alpha)$ on the arc. Since the complementarity row gives $\mathcal{B}(W_k, D; W_k) = -\Delta_k^{\text{comp}}(\alpha, \rho)$, a quadratic expansion of Φ with the above $U_k(\alpha)$ gives

$$\Phi(W_k(\alpha) + D; W_k) - T_k(\alpha, \rho) = \underbrace{\mathcal{B}(U_k(\alpha), D; W_k)}_{\text{arc-correction cross term}} + \underbrace{\Phi(D; W_k)}_{\text{quadratic correction term}}. \quad (21)$$

Thus, the current-iterate KKT correction exactly removes the complementarity mismatch linearized at the current iterate. The only remaining terms in (21) are the cross term between the arc increment and the correction and the quadratic term of the correction.

For a candidate angle α , define the target average complementarity by $\mu_{\text{tar}} := (1 - \sigma)\mu(W_k)$. For fixed (α, ρ) , if

$$\mu(W_k(\alpha) + D_k^{\text{cor}}(\alpha, \rho) + \zeta D_k^{\text{unit}}) = \mu_{\text{tar}} \quad (22)$$

has a unique solution, denote it by $\zeta_k(\alpha, \rho)$. The scalar ζ_k is not an independent candidate parameter; it is an auxiliary scalar for adjusting the average complementarity. Both D_k^{cor} and D_k^{unit} are HSD residual-invariant directions. Lemma 4.1(ii) directly shows the orthogonality of residual-zero directions. Hence, both the quadratic complementarity term of each residual-invariant direction and the cross complementarity term between two such directions vanish. Therefore, the average complementarity μ is exactly linear in ζ . With $V_0 := W_k(\alpha) + D_k^{\text{cor}}(\alpha, \rho)$, we can then compute

$$\zeta_k(\alpha, \rho) = \frac{\mu_{\text{tar}} - \mu(V_0)}{\mu(V_0 + D_k^{\text{unit}}) - \mu(V_0)} \quad (23)$$

directly when the denominator is nonzero. Section 4 verifies this linearity and uniqueness. We further define $D_k^{\text{full}}(\alpha, \rho) := D_k^{\text{cor}}(\alpha, \rho) + \zeta_k(\alpha, \rho)D_k^{\text{unit}}$ and use

$$W^+(\alpha, \rho, \beta) := W_k(\alpha) + \beta D_k^{\text{full}}(\alpha, \rho), \quad 0 < \beta \leq 1 \quad (24)$$

as the corrected candidate. The choice $\beta = 1$ gives the full correction, while $0 < \beta < 1$ gives a damped correction. Thus, β controls only the amount of correction.

Candidate selection uses only quantities that can be evaluated directly from an actual candidate point. First, set $\omega_0 := 1$ and $\omega_{k+1} := (1 - \sin \alpha_k)\omega_k$. For an HSD interior point W and $\omega > 0$, define $\theta(W, \omega) := \mu(W)/\omega$ and $\mathcal{J}(W, \omega) := \max\{\omega, \omega/\theta(W, \omega), \omega/\theta(W, \omega)^2\}$. The quantity $\mathcal{J}$ is computed directly from the candidate point and the known value of ω . Section 4 also uses it to control the scale of the projective normalization. For a candidate angle with $\sigma = \sin \alpha$, set $\omega^+ := (1 - \sigma)\omega_k$.

Fix $\sigma_{\text{max}} \in (0, 1)$, and let $\sigma^{\text{pre}} \in (0, \sigma_{\text{max}}]$ be the minimum value of $\sigma = \sin \alpha$ allowed for preliminary candidates. The threshold σ^{pre} is a fixed positive constant, independent of ν , used to separate preliminary candidates from the successive-halving fallback search. The analysis requires only that this threshold be positive, not any particular numerical value. We also set the centering-recovery parameter used in the successive-halving fallback search to $\rho_0 := 1/8$, and define $\mathcal{P}_\rho := \{1, 1/2, 1/4, \rho_0\}$ and $\mathcal{W}_{\text{HSD}} := \mathbb{S}_+^n \times \mathbb{R}^m \times \mathbb{S}_+^n \times \mathbb{R}_+ \times \mathbb{R}_+$. A candidate triple (α, ρ, β) is called *OFAS-admissible* if all of the following four conditions hold.

(C1) **Well-definedness:** $\mu_{\text{tar}} > 0$, and (22) has a unique solution $\zeta_k(\alpha, \rho)$.

(C2) **Interiority:** $W^+(\alpha, \rho, \beta) \in \text{int } \mathcal{W}_{\text{HSD}}$ and $\mu(W^+) > 0$.

(C3) **Neighborhood condition:** $\lambda_{\min}(\Phi(W^+; W_k)) \geq \gamma\mu(W^+)$.

(C4) **Local progress:**

$$\mathcal{Q}(W^+) \leq \left(1 - \frac{\sigma}{4}\right) \mathcal{Q}(W_k), \quad \mathcal{J}(W^+, \omega^+) \leq \left(1 - \frac{\sigma}{4}\right) \mathcal{J}(W_k, \omega_k).$$

Condition (C3), together with Lemma 4.2, guarantees that $W^+ \in \mathcal{N}_{\text{long}}(\gamma)$. Condition (C4) is a local progress condition introduced to obtain a uniform contraction factor and an iteration bound. The three-term decomposition, the current-iterate correction basis, and the candidate construction in (24) do not depend on (C4). Moreover, the one-factorization reconstruction itself does not require evaluation of $\mathcal{J}$. The numerical implementation in Section 5 uses the same candidate formulas and conditions (C1)–(C4), together with ordered preliminary candidates and a finite safeguarded version of the fallback search. Section 5 lists the numerical search safeguards and fixed parameter values used to make the candidate search more efficient. Since (C4) uses only quantities computed from the candidate point, checking it requires neither an additional KKT factorization nor an additional algorithm-specific linear-system solve.

Candidate selection has two stages. First, we evaluate at most $N^{\text{pre}} = O(1)$ ordered preliminary candidates satisfying $\sigma^{\text{pre}} \leq \sin \alpha \leq \sigma_{\max}$, $\rho \in \mathcal{P}_\rho$, and $0 < \beta \leq 1$, and accept the first admissible candidate. The preliminary rule is implementation-oriented; the worst-case analysis below uses only that at most $N^{\text{pre}} = O(1)$ preliminary candidates are tested before the fallback search. If no such candidate is accepted, we fix $\rho = \rho_0$ and use the dyadic fallback grid

$$\sigma_j := \sigma_{\max} 2^{-j}, \quad \alpha_j := \arcsin \sigma_j, \quad \beta_\ell := 2^{-\ell}, \quad j, \ell = 0, 1, \dots$$

The fallback candidates are examined in stages as specified in Algorithm 1. In Section 4, we show that this search reaches an admissible candidate with $\sigma = \Theta(1/\nu)$, $\beta = \Theta(1/\nu)$, and $\rho = \rho_0$.

3.4 The OFAS Algorithm

Algorithm 1 summarizes OFAS with the candidate-selection rule above. In Section 5, we implement and evaluate this algorithm on the HSD interior-point framework of Clarabel.jl. For the original SDP, termination is based on the accuracy test $\mathcal{Q}_{\text{SDP}}(W_k) \leq \delta$. Under Assumption 2.2, Subsection 4.5 shows that this condition is satisfied after finitely many iterations. In contrast, neither the HSD candidate selection nor the contraction analysis of the homogeneous progress measure $\mathcal{Q}$ uses this assumption.

Algorithm 1 One-Factorization Arc-Search Method (OFAS)

Require: SDP data $(\mathcal{A}, b, C)$, tolerance $\delta > 0$, $\gamma \in (0, 1)$, $0 < \sigma^{\text{pre}} \leq \sigma_{\max} < 1$, and a preliminary candidate rule

- 1: $W_0 \leftarrow (I, 0, I, 1, 1)$, $\omega_0 \leftarrow 1$, $k \leftarrow 0$
 - 2: **while** $\mathcal{Q}_{\text{SDP}}(W_k) > \delta$ **do**
 - 3: Compute the current-iterate NT scaling and factorize the corresponding KKT matrix once
 - 4: Reuse the same factorization for the five right-hand sides defining $\dot{W}_k, \ddot{W}_k, D_k^{\text{unit}}, D_k^{\text{mix}}, D_k^{\text{rem}}$; see (8)–(9), (18), and (17).
 - 5: Reconstruct D_k^{cen} from (19)
 - 6: Test the ordered preliminary candidates from Subsection 3.3
 - 7: If none is admissible, set $\rho = \rho_0$ and, for $m = 0, 1, \dots$, test the fallback pairs (j, ℓ) satisfying $\max\{j, \ell\} = m$, in increasing order of j and then ℓ , until an admissible candidate is found
 - 8: Set the accepted triple to $(\alpha_k, \rho_k, \beta_k)$ and $W_{k+1} \leftarrow W^+(\alpha_k, \rho_k, \beta_k)$
 - 9: $\omega_{k+1} \leftarrow (1 - \sin \alpha_k)\omega_k$, $k \leftarrow k + 1$
 - 10: **return** W_k
-

3.5 CA-OFAS: Curvature-Amplified Candidates

Fix an iteration k . For quantities specific to the OFAS and amplified candidates, we omit the iteration index k when there is no risk of confusion. We retain k for directions computed at the current iterate, complementarity targets, and quantities used to decide whether to evaluate the amplification branch, so that their relation to the current iterate remains explicit.

At each iteration, CA-OFAS first obtains an OFAS candidate $(\alpha_{\text{OFAS}}, \rho_{\text{OFAS}}, \beta_{\text{OFAS}})$ accepted by the OFAS candidate-selection rule, together with the corresponding point W_{OFAS}^+ . It then evaluates a different family of curvature-amplified arcs using the same current-iterate KKT factorization and correction basis. Here, α_{OFAS} is the accepted angle obtained from the OFAS candidate-selection rule; it is not the maximum admissible angle on the regular arc $W_k(\alpha)$. Therefore, the progress guarantee of OFAS is preserved even if no amplified candidate is accepted.

Let $\sigma_{\text{OFAS}} := \sin \alpha_{\text{OFAS}}$. For a fixed amplification factor $\eta > 1$ and increment $\Delta_\sigma > 0$, define

$$\sigma_{\text{CA}} := \min\{\sigma_{\text{max}}, \sigma_{\text{OFAS}} + \Delta_\sigma\}, \quad \alpha_{\text{CA}} := \arcsin \sigma_{\text{CA}}, \quad \psi_{\text{CA}} := 1 - \cos \alpha_{\text{CA}}.$$

The value σ_{CA} is a trial value for testing an amplified candidate with greater progress than the OFAS candidate. The amplified candidate does not simply extend α_{CA} farther along the regular arc; instead, it uses a different arc family in which the second-order direction is amplified by $\eta > 1$. We represent the decision to evaluate the amplification branch by $e_k^{\text{amp}} \in \{0, 1\}$. This quantity is determined only from values already computed at the current iterate. The activation condition is an implementation choice and does not enter the complexity analysis.

If $e_k^{\text{amp}} = 1$ and $\sigma_{\text{CA}} > \sigma_{\text{OFAS}}$, define the amplified arc by

$$W_{\text{arc,CA}} := W_k - \sigma_{\text{CA}} \dot{W}_k + \eta \psi_{\text{CA}} \ddot{W}_k. \quad (25)$$

This curve is not obtained by simply extending the regular arc $W_k(\alpha)$ beyond α_{OFAS} . Since $\ddot{W}_k$ is an HSD residual-invariant direction, this change requires no additional KKT linear-system solve. The residual reduction factor of the candidate remains $1 - \sigma_{\text{CA}}$. For the centering-recovery parameter, we use the OFAS value ρ_{OFAS} .

When $\eta \neq 1$, simply scaling the OFAS three-term decomposition by η is not sufficient because an additional complementarity mismatch proportional to $\Phi(-\dot{W}_k; W_k)$ appears. Define

$$T_{k,\text{CA}} := (1 - \sigma_{\text{CA}})\{(1 - \rho_{\text{OFAS}})\Phi(W_k; W_k) + \rho_{\text{OFAS}}\mu(W_k)\bar{I}\} \quad (26)$$

and $\Delta_{\text{CA}}^{\text{comp}} := \Phi(W_{\text{arc,CA}}; W_k) - T_{k,\text{CA}}$. Also define $c_{\text{CA}}^{\text{amp}} := (\eta - 1)\psi_{\text{CA}}\{(\eta + 1)\psi_{\text{CA}} - 2\}$. The following proposition gives the additional term caused by the amplification.

Proposition 3.6 (Exact decomposition of the amplified complementarity mismatch). *Using Δ_k^{cen} , Δ_k^{mix} , and Δ_k^{rem} defined in (16), we have*

$$\begin{aligned} \Delta_{\text{CA}}^{\text{comp}} &= \rho_{\text{OFAS}}(1 - \sigma_{\text{CA}})\Delta_k^{\text{cen}} + \eta\sigma_{\text{CA}}\psi_{\text{CA}}\Delta_k^{\text{mix}} + \eta^2\psi_{\text{CA}}^2\Delta_k^{\text{rem}} + \Delta_{\text{CA}}^{\text{amp}}, \\ \Delta_{\text{CA}}^{\text{amp}} &:= c_{\text{CA}}^{\text{amp}}\Phi(-\dot{W}_k; W_k). \end{aligned} \quad (27)$$

Therefore, the only additional term beyond the OFAS three-term decomposition is $\Delta_{\text{CA}}^{\text{amp}}$. In particular, when $\eta = 1$, we have $c_{\text{CA}}^{\text{amp}} = 0$, and the expression reduces to the OFAS three-term decomposition for the same $(\alpha_{\text{CA}}, \rho_{\text{OFAS}})$.

Proof. Expanding (25) quadratically and using (11) and (16) gives

$$\begin{aligned} \Phi(W_{\text{arc,CA}}; W_k) &= (1 - \sigma_{\text{CA}})\Phi(W_k; W_k) + \eta\sigma_{\text{CA}}\psi_{\text{CA}}\Delta_k^{\text{mix}} + \eta^2\psi_{\text{CA}}^2\Delta_k^{\text{rem}} \\ &\quad + \{\sigma_{\text{CA}}^2 - 2\eta\psi_{\text{CA}} + \eta^2\psi_{\text{CA}}^2\}\Phi(-\dot{W}_k; W_k). \end{aligned}$$

Since $\sigma_{\text{CA}}^2 = 2\psi_{\text{CA}} - \psi_{\text{CA}}^2$, the last coefficient is $(\eta - 1)\psi_{\text{CA}}\{(\eta + 1)\psi_{\text{CA}} - 2\} = c_{\text{CA}}^{\text{amp}}$. On the other hand, (26) can be written as $T_{k,\text{CA}} = (1 - \sigma_{\text{CA}})\Phi(W_k; W_k) - \rho_{\text{OFAS}}(1 - \sigma_{\text{CA}})\Delta_k^{\text{cen}}$. Taking the difference gives (27). $\square$

By Proposition 3.6, the same directions D_k^{cen} , D_k^{mix} , and D_k^{rem} can be reused for the amplified correction. Thus, the only new algebraic quantity in CA-OFAS is the response to the amplification-specific term $\Delta_{\text{CA}}^{\text{amp}}$. This response can also be reconstructed from the already computed direction $\ddot{W}_k$.

Lemma 3.7 (Algebraic reconstruction of the amplification-specific correction).

$$D_{k,\text{CA}}^{\text{amp}} := \mathcal{G}(\Delta_{\text{CA}}^{\text{amp}}; W_k) = \frac{c_{\text{CA}}^{\text{amp}}}{2} \ddot{W}_k. \quad (28)$$

Therefore, no additional linear-system solve is required for $D_{k,\text{CA}}^{\text{amp}}$.

Proof. Equation (9) gives $\mathcal{L}(W_k)[\ddot{W}_k] = (O_{\mathcal{F}}, -2\Phi(\dot{W}_k; W_k))$. Using $\Phi(-\dot{W}_k; W_k) = \Phi(\dot{W}_k; W_k)$ and the linearity of $\mathcal{L}(W_k)$, (9) gives

$$\mathcal{L}(W_k) \left[\frac{c_{\text{CA}}^{\text{amp}}}{2} \ddot{W}_k \right] = (O_{\mathcal{F}}, -c_{\text{CA}}^{\text{amp}} \Phi(-\dot{W}_k; W_k)) = (O_{\mathcal{F}}, -\Delta_{\text{CA}}^{\text{amp}}).$$

Therefore, the uniqueness of the correction-response map gives (28). $\square$

Combining the correction directions corresponding to the components of the amplified complementarity mismatch with the same coefficients gives the amplified corrector direction

$$D_{k,\text{CA}}^{\text{cor}} = \rho_{\text{OFAS}}(1 - \sigma_{\text{CA}})D_k^{\text{cen}} + \eta\sigma_{\text{CA}}\psi_{\text{CA}}D_k^{\text{mix}} + \eta^2\psi_{\text{CA}}^2D_k^{\text{rem}} + \frac{c_{\text{CA}}^{\text{amp}}}{2}\ddot{W}_k. \quad (29)$$

By linearity and Lemma 3.7, $\mathcal{L}(W_k)[D_{k,\text{CA}}^{\text{cor}}] = (O_{\mathcal{F}}, -\Delta_{\text{CA}}^{\text{comp}})$. Therefore, the amplification branch reuses the current-iterate KKT factorization and the five linear-system solves already performed for OFAS.

The average complementarity is adjusted in the same way as in OFAS. First, define $\mu_{\text{tar},\text{CA}} := (1 - \sigma_{\text{CA}})\mu(W_k)$. If the equation in the scalar ζ

$$\mu(W_{\text{arc},\text{CA}} + D_{k,\text{CA}}^{\text{cor}} + \zeta D_k^{\text{unit}}) = \mu_{\text{tar},\text{CA}} \quad (30)$$

has a unique solution, denote it by ζ_{CA} . The scalar ζ_{CA} can also be computed directly by the same two-point linear formula as in (23), with no additional KKT linear-system solve. If the solution is not unique, we do not construct the amplified candidate. Next, for a damping factor $\beta_{\text{CA}} \in (0, 1]$, define the corrected amplified candidate by

$$W_{\text{CA}}^+(\beta_{\text{CA}}) := W_{\text{arc},\text{CA}} + \beta_{\text{CA}} (D_{k,\text{CA}}^{\text{cor}} + \zeta_{\text{CA}} D_k^{\text{unit}}).$$

For the amplification branch, conditions (C1)–(C4) are evaluated by replacing the corresponding OFAS quantities with $\mu_{\text{tar},\text{CA}}$, ζ_{CA} , $W_{\text{CA}}^+(\beta_{\text{CA}})$, and $\omega_{\text{CA}}^+ := (1 - \sigma_{\text{CA}})\omega_k$. We evaluate, in a prescribed order, at most $N^{\text{amp}} = O(1)$ damping values fixed in advance, and accept the first β_{CA} for which the amplified candidate satisfies conditions (C1)–(C4). If no such β_{CA} exists, we use the already accepted OFAS point W_{OFAS}^+ .

In summary, CA-OFAS first obtains the guaranteed OFAS candidate of Algorithm 1 and then tests the amplification branch before updating W_{OFAS}^+ . An amplified candidate is accepted only if it also satisfies conditions (C1)–(C4); otherwise, the method uses the OFAS candidate. Therefore, Section 4 can establish the same invariant conditions and iteration bound as for OFAS. The numerical implementation in Section 5 combines the same amplified-candidate construction with implementation-oriented search rules and safeguards. Rather than repeating the complete procedure, Algorithm 2 lists only the changes from OFAS.

Algorithm 2 CA-OFAS: Changes from OFAS

Require: The inputs of Algorithm 1, an amplification factor $\eta > 1$, an increment $\Delta_\sigma > 0$, an activation condition, and at most $N^{\text{amp}} = O(1)$ prescribed damping values for amplification

- 1: Run the current OFAS iteration until an admissible OFAS candidate $(\alpha_{\text{OFAS}}, \rho_{\text{OFAS}}, \beta_{\text{OFAS}})$ and W_{OFAS}^+ are obtained, and postpone the update
 - 2: Set $\sigma_{\text{OFAS}} \leftarrow \sin \alpha_{\text{OFAS}}$ and $\sigma_{\text{CA}} \leftarrow \min\{\sigma_{\text{max}}, \sigma_{\text{OFAS}} + \Delta_\sigma\}$, and evaluate the amplification activation condition
 - 3: **if** amplification is activated and $\sigma_{\text{CA}} > \sigma_{\text{OFAS}}$ **then**
 - 4: Construct $W_{\text{arc,CA}}$ and $D_{k,\text{CA}}^{\text{cor}}$ from the directions already computed in the OFAS iteration using (25), (28), and (29)
 - 5: If (30) has a unique solution ζ_{CA} , evaluate the prescribed amplification damping values β_{CA} in order
 - 6: If any $W_{\text{CA}}^+(\beta_{\text{CA}})$ satisfies (C1)–(C4), accept the first such candidate as W_{k+1}
 - 7: If no amplified candidate is accepted, set $W_{k+1} \leftarrow W_{\text{OFAS}}^+$
 - 8: If an amplified candidate was accepted, set $\omega_{k+1} \leftarrow (1 - \sigma_{\text{CA}})\omega_k$; otherwise set $\omega_{k+1} \leftarrow (1 - \sigma_{\text{OFAS}})\omega_k$, and proceed to the next iteration
-

4 Long-Step Analysis of OFAS and CA-OFAS

In this section, we give a long-step analysis of OFAS and CA-OFAS using the candidate-selection rule defined in Algorithms 1 and 2 of Section 3. Throughout this section, $\varepsilon > 0$ denotes the target accuracy for the homogeneous progress measure, and $\delta > 0$ denotes the normalized KKT accuracy for the original SDP. Our goal is to preserve strict interiority, membership in the wide long-step neighborhood, and uniform progress without solving a new corrector system at the accepted point. The candidate-dependent correction is instead reconstructed in the NT coordinates of the current iterate. The homogeneous analysis has four parts. We first derive exact identities for the OFAS update and a certificate in the current-iterate NT coordinates. We then establish NT direction estimates in the wide long-step neighborhood. Next, we construct a uniform admissible window with $\sin \alpha = \Theta(1/\nu)$ and $\beta = \Theta(1/\nu)$. Finally, we prove finite termination of the candidate search and an iteration bound of $O(\nu \log(1/\varepsilon))$. We then use projective normalization to evaluate the accuracy for the original SDP.

We use the notation of Sections 2–3. Assumption 2.1 is used for unique solvability of the KKT operator at the current iterate. Assumption 2.2, however, is not used until Subsection 4.5; it is imposed only when recovering a solution of the original SDP by projective normalization. For the standard initial point W_0 , let $\mathcal{Q}_0 := \mathcal{Q}(W_0)$. The condition $\mathcal{Q}(W) \leq \varepsilon$ measures the accuracy of the homogeneous HSD quantities and is distinguished from the stopping condition $\mathcal{Q}_{\text{SDP}}(W) \leq \delta$ for the original SDP.

To construct the uniformly guaranteed window, we use $\rho_0 = 1/8$ defined in Subsection 3.3. If all preliminary candidates are rejected, the method fixes $\rho = \rho_0$ and moves to the successive-halving fallback search in the angle and damping. The particular outcomes for the preliminary values $\rho \in \mathcal{P}_\rho = \{1, 1/2, 1/4, \rho_0\}$ therefore do not affect the worst-case argument. Therefore, for the worst-case analysis, it is enough to show that this successive-halving fallback search contains an admissible candidate with $\sin \alpha = \Theta(1/\nu)$ and $\beta = \Theta(1/\nu)$.

Below, we analyze the number of outer iterations and use the results of Section 3 for the numbers of matrix factorizations and linear-system solves per iteration. As in Yang and Yamashita [15], the basic proof structure consists of bounds on products of directions, the existence of an $O(1/\nu)$ step, and a uniform cumulative decrease rate.

4.1 Exact Identities for the OFAS Update and a Current-Iterate Certificate

In the following local calculations, we focus on iteration k and write $\dot{W} := \dot{W}_k$ and $\ddot{W} := \ddot{W}_k$ for brevity. We also set $\sigma := \sin \alpha$ and $\psi := 1 - \cos \alpha$. The HSD residual map Ψ is linear. The first- and second-order directions satisfy $\Psi(\dot{W}) = \Psi(W)$ and $\Psi(\ddot{W}) = 0$, respectively. Define the total complementarity, combining the PSD block and the scalar HSD block, by $\mathcal{C}(W) := X \bullet S + \tau\kappa = \nu\mu(W)$. For HSD variations $U = (U_X, u_y, U_S, u_\tau, u_\kappa)$ and $V = (V_X, v_y, V_S, v_\tau, v_\kappa)$, define the polarization of $\mathcal{C}$ by

$$\Gamma(U, V) := U_X \bullet V_S + V_X \bullet U_S + u_\tau v_\kappa + v_\tau u_\kappa.$$

Then $\mathcal{C}(U + V) = \mathcal{C}(U) + \Gamma(U, V) + \mathcal{C}(V)$. Here, Γ is the cross term of the scaling-independent scalar complementarity $\mathcal{C}$. This is different from the Jordan-product cross term $\mathcal{B}(\cdot, \cdot; W_k)$ in the current-iterate NT coordinates; see Section 2.

Lemma 4.1 (Exact identities for the OFAS update). *The following statements hold.*

- (i) *The predictor arc $W_k(\alpha) = W - \sigma\dot{W} + \psi\ddot{W}$ satisfies $\Psi(W_k(\alpha)) = (1 - \sigma)\Psi(W)$. This identity is unchanged if the second-order direction is multiplied by any scalar.*
- (ii) *For any residual-invariant direction $D = (D_X, d_y, D_S, d_\tau, d_\kappa)$ satisfying $\Psi(D) = 0$, $D_X \bullet D_S + d_\tau d_\kappa = 0$. Moreover, for any two residual-invariant directions D_1, D_2 , we have $\Gamma(D_1, D_2) = 0$. Here D denotes an arbitrary residual-invariant direction.*
- (iii) *Suppose that $D_k^{\text{cor}}(\alpha, \rho)$ and D_k^{unit} are residual-invariant directions, and define*

$$\begin{aligned}\mu_{\text{tar}} &:= (1 - \sigma)\mu(W), \\ D_\zeta &:= \Gamma(W_k(\alpha), D_k^{\text{unit}}), \\ N_\zeta &:= \nu\mu_{\text{tar}} - \mathcal{C}(W_k(\alpha)) - \Gamma(W_k(\alpha), D_k^{\text{cor}}(\alpha, \rho)).\end{aligned}$$

If $D_\zeta \neq 0$, then $\zeta = \frac{N_\zeta}{D_\zeta}$ is the unique scalar for which the fully corrected point corresponding to $\beta = 1$ in (24), $W_k(\alpha) + D_k^{\text{cor}}(\alpha, \rho) + \zeta D_k^{\text{unit}}$, satisfies $\mu(W_k(\alpha) + D_k^{\text{cor}}(\alpha, \rho) + \zeta D_k^{\text{unit}}) = \mu_{\text{tar}}$.

(iv) *Let*

$$D_k^{\text{full}}(\alpha, \rho) := D_k^{\text{cor}}(\alpha, \rho) + \zeta D_k^{\text{unit}}, \quad W^+(\beta) := W_k(\alpha) + \beta D_k^{\text{full}}(\alpha, \rho), \quad 0 \leq \beta \leq 1.$$

Then

$$\Psi(W^+(\beta)) = (1 - \sigma)\Psi(W), \tag{31}$$

$$\mu(W^+(\beta)) = (1 - \beta)\mu(W_k(\alpha)) + \beta(1 - \sigma)\mu(W). \tag{32}$$

Proof. For (i), linearity of Ψ gives

$$\Psi(W_k(\alpha)) = \Psi(W) - \sigma\Psi(\dot{W}) + \psi\Psi(\ddot{W}) = (1 - \sigma)\Psi(W).$$

Multiplying $\ddot{W}$ by any scalar does not change this identity because $\Psi(\ddot{W}) = 0$.

For (ii), the orthogonality relation follows from the skew-adjoint structure of the HSD constraint system. The same structure appears in the HSD formulation for LP [10], but we derive it directly here from the SDP-HSD residual equations used in this paper. The condition $\Psi(D) = 0$ means $\mathcal{A}(D_X) = bd_\tau$, $\mathcal{A}^*(d_y) + D_S = Cd_\tau$, and $C \bullet D_X - b^T d_y + d_\kappa = 0$. Taking the inner product of the second equation with D_X and using the first equation gives

$$D_X \bullet D_S = d_\tau(C \bullet D_X - b^T d_y) = -d_\tau d_\kappa.$$

Hence, $\mathcal{C}(D) = 0$. Since the residual-invariance equations are linear, $D_1 + D_2$ is also residual-invariant. Therefore, $0 = \mathcal{C}(D_1 + D_2) = \mathcal{C}(D_1) + \Gamma(D_1, D_2) + \mathcal{C}(D_2)$, which gives $\Gamma(D_1, D_2) = 0$. This proof uses only the matrix inner product and adjoint relations; it is not specific to LP.

For (iii), both $D_k^{\text{cor}}(\alpha, \rho)$ and D_k^{unit} are residual-invariant. Part (ii) therefore gives

$$\mathcal{C}(D_k^{\text{cor}}(\alpha, \rho)) = \mathcal{C}(D_k^{\text{unit}}) = 0, \quad \Gamma(D_k^{\text{cor}}(\alpha, \rho), D_k^{\text{unit}}) = 0.$$

Thus, for any scalar ζ ,

$$\mathcal{C}(D_k^{\text{cor}}(\alpha, \rho) + \zeta D_k^{\text{unit}}) = 0.$$

Moreover, the same quadratic expansion gives

$$\begin{aligned}\mathcal{C}(W_k(\alpha) + D_k^{\text{cor}}(\alpha, \rho) + \zeta D_k^{\text{unit}}) \\ = \mathcal{C}(W_k(\alpha)) + \Gamma(W_k(\alpha), D_k^{\text{cor}}(\alpha, \rho)) + \zeta D_\zeta.\end{aligned}$$

Hence, the average complementarity after the full correction is exactly affine in ζ . Equating it to $\nu\mu_{\text{tar}}$ gives $\zeta = N_\zeta/D_\zeta$.

For (iv), write $D^{\text{full}} := D_k^{\text{full}}(\alpha, \rho)$. By part (ii), $\mathcal{C}(D^{\text{full}}) = 0$. Hence, for every $0 \leq \beta \leq 1$,

$$\mathcal{C}(W_k(\alpha) + \beta D^{\text{full}}) = \mathcal{C}(W_k(\alpha)) + \beta \Gamma(W_k(\alpha), D^{\text{full}}).$$

At $\beta = 1$, part (iii) gives

$$\mathcal{C}(W_k(\alpha)) + \Gamma(W_k(\alpha), D^{\text{full}}) = \nu\mu_{\text{tar}}.$$

Eliminating the cross term between these two identities yields

$$\mathcal{C}(W_k(\alpha) + \beta D^{\text{full}}) = (1 - \beta)\mathcal{C}(W_k(\alpha)) + \beta\nu\mu_{\text{tar}}.$$

Dividing by ν gives (32). Likewise, linearity of Ψ and $\Psi(D^{\text{full}}) = 0$ give $\Psi(W_k(\alpha) + \beta D^{\text{full}}) = \Psi(W_k(\alpha))$, which is (31) by part (i). $\square$

This lemma does not require strict interiority of the predictor arc, so the same algebra also applies to the amplified arc of CA-OFAS. However, positivity and membership in the wide long-step neighborhood for an amplified candidate are checked separately by the acceptance conditions. Also, when $\beta < 1$, in general $\mu(W^+) \neq (1 - \sigma)\mu(W)$. The iteration analysis below therefore does not rely on this equality.

The next lemma connects the current-iterate NT certificate in acceptance condition (C3) of Section 3 with the scaling-independent wide long-step neighborhood.

Lemma 4.2 (Wide Long-Step Neighborhood Certificate in the Current-Iterate NT Coordinates). *For a strictly interior candidate W_{cand} , if $\lambda_{\min}(\Phi(W_{\text{cand}}; W_k)) \geq \gamma\mu(W_{\text{cand}})$, then $W_{\text{cand}} \in \mathcal{N}_{\text{long}}(\gamma)$.*

Proof. Write $W_{\text{cand}} = (X_{\text{cand}}, y_{\text{cand}}, S_{\text{cand}}, \tau_{\text{cand}}, \kappa_{\text{cand}})$. Apply the current-iterate NT coordinates in (3) to W_{cand} and define the PSD blocks by $P_X = M_k X_{\text{cand}} M_k^T$ and $R_X = M_k^{-T} S_{\text{cand}} M_k^{-1}$. Then $P_X R_X = M_k X_{\text{cand}} S_{\text{cand}} M_k^{-1}$, so $P_X R_X$ is similar to $X_{\text{cand}}^{1/2} S_{\text{cand}} X_{\text{cand}}^{1/2}$ and therefore has the same positive real eigenvalues. Symmetrization of the noncommutative complementarity product is standard in the analysis of Monteiro–Zhang-type search directions for SDP, and the NT direction is also treated in this framework [31, 32]. Let λ_* be the smallest eigenvalue of $P_X R_X$, and let v be a corresponding real unit eigenvector. Since P_X and R_X are symmetric, $v^T (P_X \circ R_X) v = v^T P_X R_X v = \lambda_*$. The Rayleigh quotient therefore gives

$$\lambda_{\min}(P_X \circ R_X) \leq \lambda_* = \lambda_{\min}(X_{\text{cand}}^{1/2} S_{\text{cand}} X_{\text{cand}}^{1/2}).$$

For the scalar HSD block, $(\xi_k \tau_{\text{cand}})(\xi_k^{-1} \kappa_{\text{cand}}) = \tau_{\text{cand}} \kappa_{\text{cand}}$. Therefore, the lower bound in (C3) also holds for the true complementarity spectrum of both the PSD block and the scalar block of W_{cand} . Hence, $W_{\text{cand}} \in \mathcal{N}_{\text{long}}(\gamma)$. $\square$

This certificate is sufficient but not necessary. For a general candidate, P_X and R_X need not commute, and the smallest eigenvalue of $P_X \circ R_X$ can be smaller than the smallest value in the true complementarity spectrum. Thus, the test can be performed entirely in the current-iterate NT coordinates, although it can be conservative.

4.2 NT Direction Estimates in the Wide Long-Step Neighborhood

This subsection collects the direction estimates needed later for the arc remainder and the guaranteed candidate. We use the product Euclidean Jordan algebra $\mathcal{E} = \mathbb{S}^n \times \mathbb{R}$ defined in Section 2, whose rank is $\nu = n + 1$. Fix an iteration and write $W := W_k \in \mathcal{N}_{\text{long}}(\gamma)$. We use the following

local NT notation:

$$\begin{aligned}
Q &:= Q^{\text{NT}}(W), & H &:= Q \circ Q = \Phi(W; W), & \mu &:= \mu(W), \\
L_Q(U) &:= Q \circ U, \\
(\dot{P}, \dot{R}) &:= (P(\dot{W}; W), R(\dot{W}; W)), \\
(\ddot{P}, \ddot{R}) &:= (P(\ddot{W}; W), R(\ddot{W}; W)), \\
(P_k^{\text{term}}, R_k^{\text{term}}) &:= (P(D_k^{\text{term}}; W_k), R(D_k^{\text{term}}; W_k)), \\
&\text{term} \in \{\text{cen}, \text{mix}, \text{rem}, \text{unit}\}.
\end{aligned} \tag{33}$$

By (3), the spectrum of the PSD block of H is $\lambda(X, S)$, while the scalar HSD block is $\tau\kappa$. Since $W \in \mathcal{N}_{\text{long}}(\gamma)$, every spectral value of H is at least $\gamma\mu$, and their sum is $X \bullet S + \tau\kappa = \nu\mu$.

We use the standard theory of NT scaling on self-scaled cones [11, 12] and derive the estimates needed here in the present notation. Similar estimates for scaled directions are also used in arc-search methods for SDP [17, 18]. We also need their relation to the orthogonality of HSD residual-invariant directions. The NT congruence transformation preserves the primal-dual cross inner product. Thus, for a direction D with NT components (P^D, R^D) , $\langle P^D, R^D \rangle = D_X \bullet D_S + d_\tau d_\kappa$. Hence, by Lemma 4.1(ii), $\langle P^D, R^D \rangle = 0$ for every HSD residual-invariant direction.

Proposition 4.3 (NT-scaled direction estimates for OFAS). *Let $W \in \mathcal{N}_{\text{long}}(\gamma)$. Then the NT components of the first- and second-order directions satisfy*

$$\begin{aligned}
\|\dot{P}\|^2 + \|\dot{R}\|^2 &= \nu\mu, & \|\dot{P} \circ \dot{R}\| &\leq \frac{\nu\mu}{2}, \\
\|\ddot{P}\|^2 + \|\ddot{R}\|^2 &\leq \frac{\nu^2}{\gamma}\mu, & \|\dot{P} \circ \ddot{R} + \ddot{P} \circ \dot{R}\| &\leq \frac{\nu^{3/2}}{\sqrt{\gamma}}\mu, \\
&& \|\ddot{P} \circ \ddot{R}\| &\leq \frac{\nu^2}{2\gamma}\mu.
\end{aligned} \tag{34}$$

For the centering right-hand side,

$$\|L_Q^{-1}(H - \mu\bar{I})\|^2 \leq \nu(\gamma^{-1} - 1)\mu.$$

Let (P_Δ, R_Δ) be the NT components of the residual-invariant solution $D = \mathcal{G}(\Delta; W)$ whose only complementarity right-hand side is $-\Delta$. Then

$$L_Q(P_\Delta + R_\Delta) = -\Delta, \quad \langle P_\Delta, R_\Delta \rangle = 0,$$

and therefore

$$\sqrt{\|P_\Delta\|^2 + \|R_\Delta\|^2} = \|L_Q^{-1}(\Delta)\|.$$

Moreover, for any two pairs (P, R) and $(\hat{P}, \hat{R})$,

$$\|P \circ \hat{R} + \hat{P} \circ R\| \leq \sqrt{\|P\|^2 + \|R\|^2} \sqrt{\|\hat{P}\|^2 + \|\hat{R}\|^2}, \quad \|P \circ R\| \leq \frac{1}{2}(\|P\|^2 + \|R\|^2).$$

In particular, the four correction-basis directions defined in Section 3 satisfy

$$\begin{aligned}
\sqrt{\|P_k^{\text{cen}}\|^2 + \|R_k^{\text{cen}}\|^2} &\leq \sqrt{\nu(\gamma^{-1} - 1)\mu}, \\
\sqrt{\|P_k^{\text{mix}}\|^2 + \|R_k^{\text{mix}}\|^2} &\leq \frac{\nu^{3/2}}{\gamma}\sqrt{\mu}, \\
\sqrt{\|P_k^{\text{rem}}\|^2 + \|R_k^{\text{rem}}\|^2} &\leq \frac{1 + \gamma}{2\gamma^{3/2}}\nu^2\sqrt{\mu}, \\
\sqrt{\|P_k^{\text{unit}}\|^2 + \|R_k^{\text{unit}}\|^2} &\leq \sqrt{\frac{\nu}{\gamma\mu}}.
\end{aligned}$$

Proof. For NT scaling on self-scaled cones, we use the standard Jordan-product bound $\|U \circ V\| \leq \|U\| \|V\|$ and the inverse-operator estimate obtained from the spectral decomposition [11, 12]. Since every spectral value of $H = Q \circ Q$ is at least $\gamma\mu$, every spectral value of Q is at least $\sqrt{\gamma\mu}$. If the spectral values of Q on the PSD block are denoted by q_i , the eigenvalues of L_Q are $(q_i + q_j)/2$; on the scalar HSD block, L_Q is multiplication by the scalar component of Q . These are eigenvalues of the linear operator $L_Q : \mathbb{S}^n \rightarrow \mathbb{S}^n$, not the ordinary matrix eigenvalues of Q_X . Indeed, if $Q_X = U \text{Diag}(q_1, \dots, q_n) U^T$, then in the induced symmetric-matrix basis the mode associated with indices (i, j) is multiplied by $(q_i + q_j)/2$ (with value q_i when $i = j$). Therefore, for every $U \in \mathcal{E}$,

$$\|L_Q^{-1}(U)\| \leq \frac{1}{\sqrt{\gamma\mu}} \|U\|. \quad (35)$$

The standard Jordan-product bound and Cauchy-Schwarz also give, for any two pairs (P, R) and $(\hat{P}, \hat{R})$,

$$\|P \circ \hat{R} + \hat{P} \circ R\| \leq \sqrt{\|P\|^2 + \|R\|^2} \sqrt{\|\hat{P}\|^2 + \|\hat{R}\|^2}, \quad \|P \circ R\| \leq \frac{1}{2}(\|P\|^2 + \|R\|^2).$$

We first establish orthogonality of the first-order direction, which does not follow directly from residual invariance. For $U = (U_X, u_y, U_S, u_\tau, u_\kappa)$ and $V = (V_X, v_y, V_S, v_\tau, v_\kappa)$, direct substitution into the HSD residual equations gives

$$\begin{aligned} \Gamma(U, V) &= U_X \bullet R_d(V) - u_y^T R_p(V) + u_\tau R_g(V) \\ &\quad + V_X \bullet R_d(U) - v_y^T R_p(U) + v_\tau R_g(U). \end{aligned}$$

Since $\Psi(\dot{W}) = \Psi(W)$, we obtain $\Gamma(W, \dot{W}) = \mathcal{C}(W) + \mathcal{C}(\dot{W})$. On the other hand, taking the inner product of the first-order complementarity equation $L_Q(\dot{P} + \dot{R}) = H$ with $\bar{I}$ gives $\Gamma(W, \dot{W}) = \langle Q, \dot{P} + \dot{R} \rangle = \langle \bar{I}, H \rangle = \mathcal{C}(W)$. Thus, $\mathcal{C}(\dot{W}) = 0$, and preservation of the cross inner product under the NT congruence transformation gives $\langle \dot{P}, \dot{R} \rangle = 0$. Also, $L_Q(\dot{P} + \dot{R}) = L_Q(Q)$ implies $\dot{P} + \dot{R} = Q$. Therefore,

$$\|\dot{P}\|^2 + \|\dot{R}\|^2 = \|Q\|^2 = \nu\mu.$$

Using $\|U \circ V\| \leq \|U\| \|V\|$ and $2ab \leq a^2 + b^2$, we obtain

$$\|\dot{P} \circ \dot{R}\| \leq \|\dot{P}\| \|\dot{R}\| \leq \frac{1}{2}(\|\dot{P}\|^2 + \|\dot{R}\|^2) = \frac{\nu\mu}{2},$$

which is the product bound in (34).

The second-order direction satisfies $\Psi(\ddot{W}) = O_{\mathcal{F}}$. Hence, Lemma 4.1(ii) gives $\langle \ddot{P}, \ddot{R} \rangle = 0$. Moreover, $L_Q(\ddot{P} + \ddot{R}) = -2(\dot{P} \circ \dot{R})$. Using (35) and (34),

$$\|\ddot{P} + \ddot{R}\| \leq \frac{2}{\sqrt{\gamma\mu}} \|\dot{P} \circ \dot{R}\| \leq \frac{\nu}{\sqrt{\gamma}} \sqrt{\mu}.$$

By orthogonality, the square of the left-hand side equals $\|\ddot{P}\|^2 + \|\ddot{R}\|^2$, which gives the stated second-order bound. The standard Jordan-product bound and Cauchy-Schwarz then give

$$\begin{aligned} \|\dot{P} \circ \ddot{R} + \ddot{P} \circ \dot{R}\| &\leq \sqrt{\|\dot{P}\|^2 + \|\dot{R}\|^2} \sqrt{\|\ddot{P}\|^2 + \|\ddot{R}\|^2} \leq \frac{\nu^{3/2}}{\sqrt{\gamma}} \mu, \\ \|\ddot{P} \circ \ddot{R}\| &\leq \frac{1}{2}(\|\ddot{P}\|^2 + \|\ddot{R}\|^2) \leq \frac{\nu^2}{2\gamma} \mu. \end{aligned}$$

We next estimate the centering right-hand side. Let the spectral decomposition of H be $H = \mu \sum_{i=1}^{\nu} t_i c_i$, where $\{c_i\}_{i=1}^{\nu}$ is a Jordan frame of mutually orthogonal primitive idempotents. Then $\bar{I} = \sum_i c_i$ and $Q = \sqrt{\mu} \sum_i \sqrt{t_i} c_i$, with $t_i \geq \gamma$ and $\sum_i t_i = \nu$. Hence,

$$\|L_Q^{-1}(H - \mu\bar{I})\|^2 = \mu \sum_{i=1}^{\nu} \frac{(t_i - 1)^2}{t_i} = \mu \left(\sum_{i=1}^{\nu} t_i^{-1} - \nu \right) \leq \nu(\gamma^{-1} - 1)\mu.$$

Finally, for the residual-invariant solution $D = \mathcal{G}(\Delta; W)$, $L_Q(P_\Delta + R_\Delta) = -\Delta$ and $\langle P_\Delta, R_\Delta \rangle = 0$. Therefore,

$$\|P_\Delta\|^2 + \|R_\Delta\|^2 = \|P_\Delta + R_\Delta\|^2 = \|L_Q^{-1}(\Delta)\|^2.$$

By the definitions in Section 3 and the product bounds above,

$$\begin{aligned} \|\Delta_k^{\text{mix}}\| &\leq \frac{\nu^{3/2}}{\sqrt{\gamma}} \mu, \\ \|\Delta_k^{\text{rem}}\| &\leq \|\ddot{P} \circ \ddot{R}\| + \|\dot{P} \circ \dot{R}\| \\ &\leq \frac{\nu^2}{2\gamma} \mu + \frac{\nu}{2} \mu \leq \left(\frac{1}{2\gamma} + \frac{1}{2} \right) \nu^2 \mu = \frac{1+\gamma}{2\gamma} \nu^2 \mu. \end{aligned}$$

The last inequality uses $\nu \geq 1$. Applying (35) to this estimate gives the factor $(1+\gamma)/(2\gamma^{3/2})$ in Proposition 4.3. Combining these estimates with (35), and applying the centering estimate above to $\Delta_k^{\text{cen}} = H - \mu \bar{I}$, gives the stated bounds for D_k^{cen} , D_k^{mix} , and D_k^{rem} . The bound for D_k^{unit} follows from $D_k^{\text{unit}} = \mathcal{G}(-\bar{I}; W)$ and $\|\bar{I}\| = \sqrt{\nu}$. $\square$

These are the scaled estimates needed later for the arc remainder and the OFAS correction. The detailed constants are collected in Appendix A.2 and are used only where needed.

4.3 Uniformly Guaranteed Candidates and Candidate Selection

Throughout this subsection, let $\sigma := \sin \alpha$ and $\psi := 1 - \cos \alpha$. At iteration k , the current iterate W has primal and dual factors $P = R = Q$ in NT coordinates. Using the same current-iterate scaling, write the primal and dual factors at a point on the predictor arc as

$$P(\alpha) = Q - \sigma \dot{P} + \psi \ddot{P}, \quad R(\alpha) = Q - \sigma \dot{R} + \psi \ddot{R}.$$

Thus, $P(\alpha) - Q$ and $R(\alpha) - Q$ are the NT-coordinate displacements from W to the predictor-arc point. Combining $\psi \leq \sigma^2$ with Proposition 4.3 yields the next arc estimate.

Lemma 4.4 (Wide-neighborhood arc displacement and remainder). *Let $\Delta P(\alpha) := P(\alpha) - Q$ and $\Delta R(\alpha) := R(\alpha) - Q$. If $0 \leq \sigma \leq \min\{1, c/\nu\}$, then*

$$\|\Delta P(\alpha)\| + \|\Delta R(\alpha)\| \leq d_\gamma(c) \sqrt{\frac{\mu(W)}{\nu}}, \quad d_\gamma(c) := \sqrt{2}c + \sqrt{\frac{2}{\gamma}}c^2.$$

Moreover,

$$P(\alpha) \circ R(\alpha) = (1 - \sigma)H + \mathcal{E}_{\text{arc}}(\alpha), \tag{36}$$

where

$$\mathcal{E}_{\text{arc}}(\alpha) = -\psi^2(\dot{P} \circ \dot{R}) - \sigma\psi(\dot{P} \circ \ddot{R} + \ddot{P} \circ \dot{R}) + \psi^2(\ddot{P} \circ \ddot{R}),$$

and

$$\|\mathcal{E}_{\text{arc}}(\alpha)\| \leq r_\gamma(c) \frac{\mu(W)}{\nu^{3/2}}, \quad r_\gamma(c) := \frac{c^3}{\sqrt{\gamma}} + \frac{c^4}{2} + \frac{c^4}{2\gamma}. \tag{37}$$

For the average complementarity, we also have

$$|\mu(W_k(\alpha)) - (1 - \sigma)\mu(W)| \leq r_\gamma(c) \frac{\mu(W)}{\nu^2}. \tag{38}$$

Proof. Throughout this proof, μ denotes $\mu(W)$, consistently with the local notation in (33). From $\psi \leq \sigma^2$ and Proposition 4.3,

$$\|\Delta P(\alpha)\| + \|\Delta R(\alpha)\| \leq \sigma \sqrt{2\nu\mu} + \sigma^2 \sqrt{2/\gamma} \nu \sqrt{\mu} \leq d_\gamma(c) \sqrt{\mu/\nu}.$$

Expanding $P(\alpha) = Q - \sigma\dot{P} + \psi\ddot{P}$ and $R(\alpha) = Q - \sigma\dot{R} + \psi\ddot{R}$, and using $Q \circ (\dot{P} + \dot{R}) = H$, $Q \circ (\ddot{P} + \ddot{R}) = -2\dot{P} \circ \dot{R}$, and $\sigma^2 - 2\psi = -\psi^2$, gives (36). Therefore, using the three product estimates in Proposition 4.3, together with $\psi \leq \sigma^2$ and $\sigma \leq c/\nu$, we obtain

$$\|\mathcal{E}_{\text{arc}}(\alpha)\| \leq \psi^2 \frac{\nu\mu}{2} + \sigma\psi \frac{\nu^{3/2}}{\sqrt{\gamma}}\mu + \psi^2 \frac{\nu^2}{2\gamma}\mu \leq r_\gamma(c) \frac{\mu}{\nu^{3/2}},$$

which proves (37). Finally, from (36), $\mathcal{C}(W_k(\alpha)) = \langle \bar{I}, P(\alpha) \circ R(\alpha) \rangle_{\mathcal{E}}$, and $\mathcal{C}(W) = \langle \bar{I}, H \rangle_{\mathcal{E}} = \nu\mu$, we have

$$\mu(W_k(\alpha)) - (1 - \sigma)\mu(W) = \frac{1}{\nu} \langle \bar{I}, \mathcal{E}_{\text{arc}}(\alpha) \rangle_{\mathcal{E}}.$$

Since $\|\bar{I}\| = \sqrt{\nu}$, the Cauchy–Schwarz inequality and (37) give

$$|\mu(W_k(\alpha)) - (1 - \sigma)\mu(W)| \leq \frac{1}{\sqrt{\nu}} \|\mathcal{E}_{\text{arc}}(\alpha)\| \leq r_\gamma(c) \frac{\mu}{\nu^2},$$

which proves (38). $\square$

To preserve strict interiority and the wide long-step neighborhood, the correction damping parameter β must be sufficiently small. At the same time, this damping must not destroy HSD progress if we are to obtain a long-step iteration bound. We first show that the arc itself gives $O(1/\nu)$ progress and that a damped correction preserves this uniform decrease.

We first establish uniform progress for a small angle. Let $W \in \mathcal{N}_{\text{long}}(\gamma)$, and let $0 < c \leq 1/2$ satisfy $r_\gamma(c) \leq c/4$ and $c/(2\nu) < \sigma \leq c/\nu$. Choose ζ so that the fully corrected point has average complementarity $(1 - \sigma)\mu(W)$. Combining Lemma 4.4 with the damping interpolation in Lemma 4.1, we obtain, for every $0 \leq \beta \leq 1$,

$$\mu(W^+(\beta)) \leq \left(1 - \sigma + \frac{r_\gamma(c)}{\nu^2}\right) \mu(W) \leq \left(1 - \frac{c}{4\nu}\right) \mu(W). \quad (39)$$

The last inequality uses $\sigma > c/(2\nu)$, $r_\gamma(c) \leq c/4$, and $\nu \geq 1$. Combining the exact residual identity (31) with the average complementarity bound (39) gives $\mathcal{Q}(W^+(\beta)) \leq (1 - c/(4\nu))\mathcal{Q}(W)$. This estimate does not use a positive lower bound on β . Later, we impose $\beta = \Theta(1/\nu)$ for the neighborhood argument, not for progress. The upper bound on β preserves strict interiority, while the positive lower bound provides enough correction to recover the wide long-step neighborhood.

We next collect the constants required for a uniformly guaranteed candidate. For the fixed branch $\rho = \rho_0 = 1/8$, Appendix A.2 explicitly constructs positive constants $b_\gamma, c_\gamma > 0$ that depend only on γ , ρ_0 , and σ_{max} , and not on ν or the problem data. They are chosen so that, for $\sigma = \Theta(1/\nu)$ and $\beta = \Theta(1/\nu)$, the well-definedness of ζ , strict interiority, recovery of the wide long-step neighborhood, and the local progress condition below all hold. The appendix gives the explicit constants and verifies the required estimates. Here, we use only the resulting uniform admissible window.

Combining these results, every candidate triple with σ and β in suitable intervals of width proportional to $1/\nu$ satisfies (C1)–(C3). It also gives a uniform decrease in both $\mathcal{Q}$ and the projective scale measure $\mathcal{J}$. The next proposition gives an admissible window of positive width that the successive-halving fallback search is guaranteed to reach.

Proposition 4.5 (Uniform admissible window for OFAS candidate selection). *Let $W \in \mathcal{N}_{\text{long}}(\gamma)$ and $\omega > 0$, and use the notation of Section 3. Let b_γ and c_γ be the positive constants constructed in Appendix A.2, specifically in (54)–(55); in particular, $0 < c_\gamma \leq \min\{\sigma_{\text{max}}, 1/2\}$. For any candidate satisfying*

$$\frac{c_\gamma}{2\nu} < \sigma \leq \frac{c_\gamma}{\nu}, \quad \rho = \rho_0 = \frac{1}{8}, \quad \frac{b_\gamma}{4\nu} \leq \beta \leq \frac{b_\gamma}{\nu},$$

we have $\mu_{\text{tar}} > 0$ and $D_\zeta \geq \nu/2 > 0$, so ζ is well-defined. The corrected point W^+ is strictly interior and satisfies $\lambda_{\min}(\Phi(W^+; W)) \geq \gamma\mu(W^+)$, and hence $W^+ \in \mathcal{N}_{\text{long}}(\gamma)$. Moreover, if $\omega^+ := (1 - \sigma)\omega$, then the local progress condition (C4) of Section 3 holds. Therefore, every candidate triple (α, ρ, β) in the three ranges above satisfies conditions (C1)–(C4).

Proof. Since $c_\gamma \leq 1/2$, we have $\mu_{\text{tar}} = (1 - \sigma)\mu(W) > 0$. The bound $D_\zeta \geq \nu/2$, strict interiority, and the current-iterate NT certificate follow from Lemma A.1, Lemma A.2, Lemma A.3, and the choices of c_γ and b_γ used there. Thus, (C1)–(C3) hold.

For (C4), the small-angle progress estimate above gives $\mathcal{Q}(W^+) \leq (1 - c_\gamma/(4\nu))\mathcal{Q}(W)$. Since $\sigma \leq c_\gamma/\nu$, we have $1 - c_\gamma/(4\nu) \leq 1 - \sigma/4$, which gives (C4) for $\mathcal{Q}$.

Next, let $a_\gamma := r_\gamma(c_\gamma)/(1 - \sigma_{\max})$. By the definition of c_γ in Appendix A.2, specifically (55),

$$r_\gamma(c_\gamma) \leq \bar{r}_\gamma c_\gamma^3, \quad c_\gamma^2 \leq \frac{1 - \sigma_{\max}}{64\bar{r}_\gamma}.$$

Hence

$$a_\gamma = \frac{r_\gamma(c_\gamma)}{1 - \sigma_{\max}} \leq \frac{\bar{r}_\gamma c_\gamma^3}{1 - \sigma_{\max}} \leq \frac{c_\gamma}{64}.$$

Using the average complementarity estimate in Lemma A.3 and $\omega^+ = (1 - \sigma)\omega$, set $\theta := \mu(W)/\omega$ and $\theta^+ := \mu(W^+)/\omega^+$. Then

$$\left| \frac{\theta^+}{\theta} - 1 \right| \leq \frac{a_\gamma}{\nu^2} \leq \frac{c_\gamma}{64\nu^2}.$$

Hence $\theta^+/\theta \geq 1 - a_\gamma/\nu^2$. Using also $\sigma > c_\gamma/(2\nu)$, the three ratios that define $\mathcal{J}$ satisfy

$$\max \left\{ \frac{\omega^+}{\omega}, \frac{\omega^+/\theta^+}{\omega/\theta}, \frac{\omega^+/(\theta^+)^2}{\omega/\theta^2} \right\} \leq \frac{1 - c_\gamma/(2\nu)}{(1 - a_\gamma/\nu^2)^2} \leq 1 - \frac{c_\gamma}{4\nu} \leq 1 - \frac{\sigma}{4}.$$

The middle inequality follows from

$$\left(1 - \frac{c_\gamma}{4\nu}\right) \left(1 - \frac{a_\gamma}{\nu^2}\right)^2 - \left(1 - \frac{c_\gamma}{2\nu}\right) \geq \frac{c_\gamma}{4\nu} - \frac{2a_\gamma}{\nu^2} > 0.$$

Thus, (C4) also holds for $\mathcal{J}$. □

From now on, let α_k be the angle of the candidate accepted by OFAS, and write $\sigma_k := \sin \alpha_k$. Define the positive constant

$$c_\gamma^{\text{fb}} := \frac{1}{2} \min\{c_\gamma, \sigma_{\max} b_\gamma\}.$$

Theorem 4.6 (Well-definedness of the OFAS search and a lower bound on the accepted angle). *Suppose that $W_k \in \mathcal{N}_{\text{long}}(\gamma)$. Even if none of the preliminary candidates is admissible, the search of Section 3 finds a candidate satisfying (C1)–(C4) after finitely many candidate tests. Moreover, if a candidate is accepted by this search, then $\sigma_k > c_\gamma^{\text{fb}}/\nu$.*

Proof. Let J_ν be the smallest nonnegative integer satisfying $\sigma_{J_\nu} \leq c_\gamma/\nu$, and let L_ν be the smallest nonnegative integer satisfying $\beta_{L_\nu} \leq b_\gamma/\nu$. Since $\sigma_j = \sigma_{\max} 2^{-j} \rightarrow 0$ and $\beta_\ell = 2^{-\ell} \rightarrow 0$, both indices are finite. Moreover, $c_\gamma \leq \sigma_{\max}$ by Proposition 4.5. The defining inequality gives $\sigma_{J_\nu} \leq c_\gamma/\nu$. If $J_\nu > 0$, minimality and the halving rule give $\sigma_{J_\nu} > c_\gamma/(2\nu)$. If $J_\nu = 0$, the same lower bound follows from $\sigma_0 \geq c_\gamma$. Hence

$$\frac{c_\gamma}{2\nu} < \sigma_{J_\nu} \leq \frac{c_\gamma}{\nu}.$$

Similarly, $b_\gamma \leq 1 = \beta_0$ implies that L_ν is finite and

$$\frac{b_\gamma}{2\nu} < \beta_{L_\nu} \leq \frac{b_\gamma}{\nu}.$$

In particular, $\beta_{L_\nu} \geq b_\gamma/(4\nu)$, and therefore Proposition 4.5 shows that $(\sigma_{J_\nu}, \rho_0, \beta_{L_\nu})$ satisfies (C1)–(C4).

Set $M_\nu := \max\{J_\nu, L_\nu\}$. By construction, the staged fallback search has tested every pair (j, ℓ) with $0 \leq j, \ell \leq M_\nu$ by the end of stage M_ν . Hence it must accept a candidate no later than the test of (J_ν, L_ν) . If (j, ℓ) is the accepted fallback pair, then $j \leq M_\nu$, and therefore $\sigma_k = \sigma_j \geq \sigma_{M_\nu}$. If

$M_\nu = J_\nu$, then $\sigma_{M_\nu} = \sigma_{J_\nu} > c_\gamma/(2\nu)$. If $M_\nu = L_\nu$, then $\sigma_{M_\nu} = \sigma_{\max} 2^{-L_\nu} = \sigma_{\max} \beta_{L_\nu} > \sigma_{\max} b_\gamma/(2\nu)$. Thus, in either case,

$$\sigma_k \geq \sigma_{M_\nu} > \frac{1}{2\nu} \min\{c_\gamma, \sigma_{\max} b_\gamma\} = \frac{c_\gamma^{\text{fb}}}{\nu}.$$

Finally,

$$J_\nu \leq \left\lceil \log_2 \frac{\sigma_{\max} \nu}{c_\gamma} \right\rceil, \quad L_\nu \leq \left\lceil \log_2 \frac{\nu}{b_\gamma} \right\rceil.$$

Hence $M_\nu = O(\log \nu)$, and at most $(M_\nu + 1)^2 = O((\log \nu)^2)$ fallback pairs are tested before termination. $\square$

4.4 Iteration Complexity of OFAS and CA-OFAS

For the homogeneous progress analysis, we temporarily omit the stopping test based on $\mathcal{Q}_{\text{SDP}}$ in Algorithms 1–2. We consider the theoretical sequence obtained by continuing the same update rule. If the actual algorithm terminates earlier, the theoretical sequence need not be continued beyond that point. In Subsection 4.5, we use this sequence to derive an iteration bound for the actual algorithm by showing that the condition $\mathcal{Q}_{\text{SDP}} \leq \delta$ is satisfied after finitely many iterations.

Set $c_\gamma^{\text{acc}} := \min\{\sigma^{\text{pre}}, c_\gamma^{\text{fb}}\}$, $\hat{c}_\gamma := c_\gamma^{\text{acc}}/4$, and $q_\gamma := 1 - \hat{c}_\gamma/\nu$, where $\sigma^{\text{pre}} > 0$ is the lower bound for the preliminary candidates in Section 3. By construction, $0 < \hat{c}_\gamma < \nu$, so $q_\gamma \in (0, 1)$.

If a preliminary candidate is accepted, then $\sigma_k \geq \sigma^{\text{pre}}$. Otherwise, Theorem 4.6 gives $\sigma_k > c_\gamma^{\text{fb}}/\nu$. Thus, in both cases,

$$\sigma_k \geq \frac{c_\gamma^{\text{acc}}}{\nu}. \quad (40)$$

Together with condition (C4), this bound gives a uniform contraction estimate.

Theorem 4.7 (Well-definedness and iteration complexity of OFAS). *Consider the theoretical sequence obtained by continuing the update rule of Algorithm 1 from the standard initial point. Every iteration is well defined, and, for all $k \geq 0$,*

$$W_k \in \mathcal{N}_{\text{long}}(\gamma), \quad \mathcal{Q}(W_k) \leq q_\gamma^k \mathcal{Q}_0, \quad \mathcal{J}(W_k, \omega_k) \leq q_\gamma^k.$$

For $0 < \varepsilon < \mathcal{Q}_0$, define

$$K_\varepsilon := \left\lceil \frac{\log(\mathcal{Q}_0/\varepsilon)}{-\log q_\gamma} \right\rceil \leq \left\lceil \frac{\nu}{\hat{c}_\gamma} \log \frac{\mathcal{Q}_0}{\varepsilon} \right\rceil$$

Then $\mathcal{Q}(W_{K_\varepsilon}) \leq \varepsilon$. Since γ , σ^{pre} , $\sigma_{\max}$, and the initial scale $\mathcal{Q}_0$ are independent of ε , we obtain $K_\varepsilon = O(\nu \log(1/\varepsilon))$ as $\varepsilon \downarrow 0$.

Proof. At $k = 0$, all complementarity eigenvalues of the standard initial point $W_0 = (I, 0, I, 1, 1)$ are equal to 1. Thus $W_0 \in \mathcal{N}_{\text{long}}(\gamma)$ and $\mu(W_0) = 1$. Since $\omega_0 = 1$, we also have $\theta(W_0, \omega_0) = 1$ and $\mathcal{J}(W_0, \omega_0) = 1$.

Suppose that $W_k \in \mathcal{N}_{\text{long}}(\gamma)$ at iteration k . If an admissible preliminary candidate exists, the algorithm accepts one. Otherwise, Theorem 4.6 guarantees that the successive-halving fallback search produces an admissible candidate. Hence every iteration is well defined. The accepted candidate satisfies (C1)–(C4). Condition (C3) and Lemma 4.2 imply $W_{k+1} \in \mathcal{N}_{\text{long}}(\gamma)$, while condition (C4) and (40) give

$$\begin{aligned} \mathcal{Q}(W_{k+1}) &\leq \left(1 - \frac{\sigma_k}{4}\right) \mathcal{Q}(W_k) \leq q_\gamma \mathcal{Q}(W_k), \\ \mathcal{J}(W_{k+1}, \omega_{k+1}) &\leq \left(1 - \frac{\sigma_k}{4}\right) \mathcal{J}(W_k, \omega_k) \leq q_\gamma \mathcal{J}(W_k, \omega_k). \end{aligned}$$

Induction then gives $\mathcal{Q}(W_k) \leq q_\gamma^k \mathcal{Q}_0$ and $\mathcal{J}(W_k, \omega_k) \leq q_\gamma^k$ for all k .

Finally, $q_\gamma = 1 - \hat{c}_\gamma/\nu$ and $-\log(1 - x) \geq x$ give

$$-\log q_\gamma \geq \frac{\hat{c}_\gamma}{\nu}, \quad \frac{1}{-\log q_\gamma} \leq \frac{\nu}{\hat{c}_\gamma}.$$

This proves the bound on K_ε . $\square$

Corollary 4.8 (Iteration-bound inheritance for CA-OFAS). *Consider the theoretical sequence obtained by continuing the update rule of Algorithm 2 from the standard initial point. Every iteration is well defined, and, for all $k \geq 0$,*

$$W_k \in \mathcal{N}_{\text{long}}(\gamma), \quad \mathcal{Q}(W_k) \leq q_\gamma^k \mathcal{Q}_0, \quad \mathcal{J}(W_k, \omega_k) \leq q_\gamma^k.$$

Hence the iteration bound in Theorem 4.7 also holds for CA-OFAS, and $K_\varepsilon = O(\nu \log(1/\varepsilon))$ as $\varepsilon \downarrow 0$.

Proof. At each iteration, CA-OFAS first applies the OFAS candidate-selection rule and obtains an OFAS candidate W_{OFAS}^+ satisfying (C1)–(C4). If a preliminary candidate is accepted, then $\sigma_{\text{OFAS}} \geq \sigma^{\text{pre}}$; otherwise, Theorem 4.6 gives $\sigma_{\text{OFAS}} > c_\gamma^{\text{fb}}/\nu$. Thus the OFAS candidate satisfies the same uniform lower bound as in (40) and preserves the three inductive invariants used in Theorem 4.7.

An amplified candidate is accepted by Algorithm 2 only if it also satisfies (C1)–(C4). The amplification branch evaluates such a candidate only when $\sigma_{\text{CA}} > \sigma_{\text{OFAS}}$. Hence the uniform lower bound (40) for the OFAS candidate also gives $\sigma_{\text{CA}} \geq c_\gamma^{\text{acc}}/\nu$. Hence Lemma 4.2 and condition (C4) show that the amplified candidate preserves the same three invariants and the same contraction factor q_γ . If no amplified candidate is accepted, CA-OFAS uses the OFAS candidate W_{OFAS}^+ . Therefore the induction in Theorem 4.7 applies in either branch and gives the same iteration bound. No additional guarantee is required concerning the admissibility of an amplified candidate, its maximum admissible angle, or the existence of an amplified angle larger than the OFAS angle. $\square$

4.5 Projective Normalization and Recovery of a Solution to the Original SDP

For an HSD point with $\tau > 0$, the standard normalization for homogeneous embeddings of SDP and general conic programs recovers the primal-dual variables of the original problem as $(X/\tau, y/\tau, S/\tau)$ [7, 29]. Here we derive the control of τ needed to translate the homogeneous progress measure $\mathcal{Q}(W)$ into a bound on the normalized KKT accuracy.

Proposition 4.9 (Finite-accuracy HSD normalization). *Suppose that $W = (X, y, S, \tau, \kappa)$ is an interior HSD point with $\tau > 0$ and $\mathcal{Q}(W) \leq \varepsilon$. Let $\bar{X} = X/\tau$, $\bar{y} = y/\tau$, and $\bar{S} = S/\tau$. Then*

$$\|\mathcal{A}(\bar{X}) - b\| \leq \frac{\varepsilon}{\tau}, \quad \|\mathcal{A}^*(\bar{y}) + \bar{S} - C\| \leq \frac{\varepsilon}{\tau}, \quad (41)$$

$$\bar{X} \bullet \bar{S} \leq \frac{\nu\varepsilon}{\tau^2}, \quad |C \bullet \bar{X} - b^T \bar{y}| \leq \frac{\varepsilon}{\tau} + \frac{\nu\varepsilon}{\tau^2}. \quad (42)$$

Consequently, $\mathcal{Q}_{\text{SDP}}(W) \leq \varepsilon/\tau + \nu\varepsilon/\tau^2$.

Proof. Since $\mathcal{Q}(W) \leq \varepsilon$, each of the three HSD residual norms is at most ε . Dividing the primal and dual residual equations by τ gives (41). Also, $X \bullet S + \tau\kappa = \nu\mu(W) \leq \nu\varepsilon$, so $\bar{X} \bullet \bar{S} \leq \nu\varepsilon/\tau^2$. Finally, the gap residual satisfies

$$C \bullet \bar{X} - b^T \bar{y} = \frac{C \bullet X - b^T y}{\tau} = \frac{R_g(W) - \kappa}{\tau}.$$

Therefore,

$$|C \bullet \bar{X} - b^T \bar{y}| \leq \frac{|R_g(W)|}{\tau} + \frac{\kappa}{\tau} \leq \frac{\varepsilon}{\tau} + \frac{\tau\kappa}{\tau^2} \leq \frac{\varepsilon}{\tau} + \frac{\nu\varepsilon}{\tau^2},$$

which proves the second inequality in (42). $\square$

For ω_k defined in Section 3, the predictor arc multiplies all three HSD residuals by $1 - \sigma_k$, whereas the correction directions leave them unchanged. Hence

$$\Psi(W_k) = \omega_k \Psi(W_0), \quad \omega_k = \prod_{i=0}^{k-1} (1 - \sigma_i). \quad (43)$$

For the remainder of this subsection, write

$$\mathcal{Q}_k := \mathcal{Q}(W_k), \quad \theta_k := \frac{\mu(W_k)}{\omega_k}, \quad \mathcal{R}_0^{\max} := \max\{\|R_p(W_0)\|, \|R_d(W_0)\|, |R_g(W_0)|\}.$$

At the standard initial point, $\mu(W_0) = 1$. Thus the initial progress scale is $\mathcal{Q}_0 = \mathcal{Q}(W_0) = \max\{1, \mathcal{R}_0^{\max}\}$. Therefore,

$$\mathcal{Q}_k = \omega_k \max(\theta_k, \mathcal{R}_0^{\max}).$$

By the definition of $\mathcal{J}$ and $\mathcal{R}_0^{\max} \leq \mathcal{Q}_0$,

$$\max\left(\frac{\mathcal{Q}_k}{\theta_k}, \frac{\mathcal{Q}_k}{\theta_k^2}\right) \leq \mathcal{Q}_0 \mathcal{J}(W_k, \omega_k) \leq \mathcal{Q}_0 q_\gamma^k.$$

Fix $Z^* = (X^*, y^*, S^*) \in \mathcal{Z}^*$ and define

$$\delta_{\text{rec}}(Z^*) := \inf \left\{ \|\mathcal{A}(D_X)\| + \|\mathcal{A}^*(d_y) + D_S\| + S^* \bullet D_X + X^* \bullet D_S : \begin{array}{l} D_X, D_S \succeq O, \quad d_y \in \mathbb{R}^m, \quad \text{tr } D_X + \text{tr } D_S = 1 \end{array} \right\}. \quad (44)$$

Appendix A.3 shows that $\delta_{\text{rec}}(Z^*) > 0$. Define also

$$\mathfrak{D}_* := 1 + \frac{\|b\| + \|C\|}{\delta_{\text{rec}}(Z^*)}, \quad \mathfrak{H}_* := \frac{3 + \|y^*\| + \|X^*\|}{\delta_{\text{rec}}(Z^*)} + 1 + \|y^*\| + \|X^*\|.$$

Both are positive constants depending only on the fixed problem and Z^* . The next lemma gives two scale estimates used to establish boundedness and accuracy after normalization.

Lemma 4.10 (Projective scale estimates). *Under Assumptions 2.1–2.2, for every k ,*

$$\text{tr } X_k + \text{tr } S_k \leq \frac{(\|b\| + \|C\|)\tau_k + (3 + \|y^*\| + \|X^*\|)\mathcal{Q}_k}{\delta_{\text{rec}}(Z^*)}, \quad (45)$$

$$\tau_k \geq \frac{\nu\theta_k - \mathfrak{H}_*\mathcal{Q}_k}{\mathfrak{D}_*}, \quad \frac{\mathcal{Q}_k}{\theta_k} \leq \frac{\nu}{2\mathfrak{H}_*} \implies \tau_k \geq \frac{\nu\theta_k}{2\mathfrak{D}_*}. \quad (46)$$

Proof. See Appendix A.3. □

For a primal-dual triple $Z = (X, y, S)$, define the product-space norm $\|Z\|_{\text{pd}} := (\|X\|_F^2 + \|y\|^2 + \|S\|_F^2)^{1/2}$. The distance from Z to the optimal solution set $\mathcal{Z}^*$ is

$$\text{dist}_{\text{pd}}(Z, \mathcal{Z}^*) := \inf_{Z^* \in \mathcal{Z}^*} \|Z - Z^*\|_{\text{pd}}.$$

Theorem 4.11 (Recovery of an original SDP solution and iteration bound after normalization). *Under Assumptions 2.1–2.2, consider the theoretical sequence of OFAS or CA-OFAS defined in the preceding subsection. Then*

$$\max\left(\frac{\mathcal{Q}_k}{\theta_k}, \frac{\mathcal{Q}_k}{\theta_k^2}\right) \leq \mathcal{Q}_0 q_\gamma^k. \quad (47)$$

Let $\bar{X}_k := X_k/\tau_k$, $\bar{y}_k := y_k/\tau_k$, $\bar{S}_k := S_k/\tau_k$, and $Z_k := (\bar{X}_k, \bar{y}_k, \bar{S}_k)$. Then $\{Z_k\}$ is bounded and

$$\text{dist}_{\text{pd}}(Z_k, \mathcal{Z}^*) \rightarrow 0. \quad (48)$$

Moreover, for $\delta > 0$, define

$$\Delta_\delta := \min \left\{ \mathcal{Q}_0, \frac{\nu}{2\mathfrak{H}_*}, \frac{\delta\nu}{4\mathfrak{D}_*}, \frac{\delta\nu}{8\mathfrak{D}_*^2} \right\}.$$

The four terms in Δ_δ have separate roles. The first keeps the threshold no larger than the initial envelope $\mathcal{Q}_0$. For the remaining three, if $\mathcal{Q}_0 q_\gamma^k \leq \Delta_\delta$, then (47) gives $\mathcal{Q}_k/\theta_k \leq \Delta_\delta$ and $\mathcal{Q}_k/\theta_k^2 \leq$

Δ_δ . The bound $\Delta_\delta \leq \nu/(2\mathfrak{H}_*)$ permits the second estimate in (46), namely $\tau_k \geq \nu\theta_k/(2\mathfrak{D}_*)$. Consequently,

$$\frac{\mathcal{Q}_k}{\tau_k} \leq \frac{2\mathfrak{D}_*}{\nu} \Delta_\delta, \quad \frac{\nu\mathcal{Q}_k}{\tau_k^2} \leq \frac{4\mathfrak{D}_*^2}{\nu} \Delta_\delta.$$

The bounds $\Delta_\delta \leq \delta\nu/(4\mathfrak{D}_*)$ and $\Delta_\delta \leq \delta\nu/(8\mathfrak{D}_*^2)$ make these two quantities at most $\delta/2$, respectively; Proposition 4.9 then gives $\mathcal{Q}_{\text{SDP}}(W_k) \leq \delta$. Then, with

$$K_\delta := \left\lceil \frac{\log(\mathcal{Q}_0/\Delta_\delta)}{-\log q_\gamma} \right\rceil \leq \left\lceil \frac{\nu}{\hat{c}_\gamma} \log \frac{\mathcal{Q}_0}{\Delta_\delta} \right\rceil, \quad (49)$$

the theoretical sequence satisfies $\mathcal{Q}_{\text{SDP}}(W_k) \leq \delta$ for every $k \geq K_\delta$. Therefore, unless they terminate earlier, Algorithms 1 and 2 satisfy their stopping criterion no later than iteration K_δ . Since $\mathcal{Q}_0$, $\mathfrak{D}_*$, and $\mathfrak{H}_*$ are independent of δ , as $\delta \downarrow 0$,

$$K_\delta = O\left(\nu \log \frac{1}{\delta}\right). \quad (50)$$

In addition, $C \bullet \bar{X}_k \rightarrow p^*$ and $b^T \bar{y}_k \rightarrow d^* = p^*$. If $\mathcal{Z}^*$ is a singleton, the normalized iterates converge to that point.

Proof. Equation (47) implies $\mathcal{Q}_k/\theta_k \rightarrow 0$ and $\mathcal{Q}_k/\theta_k^2 \rightarrow 0$. Thus (46) applies for all sufficiently large k , and

$$\begin{aligned} \frac{\mathcal{Q}_k}{\tau_k} &\leq \frac{2\mathfrak{D}_*}{\nu} \frac{\mathcal{Q}_k}{\theta_k} \rightarrow 0, \\ \frac{\nu\mathcal{Q}_k}{(\tau_k)^2} &\leq \frac{4\mathfrak{D}_*^2}{\nu} \frac{\mathcal{Q}_k}{\theta_k^2} \rightarrow 0. \end{aligned}$$

Proposition 4.9 then shows that the primal residual, dual residual, complementarity, and duality gap of the normalized point all converge to zero.

Next, dividing (45) by τ_k gives

$$\text{tr } \bar{X}_k + \text{tr } \bar{S}_k \leq \frac{\|b\| + \|C\|}{\delta_{\text{rec}}(\mathcal{Z}^*)} + \frac{3 + \|y^*\| + \|X^*\|}{\delta_{\text{rec}}(\mathcal{Z}^*)} \frac{\mathcal{Q}_k}{\tau_k}.$$

Hence $\{\bar{X}_k\}$ and $\{\bar{S}_k\}$ are bounded. Moreover, $\mathcal{A}^*(\bar{y}_k) = C - \bar{S}_k + R_d(W_k)/\tau_k$ has a bounded right-hand side. Assumption 2.1 makes $\mathcal{A}^*$ injective, so $\{\bar{y}_k\}$ is also bounded. Thus $\{Z_k\}$ is bounded.

By closedness of the PSD cone and convergence of the KKT residuals, every accumulation point is primal-dual feasible and has zero duality gap. Hence every accumulation point belongs to $\mathcal{Z}^*$. A standard subsequence argument for bounded sequences then gives (48).

Now suppose that $\mathcal{Q}_0 q_\gamma^k \leq \Delta_\delta$. Equation (47) and $\Delta_\delta \leq \nu/(2\mathfrak{H}_*)$ allow us to apply (46), giving

$$\frac{\mathcal{Q}_k}{\tau_k} \leq \frac{2\mathfrak{D}_*}{\nu} \Delta_\delta \leq \frac{\delta}{2}, \quad \frac{\nu\mathcal{Q}_k}{(\tau_k)^2} \leq \frac{4\mathfrak{D}_*^2}{\nu} \Delta_\delta \leq \frac{\delta}{2}.$$

Proposition 4.9 then gives $\mathcal{Q}_{\text{SDP}}(W_k) \leq \delta$. The definition of K_δ in (49) guarantees $\mathcal{Q}_0 q_\gamma^k \leq \Delta_\delta$ for every $k \geq K_\delta$. The upper bound in the same equation follows from $q_\gamma = 1 - \hat{c}_\gamma/\nu$ and $-\log(1-x) \geq x$. For a fixed problem, $\Delta_\delta = \Theta(\delta)$ as $\delta \downarrow 0$, which proves (50).

Finally, by (48) and compactness of $\mathcal{Z}^*$, choose a nearest point $Z_k^* \in \mathcal{Z}^*$ to Z_k . Then $\|Z_k - Z_k^*\|_{\text{pd}} \rightarrow 0$, which gives convergence of the objective values. If $\mathcal{Z}^*$ is a singleton, the normalized iterates converge to its unique point. $\square$

Remark 4.12 (Multiple PSD blocks). For simplicity, this paper uses a single PSD block, but the same analysis extends to multiple PSD blocks through the standard product-cone treatment. NT scaling and the Jordan-algebra operations are applied blockwise [11, 12]. If the J PSD blocks have sizes $n_1, \dots, n_J$, the rank of the product cone is $\nu = 1 + \sum_{j=1}^J n_j$. Hence the analysis in this section and the iteration bound $O(\nu \log(1/\varepsilon))$ extend directly to the multiple-block case.

5 Numerical Experiments

The numerical experiments address two questions. First, we examine the computational effects of the post-arc corrector, the one-factorization structure, and curvature amplification. Second, we compare the resulting CA-OFAS method with the standard Mehrotra-type predictor-corrector method in Clarabel.jl [7, 27] and a Liu-type method, namely, our HSD/NT implementation based on the wide-neighborhood second-order corrector method of Liu-Liu (2012) [28]. We use 54 test problems in total: 18 SDPLIB problems and 36 NNV-SDP problems. We report the number of successful problems including Low Accuracy, the number reaching Full Accuracy, computation times, outer iteration counts, and performance profiles.

5.1 Experimental Setting and Compared Methods

Our implementation is built on the HSD interior-point framework of Clarabel.jl. We compare six methods: OFAS, CA-OFAS, Two-factorization, Corrector-free, Mehrotra, and the Liu-type method. Two-factorization follows essentially the same arc-search rule as OFAS, but after selecting the OFAS arc point, it updates the NT scaling and refactorizes the KKT coefficient matrix at that point before computing the post-arc corrector. Corrector-free is a one-factorization reference method without an independent post-arc corrector, and its centering parameter is fixed at 0.10. For the overall comparison, we use Mehrotra, the standard method in Clarabel.jl, and an HSD/NT adaptation of the wide-neighborhood second-order corrector method of Liu-Liu (2012). We refer to the latter as the Liu-type method throughout this section.

The experiments were run with Julia 1.12.6 on a Linux node equipped with two 2.90 GHz Intel Xeon Gold 6226R processors and 192 GB RAM. Julia itself ran with one thread, whereas BLAS and OpenMP each used eight threads. We used CHOLMOD for direct linear solves and disabled chordal decomposition. The six methods shared the same high-accuracy gap and feasibility tolerances of 10^{-8} and a 3600-second limit per instance.

The computational experiments are designed to isolate the effect of the proposed step-selection strategy within a common implementation framework. All Clarabel-based methods therefore use the same linear-algebra backend and solver settings, and the reported computation times are intended for relative comparison among these methods.

We use MathOptInterface (MOI) termination statuses. Following the terminology used in the numerical evaluation of Clarabel [7], we classify `OPTIMAL` as Full Accuracy and `ALMOST_OPTIMAL` as Low Accuracy. Low Accuracy or better is regarded as success, while pairwise comparisons of computation time and outer iterations use only problems for which both methods reached Full Accuracy. Detailed termination criteria are given in Appendix B.

To reduce timing variability, all six methods were run three times on each eligible problem, and we report the median computation time and outer iteration count. Repetition was used only when all six methods succeeded in the first run, all computation times exceeded one second, and the maximum computation time was below 120 seconds; otherwise, only one run was used. We also prespecified that a problem would be excluded from the computation-time comparison if any successful method finished in one second or less. None of the final 54 problems met this exclusion condition.

The computation times and outer iteration counts in Tables 3 and 4 are geometric means over the problems successfully solved by each method, including Low Accuracy results. These values summarize the overall behavior on each problem set. For pairwise speed comparisons without mixing different accuracy levels, we instead use geometric means of paired ratios over problems for which both methods reached Full Accuracy. In the following, t_A/t_B and k_A/k_B denote the geometric means of the computation-time ratio and outer-iteration ratio, respectively, over these matched Full Accuracy problems. In mathematical subscripts, CA denotes CA-OFAS. Problem-by-problem computation times, outer iteration counts, and termination states are given in Appendix B.

For OFAS and CA-OFAS, we set the wide long-step neighborhood parameter to $\gamma = 0.01$. For CA-OFAS, we additionally used $\eta = 1.75$ and $\Delta_\sigma = 0.10$ for all benchmark problems. The

CA-OFAS amplification parameters were selected using a 12-problem NNV-SDP core set and then kept unchanged for the other reported problems.

The OFAS and CA-OFAS implementations use the same candidate formulas and acceptance conditions (C1)–(C4) as in Section 3. The theoretical algorithms use the successive-halving fallback search without an a priori limit in order to obtain the worst-case guarantee, whereas the numerical code uses a finite safeguarded version. The implementation details of this search are given in Appendix B.

5.2 Test Problem Sets

We use two types of semidefinite programming test problems. The first consists of 18 standard benchmark problems from SDPLIB [33]. These 18 problems were fixed before the reported benchmark runs by selecting problems whose preliminary computation times were between 1 and 3600 seconds.

The second consists of 36 NNV-SDP problems arising from neural network verification. The SDP construction and the ellipsoid and rectangle input sets are taken from Azuma–Kim–Yamashita [34]. Let $r = n_0 = n_1$ denote the network width used to generate each problem. The order of the main PSD block is $n_{\text{PSD}} = 1 + 2r$. In SDPA standard form, the number of equality constraints is $m = 3r + 2$ for the ellipsoid family and $m = 4r + 1$ for the rectangle family. We use both dense and identity settings for the weight matrices. The ranges of (n_{PSD}, m) are $(81, 122)$ – $(113, 170)$ for ellipsoid/dense, $(65, 129)$ – $(97, 193)$ for rectangle/dense, and $(81, 161)$ – $(129, 257)$ for rectangle/identity. Problem-by-problem specifications and results are reported in Appendix B.

5.3 Numerical Comparison of Algorithmic Components

Table 1 compares computation time and outer iterations on the instances where both methods attained Full Accuracy in each component experiment. On these identical instance sets, we additionally record KKT factorization counts, linear-solve counts, and the time used by those two linear-algebra operations.

We first compare OFAS with Corrector-free to examine the role of the post-arc corrector. The corrector requires an additional linear-system solve in each outer iteration. On both problem sets, OFAS consequently used more linear-system solves than Corrector-free, but required fewer outer iterations and KKT factorizations.

For the matched Full Accuracy problems, these reductions were accompanied by lower total computation times. The geometric-mean computation time of OFAS was about 37% lower on SDPLIB and about 42% lower on NNV-SDP; OFAS was faster on 10 of the 11 matched SDPLIB problems and on all 25 matched NNV-SDP problems. Thus, on the tested problems, the additional solve per outer iteration did not increase the overall computational cost, while the number of outer iterations was reduced.

We next compare OFAS with Two-factorization to evaluate the one-factorization structure. Although OFAS required slightly more outer iterations on SDPLIB and more outer iterations on NNV-SDP, it was faster on every matched problem in both problem sets. The geometric-mean computation time was reduced by about 44% on SDPLIB and about 38% on NNV-SDP.

This difference is also reflected in the linear-algebra operations. Relative to Two-factorization, OFAS reduced the number of KKT factorizations and the factorization time by about 48% on SDPLIB and about 40% on NNV-SDP. The number of linear-system solves was about 13% smaller on SDPLIB and nearly unchanged on NNV-SDP. These results support the computational benefit of reusing a single KKT factorization for the post-arc corrector.

Finally, we compare CA-OFAS with OFAS to assess the additional effect of curvature amplification. On SDPLIB, the change was modest: the geometric-mean ratios for outer iterations and computation time were 0.964 and 0.963, respectively. On NNV-SDP, the corresponding ratios were 0.808 and 0.811, showing reductions of about 19% in both measures. CA-OFAS was faster on 7 of the 11 matched SDPLIB problems and on 28 of the 31 matched NNV-SDP problems. The similar

Table 1: Paired comparisons on problems solved to Full Accuracy by both methods

Problem set	Comparison A/B	Matched	t_A/t_B	k_A/k_B
SDPLIB 18	OFAS/Corrector-free	11	0.629	0.597
	OFAS/Two-factorization	11	0.557	1.048
	CA-OFAS/OFAS	11	0.963	0.964
NNV-SDP 36	OFAS/Corrector-free	25	0.575	0.548
	OFAS/Two-factorization	26	0.623	1.202
	CA-OFAS/OFAS	31	0.811	0.808

Matched gives the number of problems for which both methods reached Full Accuracy. t_A/t_B and k_A/k_B are the geometric means of the computation-time ratio and iteration-count ratio, respectively, over these problems. A ratio below one favors the method in the numerator.

iteration and time ratios on NNV-SDP suggest that most of the observed reduction in computation time came from fewer outer iterations, while the average cost per outer iteration remained similar.

5.4 Comparison with Existing Methods

We next compare CA-OFAS with Mehrotra and the Liu-type method to assess its overall performance relative to these methods. Table 2 reports the paired computation-time and outer-iteration ratios on problems for which both methods reached Full Accuracy.

Table 2: Paired comparisons of CA-OFAS with Mehrotra and the Liu-type method on problems for which both methods reached Full Accuracy

Problem set	Comparison A/B	Matched	t_A/t_B	k_A/k_B
SDPLIB 18	CA-OFAS/Mehrotra	14	0.830	0.804
	CA-OFAS/Liu-type	14	0.797	0.766
NNV-SDP 36	CA-OFAS/Mehrotra	31	0.869	0.848
	CA-OFAS/Liu-type	33	0.882	0.864

Matched gives the number of problems for which both methods reached Full Accuracy. t_A/t_B and k_A/k_B are the geometric means of the computation-time ratio and iteration-count ratio, respectively, over these problems. A ratio below one favors CA-OFAS.

5.4.1 SDPLIB

Table 3 summarizes the results for CA-OFAS, Mehrotra, and the Liu-type method. All three methods succeeded on all 18 SDPLIB problems, so their geometric means of computation time and outer iteration count are based on the same problem set. CA-OFAS reached Full Accuracy on 16 problems, while Mehrotra and the Liu-type method each reached Full Accuracy on 15. The geometric-mean computation time was 151.6 seconds for CA-OFAS, compared with 178.7 seconds for Mehrotra and 249.9 seconds for the Liu-type method. On the matched Full Accuracy problems, CA-OFAS reduced the geometric-mean outer iteration count by about 20% relative to Mehrotra and about 23% relative to the Liu-type method, with corresponding computation-time reductions of about 17% and 20%, respectively. It was also faster than the Liu-type method on all 14 matched problems. Thus, for these SDPLIB problems, the lower computation times of CA-OFAS are accompanied by fewer outer iterations.

5.4.2 NNV-SDP

Table 4 summarizes the results for the 36 NNV-SDP problems. The Liu-type method succeeded on 35 of the 36 problems, compared with 34 for both CA-OFAS and Mehrotra. The geometric-mean computation time over each method’s successful problems was 605.7 seconds for CA-OFAS, 734.7 seconds for Mehrotra, and 709.9 seconds for the Liu-type method; because the success sets are not

Table 3: Aggregated results for 18 SDPLIB problems

Method	Success	Full Accuracy	Time (s)	Iterations
CA-OFAS	18/18	16/18	151.6	16.21
Mehrotra	18/18	15/18	178.7	20.06
Liu-type	18/18	15/18	249.9	28.44

Success includes Low Accuracy, and Full Accuracy gives the number of problems with status **OPTIMAL**. Time and Iterations are geometric means over the problems successfully solved by each method, including Low Accuracy results. For SDPLIB, the three methods have the same success set, so these geometric means are based on the same 18 problems.

Table 4: Aggregated results for 36 NNV-SDP problems

Method	Success	Full Accuracy	Time (s)	Iterations
CA-OFAS	34/36	33/36	605.7	27.37
Mehrotra	34/36	33/36	734.7	32.50
Liu-type	35/36	35/36	709.9	32.15

Success includes Low Accuracy, and Full Accuracy gives the number of problems with status **OPTIMAL**. Time and Iterations are geometric means over the problems successfully solved by each method, including Low Accuracy results. Pairwise comparisons with Mehrotra and the Liu-type method use the matched Full Accuracy ratios in Table 2.

identical, these aggregate times are descriptive and are not used for direct pairwise comparison. On the matched Full Accuracy problems, CA-OFAS reduced the geometric-mean outer iteration count by about 15% relative to Mehrotra and about 14% relative to the Liu-type method. The corresponding computation-time reductions were about 13% and 12%, and CA-OFAS was faster on every matched problem in both comparisons. The similar reductions in iteration count and computation time indicate that the time advantage on the matched NNV-SDP problems is closely associated with the smaller number of outer iterations.

5.5 Performance Profiles for Iteration Count and Computation Time

We complement the paired tables with performance profiles in the sense of Dolan–Moré [35]. An earlier ratio-to-best display of Tits and Yang [36, pp. 1442–1443] plotted cumulative counts against a logarithmic performance-ratio axis. Dividing those counts by the number of test instances yields the same empirical cumulative form as a performance profile. The paired tables compare identical Full Accuracy instance sets directly; the profiles instead summarize iteration and time ratios across the entire test set and also incorporate Low Accuracy successes. All six implementations are included. Let $\mathcal{I}$ be the problem set, and let $\mathcal{S}_p$ be the set of methods that reached Low Accuracy or better on problem $p \in \mathcal{I}$. Let $k_{p,s}$ and $t_{p,s}$ denote the outer iteration count and computation time, respectively, for method s . When $\mathcal{S}_p \neq \emptyset$, define

$$r_{p,s}^{\text{iter}} = \begin{cases} \frac{k_{p,s}}{\min_{s' \in \mathcal{S}_p} k_{p,s'}}, & s \in \mathcal{S}_p, \\ +\infty, & s \notin \mathcal{S}_p, \end{cases} \quad r_{p,s}^{\text{time}} = \begin{cases} \frac{t_{p,s}}{\min_{s' \in \mathcal{S}_p} t_{p,s'}}, & s \in \mathcal{S}_p, \\ +\infty, & s \notin \mathcal{S}_p. \end{cases}$$

If all methods fail on a problem, we set all performance ratios to $+\infty$ and still include that problem in $|\mathcal{I}|$. For $q \in \{\text{iter}, \text{time}\}$, the performance profile is

$$\varrho_s^q(\vartheta) = \frac{1}{|\mathcal{I}|} |\{p \in \mathcal{I} : r_{p,s}^q \leq \vartheta\}|, \quad \vartheta \geq 1.$$

Hence $\varrho_s^q(1)$ gives the share of instances where method s attains the best observed value of measure q . As ϑ becomes large, the profile approaches the success fraction of that method.

Figure 1 shows the profiles for the 18 SDPLIB problems. For CA-OFAS, the time profile is more concentrated near a performance ratio of one than the iteration profile, indicating that differences

in outer iteration count do not translate directly into the same differences in total computation time. CA-OFAS was fastest on 9 of the 18 problems and required no more than 1.11 times the best time on every problem. Thus, the profile complements the paired comparisons by showing that its computation time remains close to the best observed performance throughout the SDPLIB set.

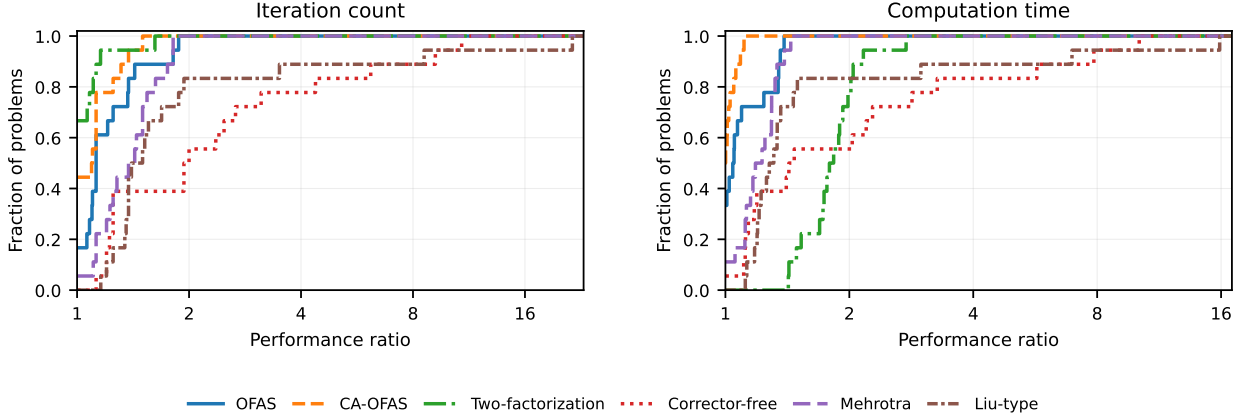


Figure 1: Performance profiles for the six implementations on 18 SDPLIB problems (success includes Low Accuracy; left: outer iteration count; right: computation time; the performance ratio is $+\infty$ for unsuccessful problems; logarithmic horizontal axis)

Figure 2 shows the profiles for the 36 NNV-SDP problems. CA-OFAS was fastest on 31 problems and required no more than 1.03 times the best time on all 34 problems that it successfully solved; its iteration profile is likewise concentrated near one. The Liu-type method succeeded on 35 problems, compared with 34 for CA-OFAS, and therefore has a slightly higher limiting profile value. These profiles complement the aggregate results by showing that CA-OFAS remains close to the best observed performance on its successful problems, while the Liu-type method succeeds on one additional problem.

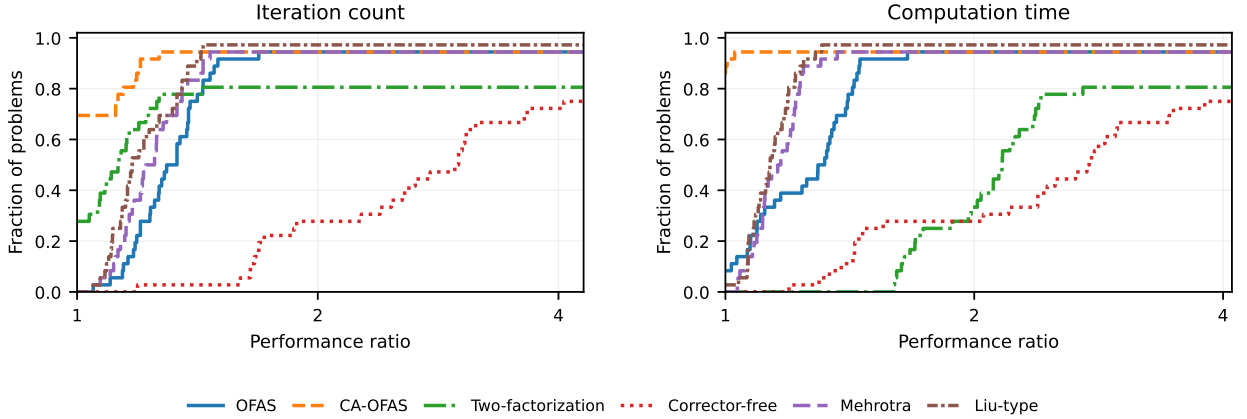


Figure 2: Performance profiles for the six implementations on 36 NNV-SDP problems (success includes Low Accuracy; left: outer iteration count; right: computation time; the performance ratio is $+\infty$ for unsuccessful problems; logarithmic horizontal axis)

The paired comparisons clarify the contributions of the proposed design. OFAS already provides substantial reductions in computation time over the structural reference methods: 44% and 38% relative to Two-factorization, and 37% and 42% relative to Corrector-free, on SDPLIB and NNV-SDP, respectively. Furthermore, CA-OFAS reduces computation time relative to OFAS by an additional 4% on SDPLIB and 19% on NNV-SDP through curvature amplification. CA-OFAS also improves on the Mehrotra-type and Liu-type methods: its computation time is reduced by 17% and 13% relative to Mehrotra, and by 20% and 12% relative to Liu, on SDPLIB and NNV-

SDP, respectively. These results show substantial gains from the one-factorization structure and post-arc correction, while curvature amplification provides a further improvement, particularly on NNV-SDP.

6 Conclusion

This work introduced OFAS, a one-factorization arc-search method that keeps both the post-arc correction and flexible candidate selection without refactorizing the KKT matrix at the accepted point. The method and its analysis are formulated within an HSD embedding. We also introduced CA-OFAS, which selectively amplifies the arc term associated with the second-order direction while keeping the same one-factorization structure. If an amplified candidate is not accepted, CA-OFAS falls back to an OFAS candidate and therefore inherits the same theoretical guarantees as OFAS.

For the theoretical analysis, we derived direction and centering estimates in a wide long-step neighborhood. Based on these estimates, we showed that the guaranteed candidate-selection rule finds an admissible candidate after finitely many halvings of the angle and damping parameter. We also established an iteration bound of $O(\nu \log(1/\varepsilon))$ for both OFAS and CA-OFAS. Under Assumption 2.2, the distance from the projectively normalized HSD sequence to the primal–dual optimal set was also shown to converge to zero. This analysis also gave a relative lower bound on the HSD scaling variable τ_k with respect to the homogeneous scale θ_k .

In the numerical experiments, we compared six methods on SDPLIB and NNV-SDP. Relative to Two-factorization, OFAS reduced both the KKT factorization count and factorization time by about 48% on SDPLIB and 40% on NNV-SDP. CA-OFAS further improved upon OFAS for many of the tested problems. These results indicate practical benefits from factorization reuse, post-arc correction, and curvature amplification.

Future work includes adaptive curvature amplification based on local curvature information, sparse SDP processing based on chordal decomposition and positive definite matrix completion, and evaluation of these ideas in high-performance implementations [37–39]. Another direction is to study the relationship between the conditioning of the HSD KKT systems and practical computational performance.

A Proofs of Technical Lemmas

A.1 Nonsingularity of the Current-Iterate KKT Operator

We prove Lemma 3.3.

Proof. Fix an arbitrary interior HSD point and denote it by W_k . We show that the homogeneous system $\mathcal{L}(W_k)[D] = (O_{\mathcal{F}}, \bar{O})$, $D = (D_X, d_y, D_S, d_\tau, d_\kappa)$ has only the zero solution $D = 0$. For brevity, set $P := P(D; W_k)$, $R := R(D; W_k)$. Since $P(W_k; W_k) = R(W_k; W_k) = Q_k^{\text{NT}}$, the complementarity row is $\mathcal{B}(W_k, D; W_k) = Q_k^{\text{NT}} \circ (P + R) = \bar{O}$. Since $Q_k^{\text{NT}} \in \mathbb{S}_{++}^n \times \mathbb{R}_{++}$, the Jordan multiplication operator $L_{Q_k^{\text{NT}}} : U \mapsto Q_k^{\text{NT}} \circ U$ is invertible. Indeed, let $q_1, \dots, q_n > 0$ be the eigenvalues of the PSD block $Q_{X,k}$. After an orthogonal diagonalization of $Q_{X,k}$, $L_{Q_k^{\text{NT}}}$ multiplies each matrix component by the positive scalar $(q_i + q_j)/2$. On the scalar block, it also acts by multiplication with a positive scalar. Therefore,

$$P + R = \bar{O}. \quad (51)$$

The three HSD residual rows give

$$\begin{aligned} \mathcal{A}(D_X) - b d_\tau &= 0, \\ \mathcal{A}^*(d_y) + D_S - C d_\tau &= 0, \\ C \bullet D_X - b^T d_y + d_\kappa &= 0. \end{aligned}$$

The second equation gives $D_S = Cd_\tau - \mathcal{A}^*(d_y)$. Hence, using the first equation and the adjoint relation,

$$\begin{aligned} D_X \bullet D_S &= d_\tau C \bullet D_X - D_X \bullet \mathcal{A}^*(d_y) = d_\tau C \bullet D_X - \mathcal{A}(D_X)^T d_y \\ &= d_\tau \{C \bullet D_X - b^T d_y\} = -d_\tau d_\kappa. \end{aligned}$$

Thus, $D_X \bullet D_S + d_\tau d_\kappa = 0$. Moreover, by the definition of the NT coordinates in (5) and cyclicity of the trace,

$$\begin{aligned} \langle P, R \rangle_{\mathcal{E}} &= \text{tr} \left(M_k D_X M_k^T M_k^{-T} D_S M_k^{-1} \right) + (\xi_k d_\tau)(\xi_k^{-1} d_\kappa) \\ &= D_X \bullet D_S + d_\tau d_\kappa = 0. \end{aligned}$$

Equation (51) gives $R = -P$, so $0 = \langle P, R \rangle_{\mathcal{E}} = -\|P\|_{\mathcal{E}}^2$. Hence, $P = R = \bar{O}$. Since M_k and ξ_k are invertible, we obtain $D_X = D_S = O$ and $d_\tau = d_\kappa = 0$. The dual residual row then gives $\mathcal{A}^*(d_y) = O$. By Assumption 2.1, $\mathcal{A}^*$ is injective, and therefore $d_y = 0$. Thus, $\ker \mathcal{L}(W_k) = \{0\}$. The current-iterate KKT system is finite-dimensional and square. Hence, injectivity implies bijectivity, and $\mathcal{L}(W_k)$ is nonsingular. $\square$

A.2 Technical Estimates for Uniformly Guaranteed OFAS Candidates

This subsection gives the technical estimates for the fixed choice $\rho_0 = 1/8$ used in Proposition 4.5. We separate the constant estimates from the main argument and verify three facts: (i) D_ζ is uniformly bounded away from zero, (ii) strict interiority is preserved for $\beta = \Theta(1/\nu)$, and (iii) the wide long-step neighborhood is recovered on the same damping scale. The constants below depend only on γ , the fixed value ρ_0 , and the prescribed upper bound on the angle. They are independent of ν and the problem data.

The candidate set in Algorithm 1 contains $\rho \in \{1, 1/2, 1/4, 1/8\}$. The choice $\rho_0 = 1/8$ is used only to construct a theoretical guarantee and does not exclude the other values of ρ . For the finite-termination proof, we fix the smallest value ρ_0 and show that the successive-halving fallback search reaches an admissible candidate when β is on the $O(1/\nu)$ scale, while retaining an $O(1/\nu)$ angle. This search requires no additional KKT solve. It uses only linear combinations of the correction directions already constructed and the unit-complementarity direction.

Throughout this subsection, fix iteration k and use the abbreviations

$$\begin{aligned} W &:= W_k, & \mu &:= \mu(W_k), & Q &:= Q_k^{\text{NT}}, & H &:= Q \circ Q, \\ \sigma &:= \sin \alpha, & \psi &:= 1 - \cos \alpha, & \mu_{\text{tar}} &:= (1 - \sigma)\mu. \end{aligned}$$

In the NT coordinates defined by the current iterate W_k , write the predictor-arc factors and their displacements as

$$\begin{aligned} P(\alpha) &:= P(W_k(\alpha); W_k), & R(\alpha) &:= R(W_k(\alpha); W_k), \\ \Delta P(\alpha) &:= P(\alpha) - Q, & \Delta R(\alpha) &:= R(\alpha) - Q. \end{aligned}$$

We also write $\Delta_{\text{arc}}(\alpha) := W_k(\alpha) - W_k$ and $H_{\text{arc}}(\alpha) := P(\alpha) \circ R(\alpha)$. The quantities D_ζ , N_ζ , and ζ are defined as in Lemma 4.1(iii) with $\rho = \rho_0$. We use the notation $d_\gamma(c)$, $r_\gamma(c)$, and $\mathcal{E}_{\text{arc}}(\alpha)$ from Lemma 4.4.

Let $0 < c \leq 1/2$ and assume $\sigma \leq c/\nu$. For the correction direction $D_k^{\text{cor}}(\alpha, \rho_0)$, define its NT components by

$$(P^{\text{cor}}, R^{\text{cor}}) := (P(D_k^{\text{cor}}(\alpha, \rho_0); W_k), R(D_k^{\text{cor}}(\alpha, \rho_0); W_k)).$$

By the definition in Section 3, this correction direction is

$$D_k^{\text{cor}}(\alpha, \rho_0) = \rho_0(1 - \sigma)D_k^{\text{cen}} + \sigma\psi D_k^{\text{mix}} + \psi^2 D_k^{\text{rem}}.$$

Therefore, using the three squared-norm bounds in Proposition 4.3,

$$\sqrt{\|P_k^{\text{cen}}\|^2 + \|R_k^{\text{cen}}\|^2} \leq \sqrt{\nu(\gamma^{-1} - 1)\mu},$$

$$\sqrt{\|P_k^{\text{mix}}\|^2 + \|R_k^{\text{mix}}\|^2} \leq \frac{\nu^{3/2}}{\gamma} \sqrt{\mu}, \quad \sqrt{\|P_k^{\text{rem}}\|^2 + \|R_k^{\text{rem}}\|^2} \leq \frac{1+\gamma}{2\gamma^{3/2}} \nu^2 \sqrt{\mu},$$

and the inequality $\|P\| + \|R\| \leq \sqrt{2}\sqrt{\|P\|^2 + \|R\|^2}$, we obtain

$$\|P^{\text{cor}}\| + \|R^{\text{cor}}\| \leq \sqrt{2} \left(\rho_0(1-\sigma)\sqrt{\nu(\gamma^{-1}-1)\mu} + \sigma\psi\frac{\nu^{3/2}}{\gamma}\sqrt{\mu} + \psi^2\frac{1+\gamma}{2\gamma^{3/2}}\nu^2\sqrt{\mu} \right).$$

Since $0 \leq \psi = 1 - \cos \alpha \leq \sin^2 \alpha = \sigma^2$, $\sigma \leq c/\nu$, and $\nu \geq 1$, it follows that

$$\|P^{\text{cor}}\| + \|R^{\text{cor}}\| \leq A_\gamma(c)\sqrt{\nu\mu},$$

where

$$A_\gamma(c) := \sqrt{2} \left(\rho_0\sqrt{\gamma^{-1}-1} + \frac{c^3}{\gamma} + \frac{1+\gamma}{2\gamma^{3/2}}c^4 \right).$$

This gives the required bound for the correction direction $D_k^{\text{cor}}(\alpha, \rho_0)$.

Let $(P_k^{\text{unit}}, R_k^{\text{unit}})$ be the NT components of the unit-complementarity direction D_k^{unit} . Define also the complementarity-row target for $D_k^{\text{cor}}(\alpha, \rho_0)$ by $T_0(\alpha) := (1-\sigma)((1-\rho_0)H + \rho_0\mu\bar{I})$. This is exactly $T_k(\alpha, \rho)$ from Section 3 with $\rho = \rho_0$. The following exact cancellations show that D_ζ remains close to ν and that ζ is small.

Lemma A.1 (Cancellation in ζ for the fixed- ρ_0 corrector). *For $D_k^{\text{cor}}(\alpha, \rho_0)$, the following identities hold exactly:*

$$\begin{aligned} D_\zeta &= \nu + \langle \Delta P(\alpha), R_k^{\text{unit}} \rangle + \langle P_k^{\text{unit}}, \Delta R(\alpha) \rangle, \\ N_\zeta &= -\langle \Delta P(\alpha), R^{\text{cor}} \rangle - \langle P^{\text{cor}}, \Delta R(\alpha) \rangle. \end{aligned}$$

Proof. For the scalar denominator D_ζ defined in Lemma 4.1(iii), the unit-complementarity direction gives $\bar{I}$ in the complementarity row in the current-iterate NT coordinates. Hence, its cross complementarity with W_k is $\Gamma(W, D_k^{\text{unit}}) = \text{Tr}_{\mathcal{E}}(\bar{I}) = \nu$. Using the arc displacement $\Delta_{\text{arc}}(\alpha)$ defined above and bilinearity, we obtain

$$\begin{aligned} D_\zeta &= \Gamma(W_k(\alpha), D_k^{\text{unit}}) = \Gamma(W, D_k^{\text{unit}}) + \Gamma(\Delta_{\text{arc}}(\alpha), D_k^{\text{unit}}) \\ &= \nu + \langle \Delta P(\alpha), R_k^{\text{unit}} \rangle + \langle P_k^{\text{unit}}, \Delta R(\alpha) \rangle. \end{aligned}$$

Thus, only the two arc-displacement cross terms perturb the denominator from ν .

By (17), the linearity of $\mathcal{G}(\cdot; W)$, and (15),

$$\begin{aligned} \mathcal{B}(W, D_k^{\text{cor}}(\alpha, \rho_0); W) &= -\Delta_k^{\text{comp}}(\alpha, \rho_0) \\ &= T_0(\alpha) - \Phi(W_k(\alpha); W). \end{aligned}$$

Taking $\text{Tr}_{\mathcal{E}}$ and using $\text{Tr}_{\mathcal{E}}(\Phi(V; W)) = \mathcal{C}(V)$, $\text{Tr}_{\mathcal{E}}(\mathcal{B}(U, V; W)) = \Gamma(U, V)$, and $\text{Tr}_{\mathcal{E}}(T_0(\alpha)) = \nu\mu_{\text{tar}}$ gives

$$\Gamma(W, D_k^{\text{cor}}(\alpha, \rho_0)) = \nu\mu_{\text{tar}} - \mathcal{C}(W_k(\alpha)).$$

On the other hand, the exact numerator of ζ is

$$\begin{aligned} N_\zeta &= \nu\mu_{\text{tar}} - \mathcal{C}(W_k(\alpha)) - \Gamma(W_k(\alpha), D_k^{\text{cor}}(\alpha, \rho_0)) \\ &= \nu\mu_{\text{tar}} - \mathcal{C}(W_k(\alpha)) - \Gamma(W, D_k^{\text{cor}}(\alpha, \rho_0)) - \Gamma(\Delta_{\text{arc}}(\alpha), D_k^{\text{cor}}(\alpha, \rho_0)). \end{aligned}$$

The first three terms cancel exactly by the identity above, and hence $N_\zeta = -\Gamma(\Delta_{\text{arc}}(\alpha), D_k^{\text{cor}}(\alpha, \rho_0))$. Expanding this cross complementarity in the current-iterate NT coordinates gives

$$\Gamma(\Delta_{\text{arc}}(\alpha), D_k^{\text{cor}}(\alpha, \rho_0)) = \langle \Delta P(\alpha), R^{\text{cor}} \rangle + \langle P^{\text{cor}}, \Delta R(\alpha) \rangle.$$

Therefore,

$$N_\zeta = -\langle \Delta P(\alpha), R^{\text{cor}} \rangle - \langle P^{\text{cor}}, \Delta R(\alpha) \rangle.$$

□

We now use the following bounds for the arc displacement and the unit-complementarity direction:

$$\|\Delta P(\alpha)\| + \|\Delta R(\alpha)\| \leq d_\gamma(c) \sqrt{\frac{\mu}{\nu}}, \quad \|P_k^{\text{unit}}\| + \|R_k^{\text{unit}}\| \leq \sqrt{\frac{2\nu}{\gamma\mu}}.$$

Applying Cauchy–Schwarz to $D_\zeta - \nu = \langle \Delta P(\alpha), R_k^{\text{unit}} \rangle + \langle P_k^{\text{unit}}, \Delta R(\alpha) \rangle$ gives $|D_\zeta - \nu| \leq d_\gamma(c) \sqrt{2/\gamma}$. Hence, if $d_\gamma(c) \sqrt{2/\gamma} \leq 1/2$, then $\nu \geq 1$ implies $D_\zeta \geq \nu - 1/2 \geq \nu/2$.

Similarly, using the full-corrector bound $\|P^{\text{cor}}\| + \|R^{\text{cor}}\| \leq A_\gamma(c) \sqrt{\nu\mu}$, and the arc-displacement bound above to $N_\zeta = -\langle \Delta P(\alpha), R^{\text{cor}} \rangle - \langle P^{\text{cor}}, \Delta R(\alpha) \rangle$ gives $|N_\zeta| \leq d_\gamma(c) A_\gamma(c) \mu$. Therefore,

$$|\zeta| \leq 2d_\gamma(c) A_\gamma(c) \frac{\mu}{\nu}.$$

Let $D^{\text{full}} := D_k^{\text{cor}}(\alpha, \rho_0) + \zeta D_k^{\text{unit}}$ be the full composite correction, and define its NT components by $(P^{\text{full}}, R^{\text{full}}) := (P(D^{\text{full}}; W_k), R(D^{\text{full}}; W_k))$. Then,

$$\|P^{\text{full}}\| + \|R^{\text{full}}\| \leq B_\gamma(c) \sqrt{\nu\mu}, \quad (52)$$

where

$$B_\gamma(c) := A_\gamma(c) \left(1 + 2d_\gamma(c) \sqrt{\frac{2}{\gamma}} \right). \quad (53)$$

For use in Lemmas A.2 and A.3, define

$$\begin{aligned} \bar{d}_\gamma &:= \sqrt{2} + \sqrt{2/\gamma}, & \bar{r}_\gamma &:= \gamma^{-1/2} + \frac{1 + \gamma^{-1}}{2}, \\ \bar{A}_\gamma &:= \sqrt{2} \left(\rho_0 \sqrt{\gamma^{-1} - 1} + \frac{1}{8\gamma} + \frac{1 + \gamma}{32\gamma^{3/2}} \right), & \bar{B}_\gamma &:= \bar{A}_\gamma \left(1 + 2\bar{d}_\gamma \sqrt{2/\gamma} \right). \end{aligned}$$

Moreover, set

$$b_\gamma := \min \left\{ 1, \frac{\sqrt{\gamma}}{8\bar{B}_\gamma}, \frac{\rho_0(1 - \gamma)}{64\bar{B}_\gamma^2} \right\} > 0, \quad (54)$$

and

$$c_\gamma := \min \left\{ \sigma_{\max}, \frac{1}{2}, \frac{\sqrt{\gamma}}{4\bar{d}_\gamma}, \frac{\rho_0(1 - \gamma)}{384\bar{d}_\gamma \bar{B}_\gamma}, \left(\frac{b_\gamma \rho_0(1 - \gamma)}{128\bar{r}_\gamma} \right)^{1/3}, \sqrt{\frac{1 - \sigma_{\max}}{64\bar{r}_\gamma}}, \frac{1}{2\sqrt{\bar{r}_\gamma}} \right\} > 0. \quad (55)$$

These constants depend only on γ , the fixed value ρ_0 , and $\sigma_{\max}$, and are independent of ν and the problem data. For $d_\gamma(c) = \sqrt{2}c + \sqrt{2/\gamma}c^2$ and $r_\gamma(c) = c^3/\sqrt{\gamma} + c^4/2 + c^4/(2\gamma)$ from Lemma 4.4, if $0 < c \leq 1/2$, then $d_\gamma(c) \leq \bar{d}_\gamma c$ and $r_\gamma(c) \leq \bar{r}_\gamma c^3$. Also, $c^3 \leq 1/8$ and $c^4 \leq 1/16$ imply $A_\gamma(c) \leq \bar{A}_\gamma$. Substituting $d_\gamma(c) \leq \bar{d}_\gamma c \leq \bar{d}_\gamma$ into $B_\gamma(c) = A_\gamma(c)(1 + 2d_\gamma(c)\sqrt{2/\gamma})$ gives $B_\gamma(c) \leq \bar{B}_\gamma$. The three bounds in (54) control, respectively, the damping range, strict interiority, and the quadratic correction term in neighborhood recovery. The seven bounds in (55) control, in order, the angle limit, the condition $c \leq 1/2$, D_ζ and strict interiority, the mixed-correction error, the cubic arc remainder, the projective quantity $\mathcal{J}$, and small-angle progress. In particular, this choice gives $r_\gamma(c_\gamma) \leq c_\gamma/4$.

Lemma A.2 (Strict interiority under fixed ρ_0 and $O(1/\nu)$ damping). *Let $W \in \mathcal{N}_{\text{long}}(\gamma)$, let $0 < \sigma \leq c_\gamma/\nu$, and set $\rho = \rho_0$. Assume further that*

$$\frac{b_\gamma}{4\nu} \leq \beta \leq \frac{b_\gamma}{\nu}.$$

Then the uncorrected arc and the damped corrected point $W^+(\beta) = W_k(\alpha) + \beta D^{\text{full}}$ are strictly interior.

Proof. The current-iterate NT factors of the uncorrected arc are $P(\alpha) = Q + \Delta P(\alpha)$, $R(\alpha) = Q + \Delta R(\alpha)$. We use the arc-displacement estimate $\|\Delta P(\alpha)\| + \|\Delta R(\alpha)\| \leq d_\gamma(c_\gamma)\sqrt{\frac{\mu}{\nu}}$ for $\sigma \leq c_\gamma/\nu$. For the full composite correction $D^{\text{full}} = D_k^{\text{cor}}(\alpha, \rho_0) + \zeta D_k^{\text{unit}}$, (52) and (53) give

$$\|P^{\text{full}}\| + \|R^{\text{full}}\| \leq B_\gamma(c_\gamma)\sqrt{\nu\mu} \leq \bar{B}_\gamma\sqrt{\nu\mu}.$$

Therefore, the two NT factors of the damped corrected point $W^+(\beta) = W_k(\alpha) + \beta D^{\text{full}}$, $P^+(\beta) = P(\alpha) + \beta P^{\text{full}}$, $R^+(\beta) = R(\alpha) + \beta R^{\text{full}}$, satisfy, using $\beta \leq b_\gamma/\nu$,

$$\begin{aligned} \|P^+(\beta) - Q\| &\leq \|\Delta P(\alpha)\| + \beta\|P^{\text{full}}\| \\ &\leq (d_\gamma(c_\gamma) + b_\gamma\bar{B}_\gamma)\sqrt{\frac{\mu}{\nu}}, \\ \|R^+(\beta) - Q\| &\leq \|\Delta R(\alpha)\| + \beta\|R^{\text{full}}\| \\ &\leq (d_\gamma(c_\gamma) + b_\gamma\bar{B}_\gamma)\sqrt{\frac{\mu}{\nu}}. \end{aligned}$$

Since $d_\gamma(c) \leq \bar{d}_\gamma c$ and $c_\gamma \leq \sqrt{\gamma}/(4\bar{d}_\gamma)$, we have $d_\gamma(c_\gamma) \leq \frac{\sqrt{\gamma}}{4}$. Also, $b_\gamma \leq \sqrt{\gamma}/(8\bar{B}_\gamma)$ gives $b_\gamma\bar{B}_\gamma \leq \frac{\sqrt{\gamma}}{8}$. Using $\nu \geq 1$, each factor perturbation is therefore bounded by

$$(d_\gamma(c_\gamma) + b_\gamma\bar{B}_\gamma)\sqrt{\frac{\mu}{\nu}} \leq \frac{3\sqrt{\gamma}}{8}\sqrt{\frac{\mu}{\nu}} \leq \frac{3}{8}\sqrt{\gamma\mu} < \frac{1}{2}\sqrt{\gamma\mu}.$$

Since $W \in \mathcal{N}_{\text{long}}(\gamma)$, the current complementarity product $H = Q \circ Q$ satisfies $\lambda_{\min}(H) \geq \gamma\mu$. Hence, the current-iterate NT factor satisfies $\lambda_{\min}(Q) \geq \sqrt{\gamma\mu}$. Applying the spectral perturbation inequality $\lambda_{\min}(Q + \Delta) \geq \lambda_{\min}(Q) - \|\Delta\|$ to $P^+(\beta) = Q + (P^+(\beta) - Q)$ and $R^+(\beta) = Q + (R^+(\beta) - Q)$ gives $\lambda_{\min}(P^+(\beta))$, $\lambda_{\min}(R^+(\beta)) > \frac{1}{2}\sqrt{\gamma\mu} > 0$. Thus, $W^+(\beta)$ is strictly interior. The same argument with $\beta = 0$ shows that $P(\alpha)$ and $R(\alpha)$ are positive definite. Hence, the uncorrected arc is also strictly interior. $\square$

Let $H^+(\beta)$ denote the complementarity product of the damped point in the current-iterate NT coordinates. To prove neighborhood recovery, we decompose $H^+(\beta)$ into a leading centering term and higher-order error terms. The full-corrector equation gives

$$H^+(\beta) = (1 - \beta)H_{\text{arc}}(\alpha) + \beta(T_0(\alpha) + \zeta\bar{I}) + \beta(\Delta P(\alpha) \circ R^{\text{full}} + P^{\text{full}} \circ \Delta R(\alpha)) + \beta^2 P^{\text{full}} \circ R^{\text{full}}, \quad (56)$$

where $T_0(\alpha) = (1 - \sigma)((1 - \rho_0)H + \rho_0\mu\bar{I})$. Define $\tilde{T}(\beta) := (1 - \sigma)((1 - \beta\rho_0)H + \beta\rho_0\mu\bar{I})$. This is the leading target after damping reduces the centering strength from ρ_0 to $\beta\rho_0$. Substituting $H_{\text{arc}}(\alpha) = (1 - \sigma)H + \mathcal{E}_{\text{arc}}(\alpha)$ from Lemma 4.4 into (56), we define the remaining error by

$$\mathcal{E}_{\text{damp}}(\beta) := H^+(\beta) - \tilde{T}(\beta).$$

The next lemma shows that this error is smaller than the positive spectral margin of $\tilde{T}(\beta)$.

Lemma A.3 (Damped neighborhood recovery with fixed ρ_0). *Let $W \in \mathcal{N}_{\text{long}}(\gamma)$, and suppose that*

$$0 < \sigma \leq \frac{c_\gamma}{\nu}, \quad \rho = \rho_0 = \frac{1}{8}, \quad \frac{b_\gamma}{4\nu} \leq \beta \leq \frac{b_\gamma}{\nu}.$$

Then

$$\|\mathcal{E}_{\text{damp}}(\beta)\| \leq \left(r_\gamma(c_\gamma) + 3b_\gamma d_\gamma(c_\gamma)\bar{B}_\gamma + \frac{1}{4}b_\gamma^2\bar{B}_\gamma^2 \right) \frac{\mu}{\nu}.$$

Moreover,

$$|\mu(W^+(\beta)) - (1 - \sigma)\mu| \leq r_\gamma(c_\gamma)\frac{\mu}{\nu^2}. \quad (57)$$

Therefore, in the NT coordinates defined by the current iterate W_k ,

$$\lambda_{\min}(\Phi(W^+(\beta); W_k)) \geq \gamma\mu(W^+(\beta)).$$

In particular, $W^+(\beta) \in \mathcal{N}_{\text{long}}(\gamma)$.

Proof. The complementarity product after damping is

$$H^+(\beta) = (1 - \beta)H_{\text{arc}}(\alpha) + \beta(T_0(\alpha) + \zeta\bar{I}) + \beta(\Delta P(\alpha) \circ R^{\text{full}} + P^{\text{full}} \circ \Delta R(\alpha)) + \beta^2 P^{\text{full}} \circ R^{\text{full}},$$

where

$$T_0(\alpha) = (1 - \sigma)((1 - \rho_0)H + \rho_0\mu\bar{I}), \quad H_{\text{arc}}(\alpha) = (1 - \sigma)H + \mathcal{E}_{\text{arc}}(\alpha).$$

Hence, using $\tilde{T}(\beta)$ defined above, we have $H^+(\beta) = \tilde{T}(\beta) + \mathcal{E}_{\text{damp}}(\beta)$, where

$$\mathcal{E}_{\text{damp}}(\beta) = (1 - \beta)\mathcal{E}_{\text{arc}}(\alpha) + \beta\zeta\bar{I} + \beta(\Delta P(\alpha) \circ R^{\text{full}} + P^{\text{full}} \circ \Delta R(\alpha)) + \beta^2 P^{\text{full}} \circ R^{\text{full}}. \quad (58)$$

We bound the four terms in (58). The arc remainder, arc displacement, full composite correction, and ζ satisfy

$$\begin{aligned} \|\mathcal{E}_{\text{arc}}(\alpha)\| &\leq r_\gamma(c_\gamma) \frac{\mu}{\nu^{3/2}}, \\ \|\Delta P(\alpha)\| + \|\Delta R(\alpha)\| &\leq d_\gamma(c_\gamma) \sqrt{\frac{\mu}{\nu}}, \\ \|P^{\text{full}}\| + \|R^{\text{full}}\| &\leq B_\gamma(c_\gamma) \sqrt{\nu\mu} \leq \bar{B}_\gamma \sqrt{\nu\mu}, \\ |\zeta| &\leq 2d_\gamma(c_\gamma) A_\gamma(c_\gamma) \frac{\mu}{\nu} \leq 2d_\gamma(c_\gamma) \bar{B}_\gamma \frac{\mu}{\nu}. \end{aligned}$$

Also, $\|\bar{I}\| = \sqrt{\nu}$ and $\beta \leq b_\gamma/\nu$. For the mixed Jordan-product term, we use

$$\|P \circ \hat{R} + \hat{P} \circ R\| \leq \sqrt{\|P\|^2 + \|R\|^2} \sqrt{\|\hat{P}\|^2 + \|\hat{R}\|^2}$$

with $(P, R) = (\Delta P(\alpha), \Delta R(\alpha))$ and $(\hat{P}, \hat{R}) = (P^{\text{full}}, R^{\text{full}})$. For the quadratic term, we use

$$\|P^{\text{full}} \circ R^{\text{full}}\| \leq \|P^{\text{full}}\| \|R^{\text{full}}\| \leq \frac{1}{4} (\|P^{\text{full}}\| + \|R^{\text{full}}\|)^2,$$

These estimates give the following bounds for the four terms in (58):

$$\begin{aligned} (1 - \beta)\|\mathcal{E}_{\text{arc}}(\alpha)\| &\leq r_\gamma(c_\gamma) \frac{\mu}{\nu}, \\ \beta|\zeta|\|\bar{I}\| &\leq 2b_\gamma d_\gamma(c_\gamma) \bar{B}_\gamma \frac{\mu}{\nu}, \\ \beta\|\Delta P(\alpha) \circ R^{\text{full}} + P^{\text{full}} \circ \Delta R(\alpha)\| &\leq b_\gamma d_\gamma(c_\gamma) \bar{B}_\gamma \frac{\mu}{\nu}, \\ \beta^2\|P^{\text{full}} \circ R^{\text{full}}\| &\leq \frac{1}{4} b_\gamma^2 \bar{B}_\gamma^2 \frac{\mu}{\nu}. \end{aligned}$$

Here we used $\nu^{-3/2} \leq \nu^{-1}$ because $\nu \geq 1$. Adding these four inequalities gives

$$\|\mathcal{E}_{\text{damp}}(\beta)\| \leq \left(r_\gamma(c_\gamma) + 3b_\gamma d_\gamma(c_\gamma) \bar{B}_\gamma + \frac{1}{4} b_\gamma^2 \bar{B}_\gamma^2 \right) \frac{\mu}{\nu}.$$

Next, we estimate the average complementarity. Since the residual-invariant full correction D^{full} satisfies $\mu(W_k(\alpha) + D^{\text{full}}) = (1 - \sigma)\mu$, damping gives the exact interpolation $\mu(W^+(\beta)) = (1 - \beta)\mu(W_k(\alpha)) + \beta(1 - \sigma)\mu$. Also,

$$H_{\text{arc}}(\alpha) = (1 - \sigma)H + \mathcal{E}_{\text{arc}}(\alpha), \quad \mu(W_k(\alpha)) = \frac{1}{\nu} \text{Tr}_{\mathcal{E}}(H_{\text{arc}}(\alpha)), \quad \text{Tr}_{\mathcal{E}}(H) = \nu\mu.$$

Therefore,

$$\mu(W^+(\beta)) - (1 - \sigma)\mu = \frac{1 - \beta}{\nu} \text{Tr}_{\mathcal{E}}(\mathcal{E}_{\text{arc}}(\alpha)).$$

Using $\text{Tr}_{\mathcal{E}}(Z) = \langle \bar{I}, Z \rangle$, $\|\bar{I}\| = \sqrt{\nu}$, and $\|\mathcal{E}_{\text{arc}}(\alpha)\| \leq r_\gamma(c_\gamma)\mu/\nu^{3/2}$, we obtain

$$|\mu(W^+(\beta)) - (1 - \sigma)\mu| \leq \frac{\sqrt{\nu}}{\nu} r_\gamma(c_\gamma) \frac{\mu}{\nu^{3/2}} = r_\gamma(c_\gamma) \frac{\mu}{\nu^2},$$

which proves (57).

Finally, since $W \in \mathcal{N}_{\text{long}}(\gamma)$ and $\lambda_{\min}(H) \geq \gamma\mu$,

$$\begin{aligned}\lambda_{\min}(\tilde{T}(\beta)) &\geq (1 - \sigma)((1 - \beta\rho_0)\gamma\mu + \beta\rho_0\mu) \\ &= (1 - \sigma)(\gamma + \beta\rho_0(1 - \gamma))\mu.\end{aligned}$$

Moreover, (57) and $\nu \geq 1$ give

$$\gamma\mu(W^+(\beta)) \leq \gamma(1 - \sigma)\mu + \gamma r_\gamma(c_\gamma) \frac{\mu}{\nu}.$$

Thus, by the spectral perturbation inequality, it is enough to show

$$\lambda_{\min}(\tilde{T}(\beta)) - \gamma(1 - \sigma)\mu \geq \|\mathcal{E}_{\text{damp}}(\beta)\| + \gamma r_\gamma(c_\gamma) \frac{\mu}{\nu}.$$

For the left-hand side, $\sigma \leq c_\gamma/\nu \leq 1/2$ and $\beta \geq b_\gamma/(4\nu)$ imply

$$\lambda_{\min}(\tilde{T}(\beta)) - \gamma(1 - \sigma)\mu \geq \frac{32}{256}b_\gamma\rho_0(1 - \gamma)\frac{\mu}{\nu}.$$

On the other hand, the definitions of b_γ and c_γ , together with $d_\gamma(c) \leq \bar{d}_\gamma c$ and $r_\gamma(c) \leq \bar{r}_\gamma c^3$, give

$$\begin{aligned}\frac{1}{4}b_\gamma^2\bar{B}_\gamma^2 &\leq \frac{b_\gamma\rho_0(1 - \gamma)}{256}, \\ 3b_\gamma\bar{d}_\gamma c_\gamma\bar{B}_\gamma &\leq \frac{b_\gamma\rho_0(1 - \gamma)}{128}, \\ (1 + \gamma)\bar{r}_\gamma c_\gamma^3 &\leq \frac{b_\gamma\rho_0(1 - \gamma)}{64}.\end{aligned}$$

Hence the right-hand side above is at most

$$\frac{7}{256}b_\gamma\rho_0(1 - \gamma)\frac{\mu}{\nu}.$$

Since $7/256 < 32/256$, the required inequality holds. Therefore,

$$\lambda_{\min}(H^+(\beta)) \geq \lambda_{\min}(\tilde{T}(\beta)) - \|\mathcal{E}_{\text{damp}}(\beta)\| \geq \gamma\mu(W^+(\beta)),$$

where $H^+(\beta) = \tilde{T}(\beta) + \mathcal{E}_{\text{damp}}(\beta) = \Phi(W^+(\beta); W_k)$.

The preceding positivity result also gives strict interiority of $W^+(\beta)$. The estimate above is the current-iterate NT certificate

$$\lambda_{\min}(\Phi(W^+(\beta); W_k)) \geq \gamma\mu(W^+(\beta)).$$

By Lemma 4.2, the true complementarity spectrum of the candidate therefore satisfies the wide long-step condition. Hence $W^+(\beta) \in \mathcal{N}_{\text{long}}(\gamma)$. $\square$

A.3 Quantitative Scale Estimates for Projective Normalization

This subsection proves the projective scale estimates used in Subsection 4.5.

By (43), $D_k^{\text{inv}} := W_k - \omega_k W_0$ is a residual-invariant direction. Applying Lemma 4.1(ii) to D_k^{inv} and using $W_0 = (I, 0, I, 1, 1)$, $\nu = n + 1$, and $\mu(W_k) = \theta_k \omega_k$, we obtain

$$\text{tr } X_k + \text{tr } S_k + \tau_k + \kappa_k = \nu(\theta_k + \omega_k). \quad (59)$$

Let $Z^* = (X^*, y^*, S^*) \in \mathcal{Z}^*$ be the point fixed in Subsection 4.5, and let $\delta_{\text{rec}}(Z^*)$ be defined by (44).

Lemma A.4 (Positive recession-direction margin). *Under Assumptions 2.1–2.2, $\delta_{\text{rec}}(Z^*) > 0$. The boundedness part of Assumption 2.2 excludes a nonzero direction along which a primal or dual optimal solution can be extended indefinitely while remaining optimal.*

Proof. The set of (D_X, D_S) satisfying $D_X, D_S \succeq O$ and $\text{tr } D_X + \text{tr } D_S = 1$ is compact. Moreover, Assumption 2.1 implies that $\mathcal{A}^*$ is injective. A finite-dimensional injective linear map is bounded away from zero on the unit sphere. Since D_S is bounded on the compact set above, it follows that $\|\mathcal{A}^*(d_y) + D_S\| \rightarrow \infty$ as $\|d_y\| \rightarrow \infty$. Hence the infimum in (44) is attained.

Suppose, to the contrary, that $\delta_{\text{rec}}(Z^*) = 0$. Then an attaining minimizer satisfies $\mathcal{A}(D_X) = 0$, $\mathcal{A}^*(d_y) + D_S = O$, $S^* \bullet D_X = 0$, and $X^* \bullet D_S = 0$. If $D_X \neq O$, dual feasibility of (y^*, S^*) gives $C = \mathcal{A}^*(y^*) + S^*$, and therefore

$$C \bullet D_X = y^{*T} \mathcal{A}(D_X) + S^* \bullet D_X = 0.$$

Together with $\mathcal{A}(D_X) = 0$ and $D_X \succeq O$, this shows that the ray $X^* + tD_X$, $t \geq 0$, remains primal feasible and satisfies $C \bullet (X^* + tD_X) = C \bullet X^*$. Every point on this ray is therefore primal optimal. If $D_X = O$, the normalization $\text{tr } D_X + \text{tr } D_S = 1$ implies $D_S \neq O$. Using primal feasibility $\mathcal{A}(X^*) = b$ and $\mathcal{A}^*(d_y) + D_S = O$, we obtain

$$b^T d_y = X^* \bullet \mathcal{A}^*(d_y) = -X^* \bullet D_S = 0.$$

Consequently, $(y^* + td_y, S^* + tD_S)$ stays dual feasible for all $t \geq 0$, while $b^T(y^* + td_y) = b^T y^*$. Hence the entire ray consists of dual optimal points. Either case contradicts the boundedness condition in Assumption 2.2. $\square$

We now prove Lemma 4.10. Set $\mathcal{T}_k^{\text{tr}} := \text{tr } X_k + \text{tr } S_k > 0$. Applying the definition of $\delta_{\text{rec}}(Z^*)$ to the normalized direction $(X_k, y_k, S_k)/\mathcal{T}_k^{\text{tr}}$ gives

$$\delta_{\text{rec}}(Z^*) \mathcal{T}_k^{\text{tr}} \leq \|\mathcal{A}(X_k)\| + \|\mathcal{A}^*(y_k) + S_k\| + S^* \bullet X_k + X^* \bullet S_k.$$

The HSD residual relations give the following bound for the first two terms: $(\|b\| + \|C\|)\tau_k + 2\mathcal{Q}_k$. Primal-dual optimality also gives

$$S^* \bullet X_k + X^* \bullet S_k = C \bullet X_k - b^T y_k - y^{*T} R_p(W_k) + X^* \bullet R_d(W_k).$$

Using $C \bullet X_k - b^T y_k = R_g(W_k) - \kappa_k$ and $\kappa_k \geq 0$, we obtain the upper bound $(1 + \|y^*\| + \|X^*\|)\mathcal{Q}_k$. This proves (45).

The same identity gives $0 \leq \kappa_k \leq (1 + \|y^*\| + \|X^*\|)\mathcal{Q}_k$. By (59) and $\omega_k \geq 0$,

$$\nu\theta_k \leq \mathcal{T}_k^{\text{tr}} + \tau_k + \kappa_k \leq \mathfrak{D}_* \tau_k + \mathfrak{H}_* \mathcal{Q}_k.$$

This proves the first lower bound in (46). If, in addition, $\mathfrak{H}_* \mathcal{Q}_k \leq \nu\theta_k/2$, the second lower bound follows immediately.

B Per-Instance Numerical Results

This appendix reports the computation time, the number of outer iterations, and the termination status for each of the 54 problems used in Section 5. Each cell shows “computation time (seconds)/number of outer iterations.” Computation times are displayed to one decimal place, while the unrounded values are used for the aggregated results.

We use MathOptInterface (MOI) termination statuses throughout the experiments. The Clarabel.jl statuses `SOLVED` and `ALMOST_SOLVED` correspond to `OPTIMAL` and `ALMOST_OPTIMAL` in MOI, respectively. See Goulart and Chen [7, Sec. 2.6] for the termination criteria of Clarabel. Under our experimental settings, `OPTIMAL` requires

$$\frac{\kappa}{\tau} \leq 1, \quad (\text{gap}_{\text{abs}} < 10^{-8} \text{ or } \text{gap}_{\text{rel}} < 10^{-8}), \quad r_p < 10^{-8}, \quad r_d < 10^{-8}.$$

Here gap_{abs} and gap_{rel} denote Clarabel.jl’s absolute and relative duality-gap measures. Its normalized residual measures for primal and dual feasibility are denoted by r_p and r_d . If these full-accuracy conditions are not satisfied before termination due to the iteration limit, time limit, numerical error, or insufficient progress, the solution is checked again using the relaxed tolerances

$$\frac{\kappa}{\tau} \leq 1, \quad (\text{gap}_{\text{abs}} < 5 \times 10^{-5} \text{ or } \text{gap}_{\text{rel}} < 5 \times 10^{-5}), \quad r_p < 10^{-4}, \quad r_d < 10^{-4}.$$

If these conditions hold, the status is `ALMOST_OPTIMAL`. Following the numerical evaluation of Clarabel [7], we classify `OPTIMAL` as Full Accuracy and `ALMOST_OPTIMAL` as Low Accuracy. The statuses `TIME_LIMIT` and `SLOW_PROGRESS` correspond to `MAX_TIME` and `INSUFFICIENT_PROGRESS` in Clarabel.jl, respectively, when the relaxed conditions are also not satisfied.

The label `SUCCESS_MIXED_ACCURACY` is not a termination status of Clarabel.jl or MOI. For a problem–method pair with repeated runs, this label is used when the runs include both `OPTIMAL` and `ALMOST_OPTIMAL`. The success count including Low Accuracy includes `OPTIMAL`, `ALMOST_OPTIMAL`, and `SUCCESS_MIXED_ACCURACY`. The Full Accuracy success count includes only `OPTIMAL`.

For the finite safeguarded candidate search used in the numerical implementation of OFAS and CA-OFAS, we set $\sigma_{\text{max}} = 0.95$. The angle backtracking factor was 0.5, with at most 50 backtracking steps. We used the centering-recovery schedule $\rho \in \{1, 1/2, 1/4, 1/8\}$ and damped the corrector by a factor of 0.5 down to $\beta_{\text{min}} = 1/64$. Warm starts, refinement, recovery, and amplification-activation safeguards were used only in the numerical search; the candidate formulas and acceptance conditions (C1)–(C4) are unchanged.

Table B.1 reports linear-algebra ratios of OFAS to Two-factorization for problems solved to Full Accuracy by both methods.

Table B.1: Linear-algebra ratios of OFAS to Two-factorization for problems solved to Full Accuracy by both methods

Problem set	Matched	Factorizations	Factorization time	Linear solves	Solve time
SDPLIB	11	0.524	0.523	0.873	0.894
NNV-SDP	26	0.601	0.606	1.001	1.053

Each entry is the geometric mean of the OFAS/Two-factorization ratio over problems solved to Full Accuracy by both methods. Ratios below one favor OFAS.

A cell without a superscript denotes `OPTIMAL`. We use the superscripts A for `ALMOST_OPTIMAL`, M for `SUCCESS_MIXED_ACCURACY`, TL for `TIME_LIMIT`, and SP for `SLOW_PROGRESS`. For `TIME_LIMIT`, the displayed time may slightly exceed 3600 seconds because an ongoing linear-algebra operation is not interrupted.

Table B.2 reports the results for the 18 SDPLIB problems, and Table B.3 reports those for the 36 NNV-SDP problems. In the NNV-SDP problem names, `ellipsoid` and `rectangle` denote the input sets considered in Sections 5.2 and 5.3 of Azuma–Kim–Yamashita [34], respectively. The labels `dense` and `identity` describe the structure of the first-layer weight matrix W^0 , where `identity` means $W^0 = I$. This corresponds to Assumption 5.7 of that paper for rectangular inputs. The label `nXX` denotes the network width $r = n_0 = n_1$, and `s701`, `s702`, and `s703` denote random seeds 701, 702, and 703, respectively. Thus, three seeds are used for each problem family and network width. In Table B.3, the problems are ordered by increasing r . For the same r , they are ordered by problem family, weight-matrix structure, and seed. In both tables, the shortest computation time and the smallest number of outer iterations among the successful methods are shown in bold for each problem.

Table B.2: Per-instance results for the 18 SDPLIB problems

Problem	OFAS	CA-OFAS	Two- factorization	Corrector- free	Mehrotra	Liu-type
control4	4.7/27	4.5/25	8.7/27	9.2/59	5.0/32	6.0/38
control5	18.3/32	20.3/35	39.6/37	52.0/100	19.3/36	23.0/43
control6	59.1/34	55.2/31	109.6/33	125.8/77	63.6/38	70.6/42
gpp100	854.4/26 ^M	634.3/19	1296.1/21 ^M	3621.4/117 ^M	708.7/21 ^M	708.9/22 ^M
gpp124-1	3496.0/30 ^M	2602.9/22	3701.3/16 ^M	3658.8/31 ^M	3435.8/29	3546.7/30
gpp124-2	2595.8/22^M	2823.3/24	3703.0/16 ^M	3702.1/32 ^M	3361.3/29	3134.5/27
gpp124-3	3452.4/29	2486.2/21	3699.1/16 ^M	3654.1/31 ^M	3317.6/28	3634.8/31
mcp100	325.9/9	313.8/9	548.0/8	310.6/9	387.4/11	371.8/11
mcp124-1	1125.0/9	1115.9/9	2102.9/9	1330.2/11	1552.0/13	1669.2/14
mcp124-2	1114.1/9	1115.3/9	1887.0/8	1246.5/10	1442.9/12	1330.1/11
mcp124-3	1116.4/9	1137.4/9	1938.4/8	1236.5/10	1609.0/13	1492.0/12
mcp124-4	1134.9/9	1110.9/9	1917.6/8	1241.1/10	1437.3/12	1356.4/11
qap6	2.6/30 ^M	1.9/21	5.3/34 ^M	4.3/56 ^M	2.3/30 ^M	13.4/180 ^M
qap7	12.2/29^M	12.9/30 ^M	18.6/24 ^M	39.9/105 ^M	14.4/37 ^M	193.6/515 ^M
theta1	4.5/10	4.5/10	8.2/10	5.1/12	5.0/12	5.8/14
theta2	382.5/11	366.1/11	661.9/10	410.8/12	349.4/10	393.8/12
truss5	4.6/17	4.5/16 ^M	7.8/16	34.2/147	4.4/18	13.0/56
truss8	417.4/20 ^M	336.6/16	679.2/18 ^M	3407.3/173	413.2/20	396.3/20

Table B.3: Per-instance results for the 36 NNV-SDP problems (in ascending order of network width r)

Problem	OFAS	CA-OFAS	Two-factorization	Corrector-free	Mehrotra	Liu-type
rectangle_dense_n32_s701	82.0/37	66.3/30	131.4/32	136.0/68	70.6/34	70.6/34
rectangle_dense_n32_s702	81.2/37	62.8/29	126.7/31	150.0/74	69.7/34	68.8/33
rectangle_dense_n32_s703	77.9/36	60.2/28	118.4/29	132.5/67	67.1/33	64.2/31
ellipsoid_dense_n40_s701	403.1/46	283.0/31	666.1/39	774.7/92	350.4/39	337.6/39
ellipsoid_dense_n40_s702	391.9/44	293.0/33	597.2/35	872.5/104	330.1/38	343.0/40
ellipsoid_dense_n40_s703	385.0/44	284.4/32	628.3/36	991.0/116	341.0/39	318.8/37
rectangle_dense_n40_s701	463.8/51	329.5/37	712.6/42	825.7/98	158.9/18 ^{SP}	378.5/43
rectangle_dense_n40_s702	440.7/48	315.4/35	754.7/43	773.0/91	380.7/44	356.2/41
rectangle_dense_n40_s703	466.8/51	332.4/37	753.8/43	957.0/113	378.1/43	361.6/41
rectangle_identity_n40_s701	165.6/18	162.9/18	268.3/15	217.7/25	176.4/20	173.5/19
rectangle_identity_n40_s702	170.3/18	165.0/18	264.3/15	214.2/24	197.4/22	181.2/20
rectangle_identity_n40_s703	147.9/16	147.9/16	247.2/14	212.0/24	178.6/20	182.1/20
rectangle_dense_n44_s701	690.3/43	550.4/34	1160.8/37	1459.5/94	618.7/39	586.5/37
rectangle_dense_n44_s702	398.5/24 ^{SP}	203.6/12 ^{SP}	438.0/14 ^{SP}	557.2/35 ^{SP}	168.9/10 ^{SP}	3205.9/207 ^{SP}
rectangle_dense_n44_s703	760.2/48	573.3/36	1209.4/39	1369.1/89	627.2/40	609.9/39
ellipsoid_dense_n48_s701	1053.2/41	772.9/29	1655.7/32	2660.2/106	901.9/35	864.6/34
ellipsoid_dense_n48_s702	1454.7/54	875.9/32^A	2371.3/46	3329.3/130	1196.1/46	1145.4/44
ellipsoid_dense_n48_s703	1077.2/39 ^A	923.6/33	1998.9/38	2545.6/99	995.8/38	1046.5/40
rectangle_dense_n48_s701	1222.6/46	842.6/32	2029.6/39	2495.3/97	1098.9/41	996.1/37
rectangle_dense_n48_s702	1519.0/56	1061.3/39	2377.7/45	3001.0/117	1227.6/47	1149.5/43
rectangle_dense_n48_s703	1217.9/44	844.2/30	2003.8/38	2359.7/92	988.8/36	962.0/35
rectangle_identity_n48_s701	534.7/20	484.7/18	821.7/16	702.9/27	602.4/22	602.7/22
rectangle_identity_n48_s702	517.3/19	530.8/19	844.8/16	714.4/27	614.4/23	627.9/23
rectangle_identity_n48_s703	519.2/19	521.7/19	830.1/15	740.7/28	537.8/20	561.1/20
ellipsoid_dense_n52_s701	2267.5/54	1686.2/39	3643.2/44 ^A	3608.9/86 ^{TL}	2043.7/49	2006.5/48
ellipsoid_dense_n52_s702	2099.2/49 ^A	1482.2/34 ^{SP}	3621.8/43 ^A	3628.9/88 ^{TL}	2002.0/48	1922.9/45
ellipsoid_dense_n52_s703	2018.7/48	1531.1/36	3621.6/43	3642.2/89 ^{TL}	1885.4/45	1755.5/41
ellipsoid_dense_n56_s701	2627.9/39	1975.0/29	3651.5/28 ^{TL}	3638.1/56 ^{TL}	2041.8/31	2048.5/31
ellipsoid_dense_n56_s702	3630.4/54 ^{TL}	3245.4/48	3665.4/28 ^{TL}	3660.8/55 ^{TL}	3609.5/54 ^A	3446.7/53
ellipsoid_dense_n56_s703	3543.4/54	2436.0/36	3614.1/27 ^{TL}	3650.8/56 ^{TL}	2992.8/45	2818.2/43
rectangle_identity_n56_s701	1411.9/21	1316.2/19	2234.2/17	1885.8/28	1469.3/22	1565.4/23
rectangle_identity_n56_s702	1511.0/22	1383.3/20	2229.3/17	2146.7/32	1534.3/23	1628.3/24
rectangle_identity_n56_s703	1362.3/20	1271.8/19	2206.5/17	1864.3/28	1546.6/23	1546.1/23
rectangle_identity_n64_s701	3296.7/22	3101.7/21	3850.8/13 ^{TL}	3702.4/25 ^A	3279.2/22	3420.5/23
rectangle_identity_n64_s702	3248.5/22	2832.7/19	3796.9/13 ^{TL}	3644.2/25 ^{TL}	3444.3/23	3640.3/25
rectangle_identity_n64_s703	3312.1/22	2968.3/20	3864.0/13 ^{TL}	3714.8/25 ^{TL}	3239.8/22	3360.6/23

Acknowledgements

ChatGPT 5.6 was used for language editing and editorial assistance. The authors reviewed the resulting manuscript and remain responsible for its content.

Statements and Declarations

Funding. The research of Makoto Yamashita was partially supported by Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Numbers 24K14836 and 26K02868.

Competing interests. The authors declare no competing interests related to this work.

Data availability. The scripts in the public repository ClarabelHSDSearch.jl can be used to obtain or generate the data for the numerical experiments.

Code availability. Implementations of the proposed methods and the numerical-experiment code are available in the public repository ClarabelHSDSearch.jl.

References

- [1] Lieven Vandenbergh and Stephen Boyd. Semidefinite programming. *SIAM Review*, 38(1): 49–95, 1996. <https://doi.org/10.1137/1038003>.
- [2] Farid Alizadeh. Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM Journal on Optimization*, 5(1):13–51, 1995. <https://doi.org/10.1137/0805002>.
- [3] Makoto Yamashita, Katsuki Fujisawa, and Masakazu Kojima. Implementation and evaluation of SDPA 6.0 (semidefinite programming algorithm 6.0). *Optimization Methods and Software*, 18(4):491–505, 2003. <https://doi.org/10.1080/1055678031000118482>.
- [4] Reha H. Tütüncü, Kim-Chuan Toh, and Michael J. Todd. Solving semidefinite-quadratic-linear programs using SDPT3. *Mathematical Programming*, 95(2):189–217, 2003. <https://doi.org/10.1007/s10107-002-0347-5>.
- [5] Jos F. Sturm. Using SeDuMi 1.02, a matlab toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11–12:625–653, 1999. <https://doi.org/10.1080/10556789908805766>.
- [6] Stefano Cipolla and Jacek Gondzio. Proximal-stabilized semidefinite programming. *Computational Optimization and Applications*, 91:573–616, 2025. <https://doi.org/10.1007/s10589-024-00614-3>.
- [7] Paul J. Goulart and Yuwen Chen. Clarabel: An interior-point solver for conic programs with quadratic objectives. *Mathematical Programming Computation*, 2026. <https://doi.org/10.1007/s12532-026-00320-7>.
- [8] Stefano Cipolla and Jacek Gondzio. Proximal stabilized interior point methods and low-frequency-update preconditioning techniques. *Journal of Optimization Theory and Applications*, 197:1061–1103, 2023. <https://doi.org/10.1007/s10957-023-02194-4>.
- [9] Alan J. Goldman and Albert W. Tucker. Theory of linear programming. In Harold W. Kuhn and Albert W. Tucker, editors, *Linear Inequalities and Related Systems*, number 38 in Annals of Mathematics Studies, pages 53–98. Princeton University Press, Princeton, NJ, 1956.

- [10] Yinyu Ye, Michael J. Todd, and Shinji Mizuno. An $O(\sqrt{n}L)$ -iteration homogeneous and self-dual linear programming algorithm. *Mathematics of Operations Research*, 19(1):53–67, 1994. <https://doi.org/10.1287/moor.19.1.53>.
- [11] Yurii Nesterov and Michael J. Todd. Self-scaled barriers and interior-point methods for convex programming. *Mathematics of Operations Research*, 22(1):1–42, 1997. <https://doi.org/10.1287/moor.22.1.1>.
- [12] Yurii Nesterov and Michael J. Todd. Primal-dual interior-point methods for self-scaled cones. *SIAM Journal on Optimization*, 8(2):324–364, 1998. <https://doi.org/10.1137/S1052623495290209>.
- [13] Yaguang Yang. A polynomial arc-search interior-point algorithm for convex quadratic programming. *European Journal of Operational Research*, 215(1):25–38, 2011. <https://doi.org/10.1016/j.ejor.2011.06.020>.
- [14] Yaguang Yang. A polynomial arc-search interior-point algorithm for linear programming. *Journal of Optimization Theory and Applications*, 158(3):859–873, 2013. <https://doi.org/10.1007/s10957-013-0281-0>.
- [15] Yaguang Yang and Makoto Yamashita. An arc-search $O(nL)$ infeasible-interior-point algorithm for linear programming. *Optimization Letters*, 12:781–798, 2018. <https://doi.org/10.1007/s11590-017-1142-9>.
- [16] Ximei Yang, Hongwei Liu, and Yinkui Zhang. An arc-search infeasible-interior-point method for symmetric optimization in a wide neighborhood of the central path. *Optimization Letters*, 11(1):135–152, 2017. <https://doi.org/10.1007/s11590-016-0997-5>.
- [17] Behrouz Kheirfam and Morteza Moslemi. On the extension of an arc-search interior-point algorithm for semidefinite optimization. *Numerical Algebra, Control and Optimization*, 8(2):261–275, 2018. <https://doi.org/10.3934/naco.2018015>.
- [18] Mingwang Zhang, Beibei Yuan, Yiyuan Zhou, Xiaoyu Luo, and Zhengwei Huang. A primal-dual interior-point algorithm with arc-search for semidefinite programming. *Optimization Letters*, 13(5):1157–1175, 2019. <https://doi.org/10.1007/s11590-019-01414-z>.
- [19] Behrouz Kheirfam. A polynomial-iteration infeasible interior-point algorithm with arc-search for semidefinite optimization. *Journal of Scientific Computing*, 88(3):Paper No. 89, 2021. <https://doi.org/10.1007/s10915-021-01609-6>.
- [20] Behrouz Kheirfam, Naser Osmanpour, and Mohammad Keyanpour. An arc-search infeasible interior-point method for semidefinite optimization with the negative infinity neighborhood. *Numerical Algorithms*, 88:143–163, 2021. <https://doi.org/10.1007/s11075-020-01033-1>.
- [21] Yaguang Yang. Two computationally efficient polynomial-iteration infeasible interior-point algorithms for linear programming. *Numerical Algorithms*, 79(3):957–992, 2018. <https://doi.org/10.1007/s11075-018-0469-3>.
- [22] Yaguang Yang. A polynomial time infeasible interior-point arc-search algorithm for convex optimization. *Optimization and Engineering*, 24(2):885–914, 2023. <https://doi.org/10.1007/s11081-022-09712-9>.
- [23] Einosuke Iida and Makoto Yamashita. An infeasible interior-point arc-search method with nesterov’s restarting strategy for linear programming problems. *Computational Optimization and Applications*, 88:643–676, 2024. <https://doi.org/10.1007/s10589-024-00561-z>.
- [24] Yaguang Yang. An arc-search interior-point algorithm for nonlinear constrained optimization. *Computational Optimization and Applications*, 90(3):969–995, 2025. <https://doi.org/10.1007/s10589-025-00648-1>.

- [25] Einosuke Iida and Makoto Yamashita. An inexact infeasible arc-search interior-point method for linear optimization problems. *Numerical Algorithms*, 2026. <https://doi.org/10.1007/s11075-026-02374-z>.
- [26] Yaguang Yang. *Arc-Search Techniques for Interior-Point Methods*. CRC Press, Boca Raton, 2020. <https://doi.org/10.1201/9781003042518>.
- [27] Sanjay Mehrotra. On the implementation of a primal-dual interior point method. *SIAM Journal on Optimization*, 2:575–601, 1992. <https://doi.org/10.1137/0802028>.
- [28] Changhe Liu and Hongwei Liu. A new second-order corrector interior-point algorithm for semidefinite programming. *Mathematical Methods of Operations Research*, 75(2):165–183, 2012. <https://doi.org/10.1007/s00186-012-0379-4>.
- [29] Etienne de Klerk, Cornelis Roos, and Tamás Terlaky. Initialization in semidefinite programming via a self-dual, skew-symmetric embedding. *Operations Research Letters*, 20(5):213–221, 1997. [https://doi.org/10.1016/S0167-6377\(97\)00011-4](https://doi.org/10.1016/S0167-6377(97)00011-4).
- [30] Florian A. Potra and Rongqin Sheng. On homogeneous interior-point algorithms for semidefinite programming. *Optimization Methods and Software*, 9(1–3):161–184, 1998. <https://doi.org/10.1080/10556789808805691>.
- [31] Renato D. C. Monteiro and Yin Zhang. A unified analysis for a class of long-step primal-dual path-following interior-point algorithms for semidefinite programming. *Mathematical Programming*, 81(3):281–299, 1998. <https://doi.org/10.1007/BF01580085>.
- [32] Michael J. Todd, Kim-Chuan Toh, and Reha H. Tütüncü. On the Nesterov–Todd direction in semidefinite programming. *SIAM Journal on Optimization*, 8(3):769–796, 1998. <https://doi.org/10.1137/S105262349630060X>.
- [33] Brian Borchers. SDPLIB 1.2, a library of semidefinite programming test problems. *Optimization Methods and Software*, 11(1–4):683–690, 1999. <https://doi.org/10.1080/10556789908805769>.
- [34] Godai Azuma, Sunyoung Kim, and Makoto Yamashita. Verifying robustness of neural networks with tight semidefinite relaxations. *Computational Optimization and Applications*, 94(1):41–72, 2026. <https://doi.org/10.1007/s10589-026-00765-5>.
- [35] Elizabeth D. Dolan and Jorge J. Moré. Benchmarking optimization software with performance profiles. *Mathematical Programming*, 91(2):201–213, 2002. <https://doi.org/10.1007/s101070100263>.
- [36] André L. Tits and Yaguang Yang. Globally convergent algorithms for robust pole assignment by state feedback. *IEEE Transactions on Automatic Control*, 41(10):1432–1452, 1996. <https://doi.org/10.1109/9.539425>.
- [37] Mituhiro Fukuda, Masakazu Kojima, Kazuo Murota, and Kazuhide Nakata. Exploiting sparsity in semidefinite programming via matrix completion i: General framework. *SIAM Journal on Optimization*, 11(3):647–674, 2001. <https://doi.org/10.1137/S1052623400366218>.
- [38] Kazuhide Nakata, Katsuki Fujisawa, Mituhiro Fukuda, Masakazu Kojima, and Kazuo Murota. Exploiting sparsity in semidefinite programming via matrix completion ii: Implementation and numerical results. *Mathematical Programming*, 95(2):303–327, 2003. <https://doi.org/10.1007/s10107-002-0351-9>.
- [39] Makoto Yamashita and Kazuhide Nakata. Fast implementation for semidefinite programs with positive matrix completion. *Optimization Methods and Software*, 30(5):1030–1049, 2015. <https://doi.org/10.1080/10556788.2015.1014554>.