LINE-SEARCH AND TRUST-REGION EQUATIONS FOR A PRIMAL-DUAL INTERIOR METHOD FOR NONLINEAR OPTIMIZATION

Philip E. Gill*Vyacheslav Kungurtsev†Daniel P. Robinson‡

UCSD Center for Computational Mathematics
Technical Report CCoM-21-4
September 1, 2021

Abstract

The approximate Newton equations for a minimizing a shifted primal-dual penalty-barrier method are derived for a nonlinearly constrained problem in general form. These equations may be used in conjunction with either a line-search or trust-region method to force convergence from an arbitrary starting point. It is shown that under certain conditions, the approximate Newton equations are equivalent to a regularized form of the conventional primal-dual path-following equations.

Key words. Nonlinear programming, nonlinear constraints, shifted penalty-barrier methods, augmented Lagrangian methods, primal-dual interior methods, path-following methods, regularized methods.

AMS subject classifications. 49J20, 49J15, 49M37, 49D37, 65F05, 65K05, 90C30

*Department of Mathematics, University of California, San Diego, La Jolla, CA 92093-0112 (pgill@ucsd.edu). Research supported in part by National Science Foundation grants DMS-1318480 and DMS-1361421. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

†Agent Technology Center, Department of Computer Science, Faculty of Electrical Engineering, Czech Technical University in Prague. (vyacheslav.kungurtsev@fel.cvut.cz) Research supported by the OP VVV project CZ.02.1.01/0.0/0.0/16 019/0000765 “Research Center for Informatics”.

‡Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, PA 18015 (dpr219@lehigh.edu). Research supported in part by National Science Foundation grant DMS-1217153. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.
1. Introduction

This note concerns that derivation of the line-search and trust-region equations for a shifted primal-dual penalty-barrier merit method for constrained optimization. These methods are intended for the minimization of a twice-continuously differentiable function subject to both equality and inequality constraints that may include a set of twice-continuously differentiable constraint functions. A description of the line-search and trust-region methods for a problem with nonlinear inequality constraints is given by Gill, Kungurtsev and Robinson [4] and Gill, Kungurtsev and Robinson [5]. The note concerns the formulation of the equations for problems written in the general form:

\[
\text{minimize}_{x \in \mathbb{R}^n, s \in \mathbb{R}^m} f(x) \quad \text{subject to} \quad \begin{cases} c(x) - s = 0, & L_X s = h_X, \\ A x = b, & E_X x = b_X, \quad \ell^x \leq E_L x, \quad E_U x \leq u^x, \end{cases} \quad \text{(NLP)}
\]

where \(A\) denotes a constant \(m_A \times n\) matrix, and \(b, h_X, b_X, \ell^s, u^s, \ell^x\) and \(u^x\) are fixed vectors of dimension \(m_A, m_X, n_X, m_L, m_U, n_L\) and \(n_U\), respectively. Similarly, \(L_X, L_L\) and \(L_U\) denote fixed matrices of dimension \(m_X \times m, m_L \times m\) and \(m_U \times m\), respectively, and \(E_X, E_L\) and \(E_U\) are fixed matrices of dimension \(n_X \times n, n_L \times n\) and \(n_U \times n\), respectively. Throughout the discussion, the functions \(c : \mathbb{R}^n \mapsto \mathbb{R}^m\) and \(f : \mathbb{R}^n \mapsto \mathbb{R}\) are assumed to be twice-continuously differentiable. The components of \(s\) may be interpreted as slack variables associated with the nonlinear constraints.

The quantity \(E_X\) denotes an \(n_X \times n\) matrix formed from \(n_X\) independent rows of \(I_n\), the identity matrix of order \(n\). This implies that the equality constraints \(E_X x = b_X\) fix \(n_X\) components of \(x\) at the corresponding values of \(b_X\). Similarly, \(E_L\) and \(E_U\) denote \(n_L \times n\) and \(n_U \times n\) matrices formed from subsets of rows of \(I_m\) such that \(E_L^T E_L = 0\), \(E_U^T E_U = 0\), i.e., a variable is either fixed or free to move, possibly bounded by an upper or lower bound. Note that an \(x_j\) may be an unrestricted variable in the sense that it is neither fixed nor subject to an upper or lower bound, in which case \(e_j^T\) is not a row of \(E_X, E_L\) or \(E_U\). Analogous definitions hold for \(L_X, L_L\) and \(L_U\) as subsets of rows of \(I_m\). However, we impose the restriction that a given \(s_j\) must be either fixed or restricted by an upper or lower bound, i.e., there are no unrestricted slacks.\(^1\) Let \(E_F\) denote the matrix of rows of \(I_n\) that are not rows of \(E_X\), and let \(L_F\) denote the matrix of rows of \(I_m\) that are not rows of \(L_X\). If \(n_F = n - n_X\) and \(m_F = m - m_X\), then \(E_F\) and \(L_F\) are \(n_F \times n\) and \(m_F \times m\) respectively. Note that \(n_L + n_U\) may be less than \(n_F\), but \(m_F\) must equal \(m_L + m_U\). The matrices \((E_X^T \quad E_F^T)\) and \((L_X^T \quad L_F^T)\) are column permutations of \(I_n\) and \(I_m\). Moreover, there are \(n \times n\) and \(m \times m\) permutation matrices \(P_x\) and \(P_s\) such that

\[
P_x = \begin{pmatrix} E_F \\ E_X \end{pmatrix} \quad \text{and} \quad P_s = \begin{pmatrix} L_F \\ L_X \end{pmatrix}, \quad (1.1)
\]

with \(E_F E_F^T = I_F, E_X E_X^T = I_X, \) and \(E_F E_X^T = 0, \) and \(L_F L_F^T = I_F, L_X L_X^T = I_X, \) and \(L_F E_F^T = 0.\)

All general inequality constraints are imposed indirectly using a shifted primal-dual barrier function. The general equality constraints \(c(x) - s = 0\) and \(A x = b\) are enforced using an primal-dual augmented Lagrangian algorithm, which implies that the

\(^{1}\)This is not a significant restriction because a “free” slack is equivalent to a unrestricted nonlinear constraint, which may be discarded from the problem. The shifted primal-dual penalty-barrier equations can be derived without this restriction, but the derivation is beyond the scope of this note.
equalities are satisfied in the limit. The exception to this is when the constraints $E_X x = b_X$, and $L_X s = h_X$ are used to fix a subset of the variables and slacks. These bounds are enforced at every iterate.

An equality constraint $c_i(x) = 0$ may be handled by introducing the slack variable $s_i$ and writing the constraint as the two constraints $c_i(x) - s_i = 0$ and $s_i = 0$. In this case the $i$th coordinate vector $e_i$ can be included as a row of $L_X$. Linear inequality constraints must be included as part of $c$. A linear equality constraint can be either included with the nonlinear equality constraints or the matrix $A$. The constraints involving $A$ may be used to temporarily fix a subset of the variables at their bounds without altering the underlying structure of the approximate Newton equations. In this case, the associated rows of $A$ are rows of the identity matrix.

The optimality conditions for problem (NLP) are given in Section 2. The shifted path-following equations are formulated in Section 3. The shifted primal-dual penalty-barrier function associated with problem is discussed in Section 4. This function serves as a merit function for both the line-search and trust-region method. The equations for a line-search modified Newton method are formulated in Sections 5 and 6, and summarized in Section 7. The analogous equations for the trust-region method are derived in Section 8 and summarized in Section 9.

**Notation.** Given vectors $x$ and $y$, the vector consisting of $x$ augmented by $y$ is denoted by $(x, y)$. The subscript $i$ is appended to vectors to denote the $i$th component of that vector, whereas the subscript $k$ is appended to a vector to denote its value during the $k$th iteration of an algorithm, e.g., $x_k$ represents the value for $x$ during the $k$th iteration, whereas $[x_k]_i$ denotes the $i$th component of the vector $x_k$. Given vectors $a$ and $b$ with the same dimension, the vector with $i$th component $a_i b_i$ is denoted by $a \cdot b$. Similarly, $\min(a, b)$ is a vector with components $\min(a_i, b_i)$. The vector $e$ denotes the column vector of ones, and $I$ denotes the identity matrix. The dimensions of $e$ and $I$ are defined by the context. The vector two-norm or its induced matrix norm are denoted by $\| \cdot \|$. The vector $\nabla f(x)$ is used to denote the gradient of $f(x)$. The matrix $J(x)$ denotes the $m \times n$ constraint Jacobian, which has $i$th row $\nabla c_i(x)^T$. Given a Lagrangian function $L(x, y) = f(x) - c(x)^T y$ with $y$ a $m$-vector of dual variables, the Hessian of the Lagrangian with respect to $x$ is denoted by $H(x, y) = \nabla^2 f(x) - \sum_{i=1}^m y_i \nabla^2 c_i(x)$. Both the line-search and trust-region equations utilize the Moore-Penrose pseudoinverse of a diagonal matrix. In particular, if $D = \text{diag}(d_1, d_2, \ldots, d_n)$, then the pseudoinverse $D^\dagger$ is diagonal with $D^\dagger_{ii} = 0$ for $d_i = 0$ and $D^\dagger_{ii} = 1/d_i$ for $d_i \neq 0$. 

2. Optimality conditions

The first-order KKT conditions for problem (NLP) are

$$\begin{align*}
\nabla f(x^*) - J(x^*)^T y^* - A^T v^* - E^T x^* z^*_x - E^T_l z^*_1 + E^T_u z^*_2 &= 0, \\
y^* - L^T_x w_x - L^T_l w^*_1 + L^T_u w^*_2 &= 0, \\
c(x^*) - s^* &= 0, \\
x^* - s^* &= u^x - E^*_ux^* \geq 0, \\
L_s s^* - \ell^x &= 0, \\
E_l x^* - \ell^x &= 0, \\
z^*_1 \cdot (E_l x^* - \ell^x) &= 0, \\
w^*_1 \cdot (L_s s^* - \ell^x) &= 0, \\
Ax^* - b &= 0, \\
E_l x^* - b^x &= 0, \\
L_s s^* - h_x &= 0, \\
z^*_2 \cdot (u^x - E^*_ux^*) &= 0, \\
w^*_2 \cdot (u^s - L^*_us^*) &= 0,
\end{align*}$$

(2.1)

where \( y^*, w^*_x, \) and \( z^*_X \) are the multipliers for the equality constraints \( c(x) - s = 0, L_s s^* = h_X \) and \( E_l x^* = b_X, \) and \( z^*_1, z^*_2, w^*_1 \) and \( w^*_2 \) may be interpreted as the Lagrange multipliers for the inequality constraints \( E_l x - \ell^x \geq 0, u^x - E^*_ux \geq 0, L_s s - \ell^s \geq 0 \) and \( u^s - L^*_us \geq 0, \) respectively. The components of \( v^* \) are the multipliers for the linear equality constraints \( Ax = b. \) If \( x_1 = E_l x - \ell^x, \) \( x_2 = u^x - E^*_ux, \) \( s_1 = L_s s - \ell^s, \) and \( s_2 = u^s - L^*_us, \) then \( z^*_1, z^*_2, w^*_1, \) and \( w^*_2 \) are the Lagrange multipliers for the inequality constraints \( x_1 \geq 0, x_2 \geq 0, s_1 \geq 0, \) and \( s_2 \geq 0, \) respectively. In the derivations that follow, the vectors \( z \) and \( w \) are defined as

$$z = E^T_x z_x + E^T_l z_1 + E^T_u z_2, \quad w = L^T_x w_x + L^T_l w_1 + L^T_u w_2,$$

(2.2)

3. The path-following equations

Let \( z^E_1, w^E_1, \) and \( w^E_2 \) denote nonnegative estimates of \( z^*_1, w^*_1, \) and \( w^*_2. \) Given small positive scalars \( \mu^E, \mu^s, \) and \( \mu^h, \) consider the perturbed optimality conditions

$$\begin{align*}
\nabla f(x) - J(x)^T y - A^T v - E^T x^* z^E_x - E^T_l z^E_1 + E^T_u z^E_2 &= 0, \\
y - L^T_x w_x - L^T_l w^E_1 + L^T_u w^E_2 &= 0, \\
c(x) - s &= \mu^E(y^E - y), \\
Ax - b &= \mu^s(v^E - v), \\
E_l x - \ell^E &= 0, \\
L_s s - \ell^s &= 0, \\
z_1 \cdot (E_l x - \ell^E) &= \mu^h(z^E_1 - z_1), \\
w_1 \cdot (L_s s - \ell^s) &= \mu^h(w^E_1 - w_1), \\
z_2 \cdot (u^x - E^*_ux) &= \mu^h(z^E_2 - z_2), \\
w_2 \cdot (u^s - L^*_us) &= \mu^h(w^E_2 - w_2),
\end{align*}$$

(3.1)
Let \( v_p \) denote the vector of variables \( v_p = (x, s, y, v, w_x, z_x, z_1, z_2, w_1, w_2) \). The primal-dual path-following equations are given by \( F(v_p) \), with

\[
F(v_p) = \begin{pmatrix}
\nabla f(x) - J(x)^T y - A^T v - E_X^T z_x - E_{11}^T z_1 + E_{12}^T z_2 \\
y - L_{v_x} w_x - L_{v_1} w_1 + L_{v_2} w_2 \\
c(x) - s + \mu^v(y - y^e) \\
A x - b + \mu^s(v - v^e) \\
E_X x - b_x \\
L_{x s} - h_x \\
z_1 \cdot (E_{x} x - \ell^x) + \mu^s(z_1 - z_1^e) \\
z_2 \cdot (u^x - E_{0} x) + \mu^s(z_2 - z_2^e) \\
w_1 \cdot (L_{s} s - \ell^s) + \mu^\nu(w_1 - w_1^e) \\
w_2 \cdot (u^s - L_{0} s) + \mu^\nu(w_2 - w_2^e)
\end{pmatrix}
\]

(To simplify the notation, the dependence of \( F \) on the parameters \( \mu^s, \mu^v, y^e, v^e, z_1^e, z_2^e, w_1^e, w_2^e \) is omitted.) Any zero \((x, s, y, v, w_x, z_x, z_1, z_2, w_1, w_2)\) of \( F \) such that \( \ell^x \leq E_{1} x < u^x, \ell^s \leq L_{1} s, L_{0} < u^s, z_{1} > 0, z_{2} > 0, w_{1} > 0, \) and \( w_{2} > 0 \) approximates a point satisfying the optimality conditions (2.1), with the approximation becoming increasingly accurate as the terms \( \mu^v(y - y^e), \mu^s(v - v^e), \mu^s(z_1 - z_1^e), \mu^s(z_2 - z_2^e), \mu^\nu(w_1 - w_1^e) \) and \( \mu^\nu(w_2 - w_2^e) \) approach zero. For any sequence of \( z_1^e, z_2^e, w_1^e, w_2^e \) and \( y^e \) such that \( z_1^e \rightarrow z_1^e, z_2^e \rightarrow z_2^e, w_1^e \rightarrow w_1^e, w_2^e \rightarrow w_2^e \), \( \mu^s \) and \( \mu^\nu \) approach zero, and it must hold that solutions \((x, s, y, z_1, z_2, w_1, w_2)\) of (3.1) will approximate a solution of (2.1) independently of the values of \( \mu^s, \mu^v \) and \( \mu^\nu \) (i.e., it is not necessary that \( \mu^v \rightarrow 0, \mu^s \rightarrow 0 \) and \( \mu^\nu \rightarrow 0 \)).

If \( v_p = (x, s, y, v, w_x, z_x, z_1, z_2, w_1, w_2) \) is a given approximate zero of \( F \) such that \( \ell^x - \mu^v < E_{1} x, E_{0} x < u^x + \mu^v, \ell^s - \mu^s < L_{1} s, L_{0} < u^s + \mu^s, z_{1} > 0, z_{2} > 0, w_{1} > 0, \) and \( w_{2} > 0 \), the Newton equations for the change in variables \( \Delta v_p = (\Delta x, \Delta s, \Delta y, \Delta v, \Delta w_x, \Delta z_x, \Delta z_1, \Delta z_2, \Delta w_1, \Delta w_2) \) are given by \( F'(v_p) \Delta v_p = -F(v_p) \), with

\[
F'(v_p) = \begin{pmatrix}
H & 0 & -J^T & -A^T & 0 & -E_X^T & -E_{11}^T & 0 & 0 \\
0 & 0 & I_m & 0 & -L_{v_x}^T & 0 & 0 & 0 & 0 \\
J & -I_m & D_v & 0 & 0 & 0 & 0 & 0 & 0 \\
A & 0 & 0 & D_{A} & 0 & 0 & 0 & 0 & 0 \\
0 & L_{x} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
E_{x} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
Z_{1} E_{L} & 0 & 0 & 0 & 0 & X_{1}^T & 0 & 0 & 0 \\
-Z_{2} E_{U} & 0 & 0 & 0 & 0 & 0 & X_{2}^T & 0 & 0 \\
0 & W_{1} L_{L} & 0 & 0 & 0 & 0 & 0 & S_{1}^\mu & 0 \\
0 & -W_{2} L_{U} & 0 & 0 & 0 & 0 & 0 & 0 & S_{2}^\mu
\end{pmatrix}
\] (3.3)
3. The path-following equations

(recall that $z = E_X^T z_X + E_L^T z_1 - E_U^T z_2$ and $w = L_X^T w_X + L_L^T w_1 - L_U^T w_2$). Any $s$ may be written as $s = L_P^T s_F + L_X^T s_X$, where $s_F$ and $s_X$ denote the components of $s$ corresponding to the “free” and “fixed” components of $s$, respectively. Similarly, any $x$ may be written as $x = E_F^T x_F + E_X^T x_X$, where $x_F$ and $x_X$ denote the free and fixed components of $x$.

The partition of $x$ into free and fixed variables induces a partition of $H$, $A$, $J$, $E_L$, and $E_U$. We use $H_F$ to denote the $n_F \times n_F$ symmetric matrix of rows and columns of $H$ associated with the free variables and $A_F$, $A_X$, $J_F$, $J_X$ to denote the free and fixed columns of $A$ and $J$. In particular,

$$H_F = E_F^T H E_F^T, \quad A_F = A E_F^T, \quad A_X = A E_X^T, \quad J_F = J E_F^T, \quad \text{and} \quad J_X = J E_X^T.$$ 

Similarly, the $n_L \times n_F$ matrix $E_L F$ and $n_U \times n_F$ matrix $E_U F$ comprise the free columns of $E_L$ and $E_U$, with

$$E_L F = E_L^T E_F^T, \quad \text{and} \quad E_U F = E_U^T E_F^T.$$ 

It follows that the components of $E_L F x_F$ are the values of the free variables that are subject to lower bounds. A similar interpretation applied for $E_U F x_F$. Analogous definitions apply for the $m_L \times m_F$ matrix $L_L F$ and $m_U \times m_F$ matrix $L_U F$.

The next step is to transform the path-following equations to reflect the structure of free and fixed variables. Consider the block-diagonal orthogonal matrix $Q = \text{diag}(P_X, P_S, I_m, I_1, I_2, I_3, I_4, I_5, I_6, I_7, I_8, I_9, I_{10}, I_{11}, I_{12})$, where $P_X$ and $P_S$ are defined in (1.1). Given the identities

$$\begin{pmatrix} \Delta x_F \\ \Delta x_X \end{pmatrix} = P_X \Delta x = \begin{pmatrix} E_F \Delta x \\ E_X \Delta x \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \Delta s_F \\ \Delta s_X \end{pmatrix} = P_S \Delta s = \begin{pmatrix} L_F \Delta s \\ L_X \Delta s \end{pmatrix},$$
and $QF'(v_p)Q^TQ \Delta v_p = QF(v_p)$, we obtain the transformed equations

\[
\begin{bmatrix}
H_F & H_X^T & 0 & 0 & -J_F & -A_F & 0 & 0 & -E_{LF}^T & E_{UF}^T & 0 & 0 \\
H_O & H_X & 0 & 0 & -J_X & -A_X & 0 & -I_X^e & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & L_F & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & L_X & 0 & -I_X^e & 0 & 0 & 0 & 0 & 0 \\
J_F & J_X & -L_F^T & -L_X^T & D_y & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
A_F & A_X & 0 & 0 & 0 & D_A & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & I_X^e & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
Z_1E_{LF} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-Z_2E_{UF} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & W_1L_{LF} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -W_2L_{UF} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\begin{bmatrix}
\Delta x_F \\
\Delta x_X \\
\Delta s_F \\
\Delta s_X \\
\Delta y \\
\Delta v \\
\Delta w_x \\
\Delta w_y \\
\Delta z_1 \\
\Delta z_2 \\
\Delta w_1 \\
\Delta w_2 \\
\end{bmatrix}
= \begin{bmatrix}
g_F - J_F^Ty - A_F^Tv - z_F \\
g_X - J_X^Ty - A_X^Tv - z_X \\
y_F - w_F \\
y_X - w_X \\
c(x) - s + \mu^p(y - y^E) \\
A_x - b + \mu^N(v - v^E) \\
E_xx - b_x \\
L_Xs - h_X \\
z_1 \cdot (E_x - \ell^x) + \mu^p(z_1 - z_1^e) \\
z_2 \cdot (w_x - E_U) + \mu^p(z_2 - z_2^e) \\
w_1 \cdot (L_s - \ell^x) + \mu^p(w_1 - w_1^e) \\
w_2 \cdot (w^x - L_u) + \mu^p(w_2 - w_2^e) \\
\end{bmatrix},
\]

where $g_F = E_Fg$, $z_F = E_Fz$ and $y_F = L_Fy$.

As the constraints $L_Xs - h_X = 0$ and $E_xx - b_x = 0$ are enforced throughout, it follows that $\Delta s_X = 0$ and $\Delta x_X = 0$, in which case $\Delta s$ and $\Delta x$ satisfy

$$\Delta s = L_F^T \Delta s_F + L_X^T \Delta s_X = L_F^T \Delta s_F \quad \text{and} \quad \Delta x = E_F^T \Delta x_F + E_X^T \Delta x_X = E_F^T \Delta x_F.$$ 

After scaling the last four blocks of equations by (respectively) $Z_1^{-1}$, $Z_2^{-1}$, $W_1^{-1}$ and $W_2^{-1}$, collecting terms and reordering the
4. A shifted primal-dual penalty-barrier function

Problem (NLP) is equivalent to

\[
\begin{align*}
\min_{x, x_1, x_2, s_1, s_2} & f(x) \\
\text{subject to} & c(x) - s = 0, \quad Ax - b = 0, \\
& E_L x - x_1 = \ell^x, \quad L_L s - s_1 = \ell^s, \quad x_1 \geq 0, \quad s_1 \geq 0, \\
& E_U x + x_2 = u^x, \quad L_U s + s_2 = u^s, \quad x_2 \geq 0, \quad s_2 \geq 0, \\
& E_X x - b_X = 0, \quad L_X s - h_X = 0.
\end{align*}
\]

Consider the shifted primal-dual penalty-barrier problem

\[
\begin{align*}
\min_{x, x_1, x_2, s, s_1, s_2, y, v, w_1, w_2; \mu, \mu^b, y^E, v^R, w_1^E, w_2^E} & M(x, x_1, x_2, s, s_1, s_2, y, v, w_1, w_2; \mu, \mu^b, y^E, v^R, w_1^E, w_2^E) \\
\text{subject to} & E_L x - x_1 = \ell^x, \quad L_L s - s_1 = \ell^s, \quad x_1 + \mu^b e > 0, \quad z_1 > 0, \quad s_1 + \mu^b e > 0, \quad w_1 > 0, \\
& E_U x + x_2 = u^x, \quad L_U s + s_2 = u^s, \quad x_2 + \mu^b e > 0, \quad z_2 > 0, \quad s_2 + \mu^b e > 0, \quad w_2 > 0, \\
& E_X x - b_X = 0, \quad L_X s - h_X = 0,
\end{align*}
\]
4. A shifted primal-dual penalty-barrier function

where \( M(x, x_1, x_2, s, s_1, y, v, z_1, z_2, w_1, w_2; \mu^p, \mu^s, y^p, v^p, z_1^p, z_2^p, w_1^p, w_2^p) \) is the shifted primal-dual penalty-barrier function

\[
\begin{align*}
    f(x) - (c(x) - s)^T y^p + \frac{1}{2\mu^p}||c(x) - s||^2 + \frac{1}{2\mu^p}||c(x) - s + \mu^p(y - y^p)||^2 \\
    - (Ax - b)^T v^p + \frac{1}{2\mu^s}||Ax - b||^2 + \frac{1}{2\mu^s}||Ax - b + \mu^s(v - v^p)||^2 \\
    - \sum_{j=1}^{n_r} \{ \mu^p [z_1^p]^j \ln \left( [z_1^p]^j [x_1 + \mu^p e]^j \right) - [z_1^p \cdot (x_1 + \mu^p e)]^j \} \\
    - \sum_{j=1}^{n_r} \{ \mu^p [z_2^p]^j \ln \left( [z_2^p]^j [x_2 + \mu^p e]^j \right) - [z_2^p \cdot (x_2 + \mu^p e)]^j \} \\
    - \sum_{i=1}^{m_l} \{ \mu^s [w_1^e]^i \ln \left( [w_1^e]^i [s_1 + \mu^e e]^i \right) - [w_1^e \cdot (s_1 + \mu^e e)]^i \} \\
    - \sum_{i=1}^{m_l} \{ \mu^s [w_2^e]^i \ln \left( [w_2^e]^i [s_2 + \mu^e e]^i \right) - [w_2^e \cdot (s_2 + \mu^e e)]^i \}. \\
\end{align*}
\] (4.2)

The gradient \( \nabla M(x, x_1, x_2, s, s_1, y, v, z_1, z_2, w_1, w_2) \) may be defined in terms of the quantities \( X_1^\mu = \text{diag}(E_{c}x - \ell^x + \mu^p e), \)
\( X_2^\mu = \text{diag}(u^x - E_{v}x + \mu^p e), \)
\( Z_1 = \text{diag}(z_1), \)
\( Z_2 = \text{diag}(z_2), \)
\( W_1 = \text{diag}(w_1), \)
\( W_2 = \text{diag}(w_2), \)
\( S_1^\mu = \text{diag}(L_{e}s - \ell^s + \mu^e e) \) and
where the quantities $D^v, \pi^v, D_A, \pi^v, D^v_1, D^v_2, \pi^w_1, \pi^w_2, D^w_1, D^w_2, \pi^w_1$, and $\pi^w_2$ are defined in (3.5).
5. Derivation of the primal-dual line-search direction

The Hessian $\nabla^2 M(x, x_1, x_2, s, s_1, s_2, y, v, z_1, z_2, w_1, w_2)$ is given by

$$
H_1 = \begin{pmatrix}
0 & 0 & 0 & -2J^T D_v^{-1} & 0 & 0 & J^T & A^T & 0 & 0 & 0 & 0 \\
0 & 2\Pi_1^*(x_1)^{-1} & 0 & 0 & 0 & -I_m & 0 & I_c^2 & 0 & 0 & 0 & 0 \\
0 & 0 & 2\Pi_2^*(x_2)^{-1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-2D_v^{-1}J & 0 & 0 & 2\Pi_1^*(S_1^w)^{-1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 2\Pi_2^*(S_2^w)^{-1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
J & 0 & 0 & -I_m & 0 & 0 & D_v & 0 & 0 & 0 & 0 & 0 \\
A & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & I_e & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & I_e & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & I_e & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & I_e & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & I_e & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & I_e & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & I_e & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & I_e & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & I_e & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & I_e & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & I_e \\
\end{pmatrix}
$$

where

$$H_1 = H(x, 2\pi^v - y) + \frac{2}{\mu^v} A^T A + \frac{2}{\mu^v} J^T J = H(x, 2\pi^v - y) + 2A^T D_a^{-1} A + 2J^T D_v^{-1} J,$$

and $I_e$, $I_e$, $I_e$, and $I_e$ denote identity matrices of dimension $n_1$, $n_2$, $m_2$, and $m_2$ respectively. The usual convention regarding diagonal matrices formed from vectors applies, with $\Pi_1^2 = \text{diag}(\pi_1^2)$, $\Pi_2^2 = \text{diag}(\pi_2^2)$, $\Pi_1^w = \text{diag}(\pi_1^w)$, and $\Pi_2^w = \text{diag}(\pi_2^w)$.

5. Derivation of the primal-dual line-search direction

The primal-dual penalty-barrier problem (4.1) may be written in the form

$$\min_{p \in \mathcal{I}} M(p) \quad \text{subject to} \quad Cp = b_c,$$

where

$$\mathcal{I} = \{ p : p = (x, x_1, x_2, s, s_1, s_2, y, v, z_1, z_2, w_1, w_2), \quad \text{with} \quad x_i + \mu^v e > 0, \quad s_i + \mu^v e > 0, \quad z_i > 0, \quad w_i > 0 \text{ for } i = 1, 2 \},$$

and

$$C = \begin{pmatrix}
E_x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
E_0 & -I_e & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
E_v & 0 & I_e & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & L_x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & L_0 & -I_e & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & L_v & 0 & I_e & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}, \quad \text{with} \quad b_c = \begin{pmatrix}
b_x \\
b_0 \\
b_v \\
b_e \\
b_e \\
b_e \\
\end{pmatrix}.$$

(5.1)
Let $p \in I$ be given such that $Cp = b_c$. The Newton direction $\Delta p$ is given by the solution of the subproblem

$$\minimize_{\Delta p} \nabla M(p)^T \Delta p + \frac{1}{2} \Delta p^T \nabla^2 M(p) \Delta p \quad \text{subject to} \quad C \Delta p = b_c - Cp = 0.$$  \hfill (5.2)

However, instead of solving (5.2), we formulate a linearly constrained approximate Newton method by approximating the Hessian $\nabla^2 M$ by a positive-definite matrix $B$. Consider the matrix defined by replacing $\pi^Y$ by $y$, $\pi^Z_1$ by $z_1$, $\pi^Z_2$ by $z_2$, $\pi^W_1$ by $w_1$, $\pi^W_2$ by $w_2$ in the matrix $\nabla^2 M(x, x_1, x_2, s, s_1, s_2, y, v, z_1, z_2, w_1, w_2)$. This gives an approximate Hessian $B(x, x_1, x_2, s, s_1, s_2, y, v, z_1, z_2, w_1, w_2)$ of the form

$$
\begin{pmatrix}
H^0 + 2A^T D^{-1} A + 2J^T D^{-1} J & 0 & 0 & -2J^T D^{-1} J & 0 & 0 & J^T & A^T & 0 & 0 & 0 & 0 \\
0 & (D^Y_1)^{-1} & 0 & 0 & 0 & 0 & 0 & 0 & I_L^x & 0 & 0 & 0 \\
0 & 0 & (D^Z_2)^{-1} & 0 & 0 & 0 & 0 & 0 & I_U^x & 0 & 0 & 0 \\
-2D^Y_1 J & 0 & 0 & 2D^Z_1 & 0 & 0 & -I_m & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 2(D^Y_1)^{-1} & 0 & 0 & 0 & 0 & 0 & I_s^x & 0 \\
0 & 0 & 0 & 0 & 0 & 2(D^Z_1)^{-1} & 0 & 0 & 0 & 0 & I_s^x & 0 \\
J & 0 & 0 & 0 & -I_m & 0 & 0 & D_y & 0 & 0 & 0 & 0 \\
A & 0 & 0 & 0 & 0 & 0 & 0 & 0 & D_A & 0 & 0 & 0 \\
0 & I_L^x & 0 & 0 & 0 & 0 & 0 & 0 & D^x_1 & 0 & 0 & 0 \\
0 & 0 & I_U^x & 0 & 0 & 0 & 0 & 0 & D^x_1 & 0 & 0 & 0 \\
0 & 0 & 0 & I_L^z & 0 & 0 & 0 & 0 & D^z_1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & I_U^z & 0 & 0 & 0 & 0 & D^z_1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & I_L^s & 0 & 0 & 0 & 0 & D^s_1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & I_U^s & 0 & 0 & 0 & 0 & D^s_1 \\
\end{pmatrix},
$$

where $H^0 \approx H(x, y)$ is chosen so that the approximate Hessian $B$ is positive definite. Given $B(p)$, with $p = (x, x_1, x_2, s, s_1, s_2, y, v, z_1, z_2, w_1, w_2)$, an approximate Newton direction is given by the solution of the QP subproblem

$$\minimize_{\Delta p} \nabla M(p)^T \Delta p + \frac{1}{2} \Delta p^T B(p) \Delta p \quad \text{subject to} \quad C \Delta p = 0.$$

Let $N$ denote a matrix whose columns form a basis for null($C$), i.e., the columns of $N$ are linearly independent and $CN = 0$. Every feasible direction $\Delta p$ may be written in the form $\Delta p = Nd$. This implies that $d$ satisfies the reduced equations
\( N^T B(p) N d = -N^T \nabla M(p) \). Consider the null-space basis defined from the columns of the matrix

\[
N = \begin{pmatrix}
E_p^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
E_{uF}^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-E_{uF}^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & L_p^T & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & L_{LF} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -L_{uF} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & I_m & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & I_A & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & I_L^x & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & I_L^y & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & I_L^z & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & I_L^v
\end{pmatrix},
\tag{5.3}
\]

where \( E_{uF} = E_u E_p^T \), \( E_{uF} = E_u E_p^T \), \( L_{LF} = L_u L_p^T \) and \( L_{uF} = L_u L_p^T \). The definition of \( N \) of (5.3) gives the reduced approximate Hessian \( N^T B(p) N \) such that

\[
\begin{pmatrix}
\hat{H}_p & -2J_p^T D_y^{-1} L_p^T & J_p^T & A_p^T & E_p^T L_{uF} & -E_p^T L_{uF} & 0 & 0 \\
-2L_p D_y^{-1} J_p & 2L_p (D_y^{-1} + D_w^T) L_p^T & -L_p & 0 & 0 & 0 & L_p^T L_{uF} & L_{uF} \\
J_p^T & -L_p & 0 & 0 & 0 & 0 & 0 & 0 \\
A_p & 0 & 0 & D_A & 0 & 0 & 0 & 0 \\
E_{LF} & 0 & 0 & 0 & D_{1L}^x & 0 & 0 & 0 \\
-E_{uF} & 0 & 0 & 0 & 0 & D_{2u}^x & 0 & 0 \\
0 & L_{LF} & 0 & 0 & 0 & 0 & D_{1u}^x & 0 \\
0 & -L_{uF} & 0 & 0 & 0 & 0 & 0 & D_{2u}^x
\end{pmatrix},
\]

where \( J_p = J E_p^T \), \( A_p = A E_p^T \), \( \hat{H}_p = E_p H^g E_p^T + 2A_p^T D_A^{-1} A_p + 2J_p D_y^{-1} J_p + 2E_{uF} D_{1u}^x E_p^T \), with

\[
D_{1L}^x = E_p^T (D_{1L}^y)^{-1} E_p + E_p^T (D_{1u}^y)^{-1} E_p, \quad \text{and} \quad D_{1u}^x = L_{uF}^T (D_{1u}^y)^{-1} L_u + L_{uF}^T (D_{2u}^y)^{-1} L_u,
\]
Similarly, the reduced gradient $N^T \nabla M(p)$ is given by

\[
\begin{pmatrix}
g_p - A_p^T (2\pi^\nu - v) - J_p^T (2\pi^\nu - y) - E_{L}^T (2\pi_1^\nu - z_1) + E_{L}^T (2\pi_2^\nu - z_2) \\
2\pi_p^\nu - y_p - L_{L}^T (2\pi_1^\nu - w_1) + L_{L}^T (2\pi_2^\nu - w_2) \\
- D_y (\pi^\nu - y) \\
- D_A (\pi^\nu - v) \\
- D_1^w (\pi_1^\nu - z_1) \\
- D_2^w (\pi_2^\nu - z_2) \\
- D_1^u (\pi_1^\nu - w_1) \\
- D_2^u (\pi_2^\nu - w_2)
\end{pmatrix}
\]

where $g_p = E_p g$, $\pi_p^\nu = L_p \pi^\nu$ and $y_p = L_p y$. The reduced approximate Newton equations $N^T B(p) N d = -N^T \nabla M(p)$ are then

\[
\begin{pmatrix}
\hat{H}_F \\
-2L_{L} D_{F}^{-1} J_F \\
J_F \\
A_F \\
E_{L_F} \\
- E_{L_F} \\
0 \\
0
\end{pmatrix}
\begin{pmatrix}
-2J_F^T D_{F}^{-1} L_F^T \\
J_F^T \\
A_F^T \\
E_{L_F}^T \\
- E_{L_F}^T \\
0 \\
0 \\
- L_{L_F}
\end{pmatrix}
\begin{pmatrix}
A_F^T \\
E_{L_F}^T \\
- E_{L_F}^T \\
0 \\
0 \\
L_{L_F} \\
- L_{L_F}
\end{pmatrix}
\begin{pmatrix}
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0
\end{pmatrix}
\begin{pmatrix}
d_1 \\
d_2 \\
d_3 \\
d_4 \\
d_5 \\
d_6 \\
d_7 \\
d_8
\end{pmatrix}
\]

where $g_p - A_p^T (2\pi^\nu - v) - J_p^T (2\pi^\nu - y) - E_{L}^T (2\pi_1^\nu - z_1) + E_{L}^T (2\pi_2^\nu - z_2) \\
2\pi_p^\nu - y_p - L_{L}^T (2\pi_1^\nu - w_1) + L_{L}^T (2\pi_2^\nu - w_2) \\
- D_y (\pi^\nu - y) \\
- D_A (\pi^\nu - v) \\
- D_1^w (\pi_1^\nu - z_1) \\
- D_2^w (\pi_2^\nu - z_2) \\
- D_1^u (\pi_1^\nu - w_1) \\
- D_2^u (\pi_2^\nu - w_2)
\]

\[(5.4)\]
Given any nonsingular matrix $R$, the direction $d$ satisfies $RN^TB(p)Nd = -RN^T\nabla M(p)$. In particular, if $R$ is the block upper-triangular matrix $R$ such that

$$
R = \begin{pmatrix}
I_p & 0 & -2J_p^T D_p^{-1} & -2A_p^T D_A^{-1} & -2E_p^T(D_2^p)^{-1} & 0 & 0 \\
I_p & 0 & 2L_p D_p^{-1} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
I_A & 0 & 0 & 0 & 0 & 0 & 0 \\
I_p^x & I_p^x & 0 & 0 & 0 & 0 & 0 \\
I_L^x & I_L^x & 0 & 0 & 0 & 0 & 0 \\
I_{UF} & I_{UF} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{pmatrix},
$$

then

$$
RN^TB(p)N = \begin{pmatrix}
H_p & 0 & -J_p^T & -A_p^T & -E_p^T & D_p & 0 & 0 \\
0 & 0 & L_p & 0 & 0 & 0 & -L_p^T & L_p^T \\
J_p & -L_p^T & D_p & 0 & 0 & 0 & 0 & 0 \\
A_p & 0 & 0 & D_A & 0 & 0 & 0 & 0 \\
E_p & 0 & 0 & 0 & D_1^p & 0 & 0 & 0 \\
-E_{UF} & 0 & 0 & 0 & 0 & D_2^p & 0 & 0 \\
0 & L_p & 0 & 0 & 0 & 0 & D_1^w & 0 \\
0 & -L_{UF} & 0 & 0 & 0 & 0 & 0 & D_2^w \\
\end{pmatrix},
$$

and

$$
RN^T\nabla M(p) = \begin{pmatrix}
g_p - J_p^T y - A_p^T v - z_p \\
y_p - w_p \\
-D_v(\pi^v - y) \\
-D_A(\pi^v - v) \\
-D_1^p(\pi_1^p - z_1) \\
-D_2^p(\pi_2^p - z_2) \\
-D_1^w(\pi_1^w - w_1) \\
-D_2^w(\pi_2^w - w_2) \\
\end{pmatrix}.
$$
The matrix $R$ is nonsingular because $Z_1$, $Z_2$ and $W$ are positive definite, which implies that we may solve the following (unsymmetric) reduced approximate Newton equations for $d$:

\[
\begin{pmatrix}
H^p & 0 & -J^T_F & -A^T_F & -E^T_{LF} & E^T_{UF} & 0 & 0 \\
0 & 0 & L^T_F & 0 & 0 & 0 & -L^T_{LF} & L^T_{UF} \\
J_F & -L^T_F & D_v & 0 & 0 & 0 & 0 & 0 \\
A_F & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
E_{LF} & 0 & 0 & 0 & 0 & D_z^2 & 0 & 0 \\
-E_{UF} & 0 & 0 & 0 & 0 & D_z^2 & 0 & 0 \\
0 & L_{LF} & 0 & 0 & 0 & 0 & D_1^w & 0 \\
0 & -L_{UF} & 0 & 0 & 0 & 0 & 0 & D_2^w \\
\end{pmatrix}
\begin{pmatrix}
d_1 \\
d_2 \\
d_3 \\
d_4 \\
d_5 \\
d_6 \\
d_7 \\
d_8 \\
\end{pmatrix}
= \begin{pmatrix}
g_F - J_F^T y - A_F^T v - z_F \\
g_y - w_y \\
\end{pmatrix}.
\]

Then, the expression $\Delta p = N d$ implies that

\[
\Delta p = \begin{pmatrix}
\Delta x \\
\Delta x_1 \\
\Delta x_2 \\
\Delta s \\
\Delta s_1 \\
\Delta s_2 \\
\Delta y \\
\Delta v \\
\Delta z_1 \\
\Delta z_2 \\
\Delta w_1 \\
\Delta w_2 \\
\end{pmatrix} = N d = \begin{pmatrix}
E^T_{F} d_1 \\
\vdots \\
E^T_{F} d_8 \\
\end{pmatrix}.
\]

These identities allow us to write equations (5.5) in the form

\[
\begin{pmatrix}
H^p & 0 & -J^T_F & -A^T_F & -E^T_{LF} & E^T_{UF} & 0 & 0 \\
0 & 0 & L^T_F & 0 & 0 & 0 & -L^T_{LF} & L^T_{UF} \\
J_F & -L^T_F & D_v & 0 & 0 & 0 & 0 & 0 \\
A_F & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
E_{LF} & 0 & 0 & 0 & 0 & D_z^2 & 0 & 0 \\
-E_{UF} & 0 & 0 & 0 & 0 & D_z^2 & 0 & 0 \\
0 & L_{LF} & 0 & 0 & 0 & 0 & D_1^w & 0 \\
0 & -L_{UF} & 0 & 0 & 0 & 0 & 0 & D_2^w \\
\end{pmatrix}
\begin{pmatrix}
\Delta x_F \\
\Delta s_F \\
\Delta y \\
\Delta v \\
\Delta z_1 \\
\Delta z_2 \\
\Delta w_1 \\
\Delta w_2 \\
\end{pmatrix} = \begin{pmatrix}
g_F - J_F^T y - A_F^T v - z_F \\
g_y - w_y \\
\end{pmatrix}.
\]
with $\Delta x = E_T^T \Delta x_p$, $\Delta s = L_T^T \Delta s_p$, $\Delta x_1 = \Delta x_p - (\ell^x - E_L x + x_1)$, $\Delta x_2 = -\Delta x_p + (u^x - E_U x - x_2)$, $\Delta s_1 = \Delta s_p - (\ell^s - L_L s + s_1)$ and $\Delta s_2 = -\Delta s_p + (u^s - L_U s - s_2)$.

The shifted penalty-barrier equations (5.7) are the same as the path following equations (3.4) except for the $(1,1)$ block, where $H(x,y)$ replaced by $H^u(x,y)$.

### 6. The shifted primal-dual penalty-barrier direction

In this section we consider the solution of the shifted primal-dual penalty-barrier equations (5.7). Collecting terms and reordering the equations and unknowns, gives

$$
\begin{pmatrix}
D_A & 0 & 0 & 0 & 0 & 0 & A_v & 0 \\
0 & D_1^w & 0 & 0 & 0 & 0 & E_{1v} & 0 \\
0 & 0 & D_2^w & 0 & 0 & 0 & -E_{2v} & 0 \\
0 & 0 & 0 & D_1^w & 0 & L_L & 0 & 0 \\
0 & 0 & 0 & 0 & D_2^w & -L_{uf} & 0 & 0 \\
0 & 0 & 0 & 0 & -L_{uf} & L_{uf} & L_L & 0 \\
-A_T & -E_{1p} & E_{1p} & 0 & 0 & 0 & H_p^T & -J_p^T \\
0 & 0 & 0 & 0 & 0 & -L_{fp}^T & J_f & D_v \\
\end{pmatrix}
\begin{pmatrix}
\Delta v \\
\Delta z_1 \\
\Delta z_2 \\
\Delta w_1 \\
\Delta w_2 \\
\Delta s_p \\
\Delta x_p \\
\Delta y \\
\end{pmatrix}
= -
\begin{pmatrix}
D_A(v - \pi^v) \\
D_1^w(z_1 - \pi^1_1) \\
D_2^w(z_2 - \pi^2_2) \\
D_1^w(w_1 - \pi^1_1) \\
D_2^w(w_2 - \pi^2_1) \\
L_p(y - w) \\
E_p(g - J^T y - A^T v - z) \\
D_v(y - \pi^v) \\
\end{pmatrix}.
$$

Consider the diagonal matrices

$$
D_w = (L_T^T (D_1^w)^{-1} L_L + L_T^T (D_2^w)^{-1} L_L)^\dagger \quad \text{and} \quad D_z = (E_T^T (D_1^w)^{-1} E_L + E_T^T (D_2^w)^{-1} E_L)^\dagger,
$$

where $(\cdot)^\dagger$ denotes the Moore-Penrose pseudoinverse of a matrix. The identity $I_m = L_L^T L_L + L_F^T L_F$ implies that the $m \times m$ matrix $D_w$ satisfies the identities

$$
L_F^T L_F D_w = D_w = D_w L_F^T L_F, \quad \text{and} \quad L_L^T L_L D_w = 0.
$$
In addition, the diagonal matrix $L_p D_A^T L_p^T$ is nonsingular if every slack is either fixed or bounded above or below. If equations (6.1) are premultiplied by the matrix

$$
\begin{pmatrix}
I_A & I_{Lp}^T & I_{Lp}^T & I_{Lp}^T & I_{Lp}^T & I_{Lp}^T \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
A^T_p D_A^{-1} & E^T_{Lp}(D_1^e)^{-1} & -E^T_{Lp}(D_2^e)^{-1} & L^T_{Lp}(D_1^w)^{-1} & -L^T_{Lp}(D_2^w)^{-1} & I_{Lp}^T \\
0 & 0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}
$$

we obtain the block upper-triangular system

$$
\begin{pmatrix}
D_A & 0 & 0 & 0 & 0 & 0 & A_p & 0 \\
0 & D_1^e & 0 & 0 & 0 & 0 & E_{Lp} & 0 \\
0 & 0 & D_2^e & 0 & 0 & 0 & -E_{Lp} & 0 \\
0 & 0 & 0 & D_1^w & 0 & L_{Lp} & 0 & 0 \\
0 & 0 & 0 & 0 & D_2^w & -L_{Lp} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & L_{Lp} D^T_A L_p^T & 0 & L_p \\
0 & 0 & 0 & 0 & 0 & 0 & \tilde{H}_p & -J_p^T \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & J_p & D_v + D_w \\
\end{pmatrix}
\begin{pmatrix}
\Delta v \\
\Delta z_1 \\
\Delta z_2 \\
\Delta w_1 \\
\Delta w_2 \\
\Delta s_p \\
\Delta x_p \\
\Delta y \\
\end{pmatrix} = -
\begin{pmatrix}
D_A (v - \pi^v) \\
D_1^e (z_1 - \pi_1^f) \\
D_2^e (z_2 - \pi_2^f) \\
D_1^w (w_1 - \pi_1^w) \\
D_2^w (w_2 - \pi_2^w) \\
L_p (y - \pi^w) \\
E_p (-J^T y - A^T \pi^v - \pi^z) \\
D_w (y - \pi^w) + D_v (y - \pi^v) \\
\end{pmatrix},
$$

where $\tilde{H}_p = H_p^p + A^T_p D_A^{-1} A_p + E_p D^T_A E_p^T$, $\pi^w = L^T_p \pi^w_1 - L^T_u \pi^w_2$ and $\pi^z = E^T_L \pi^z_1 - E^T_u \pi^z_2$. Using block back-substitution, $\Delta x_p$ and $\Delta y$ can be computed by solving the equations

$$
\begin{pmatrix}
\tilde{H}_p & -J_p^T \\
J_p & D_v + D_w \\
\end{pmatrix}
\begin{pmatrix}
\Delta x_p \\
\Delta y \\
\end{pmatrix} = -
\begin{pmatrix}
E_p (-J^T y - A^T \pi^v - \pi^z) \\
D_w (y - \pi^w) + D_v (y - \pi^v) \\
\end{pmatrix}.
$$

The full vector $\Delta x$ is then computed as $\Delta x = E_p^T \Delta x_p$. Similarly, substitution of the identity $\Delta s = L_p^T \Delta s_p$ in the sixth block of equations gives

$$
\Delta s = -D_w (y + \Delta y - \pi^w).
$$

There are several ways of computing $\Delta w_1$ and $\Delta w_2$. Instead of using the block upper-triangular system above, we use the last two blocks of equations of (3.4) to give

$$
\Delta w_1 = -(S_1^u)^{-1} \left( w_1 - (L_4 (s + \Delta s) - \ell^e + \mu^\ell e) - \mu^w w_1^e \right) \quad \text{and} \quad \Delta w_2 = -(S_2^u)^{-1} \left( w_2 - (L_4 (s + \Delta s) + \mu^\ell e) - \mu^w w_2^e \right).
$$
7. Summary: equations for the line-search direction

The results of the preceding section imply that the solution of the path-following equations \( F'(v_p)\Delta v_p = -F(v_p) \) with \( F \) and \( F' \) given by (3.2) and (3.3) may be computed as follows. Let \( x \) and \( s \) be given primal variables and slack variables such that \( E_x x = b_x, L_x s = h_x \) with \( \ell^s - \mu^u < E_c x, E_v x < u^s + \mu^u, \ell^p - \mu^u < L_i s, L_o s < u^s + \mu^u \). Similarly, let \( z_1, z_2, w_1, w_2 \) and \( y \) denote dual variables such that \( w_1 > 0, w_2 > 0, z_1 > 0, \) and \( z_2 > 0 \). Consider the diagonal matrices \( X_1^u = \text{diag}(E_x x - \ell^x + \mu^u e), X_2^u = \text{diag}(u^x - E_v x + \mu^u e), Z_1 = \text{diag}(z_1), Z_2 = \text{diag}(z_2), W_1 = \text{diag}(w_1), W_2 = \text{diag}(w_2), S_1^u = \text{diag}(L_i s - \ell^s + \mu^u e) \) and \( S_2^u = \text{diag}(u^s - L_o s + \mu^u e) \). Consider the quantities

\[
D_v = \mu^v I_m, \quad \pi^v = y^v - \frac{1}{\mu^v} (c - s),
\]
\[
D_A = \mu^A I_A, \quad \pi^A = v^A - \frac{1}{\mu^A} (Ax - b),
\]
\[
(D_1^x)^{-1} = (X_1^u)^{-1} Z_1, \quad (D_1^x)^{-1} = (X_2^u)^{-1} Z_2,
\]
\[
(D_2^x)^{-1} = (X_1^u)^{-1} Z_1, \quad (D_2^x)^{-1} = (X_2^u)^{-1} Z_2,
\]
\[
(D_z^x)^{-1} = (E_x^T X_1^u E_c + E_v^T X_2^u E_v)^{-1} E_u, \quad (D_z^x)^{-1} = (E_x^T X_1^u E_c + E_v^T X_2^u E_v)^{-1} E_u,
\]
\[
\pi_1^x = \mu^x (X_1^u)^{-1} z_1^x, \quad \pi_1^x = \mu^x (X_1^u)^{-1} z_1^x,
\]
\[
\pi_2^x = \mu^x (X_2^u)^{-1} z_2^x, \quad \pi_2^x = \mu^x (X_2^u)^{-1} z_2^x,
\]
\[
\pi^x = E_x^T \pi_1^x - E_v^T \pi_2^x,
\]
\[
\pi^v = y^v - \frac{1}{\mu^v} (c - s), \quad \pi^A = v^A - \frac{1}{\mu^A} (Ax - b),
\]
\[
\pi^v = y^v - \frac{1}{\mu^v} (c - s), \quad \pi^A = v^A - \frac{1}{\mu^A} (Ax - b),
\]
\[
(D_v)^{-1} = (S_1^u)^{-1} W_1, \quad (D_A)^{-1} = (S_2^u)^{-1} W_2,
\]
\[
(D_w)^{-1} = (S_1^u)^{-1} W_1, \quad (D_w)^{-1} = (S_2^u)^{-1} W_2,
\]
\[
D_A = (L_i^T S_1^u L_i + L_o^T S_2^u L_o)^{-1} L_v,
\]
\[
\pi_1^w = \mu^w (S_1^u)^{-1} w_1^e, \quad \pi_2^w = \mu^w (S_2^u)^{-1} w_2^e,
\]
\[
\pi^w = L_i^T \pi_1^w - L_o^T \pi_2^w.
\]

Choose \( H^u(x, y) \) so that \( H^u(x, y) \) approximates \( H(x, y) \) and the KKT matrix

\[
\begin{pmatrix}
H^u_p(x, y) + A_p D_x^{-1} A_p + E_p D_x^T E_p & J_p(x)^T \\
J_p(x) & -D_v + D_w
\end{pmatrix}
\]

is nonsingular with \( m \) negative eigenvalues. Solve the KKT system

\[
\begin{pmatrix}
H^u_p(x, y) + A_p D_x^{-1} A_p + E_p D_x^T E_p & -J_p(x)^T \\
J_p(x) & D_v + D_w
\end{pmatrix}
\begin{pmatrix}
\Delta x_p \\
\Delta y
\end{pmatrix} = -\begin{pmatrix}
E_p (\nabla f(x) - J(x)^T y - A_p \pi^v - \pi^x) \\
J_p(y - \pi^v) + D_w (y - \pi^w)
\end{pmatrix},
\]

(7.1)
and set
\[ \Delta x = E_p^T \Delta x_p, \quad \tilde{x} = x + \Delta x, \]
\[ \Delta z_1 = - (X_1^\mu)^{-1} (z_1 \cdot (E \tilde{x} - \tilde{\ell}^\mu + \mu^p e) - \mu^p z_1^\mu), \]
\[ \Delta z_2 = - (X_2^\mu)^{-1} (z_2 \cdot (w^s - E \tilde{x} + \mu^p e) - \mu^p z_2^\mu), \]
\[ \tilde{y} = y + \Delta y, \quad \Delta s = - D_w (\tilde{y} - \pi^w), \]
\[ \tilde{s} = s + \Delta s, \quad \Delta w_1 = - (S_1^\mu)^{-1} (w_1 \cdot (L \tilde{s} - \tilde{\ell}^s + \mu^s e) - \mu^s w_1^\mu), \]
\[ \Delta w_2 = - (S_2^\mu)^{-1} (w_2 \cdot (w^s - L \tilde{s} + \mu^s e) - \mu^s w_2^\mu), \]
\[ \hat{\pi}^V = v^s - \frac{1}{\mu^s} (A \tilde{x} - b), \quad \Delta v = \hat{\pi}^V - v, \]
\[ w = L_x^T w_x + L_u^T w_1 - L_v^T w_2, \quad z = E_x^T z_x + E_v^T z_1 - E_u^T z_2, \]
\[ \Delta w_x = \Delta \hat{y} - \Delta w, \quad \Delta z_x = \nabla f(x) + H^p(x, y) \Delta x - J(x)^T \hat{y} - A^T \hat{v} - z \big|_x. \]

As \((x, s) \to (x^*, s^*)\) it holds that \(\|D^1_w\|\) and \(\|D^2_w\|\) are bounded, but \(\|D_w\| \to \infty\) and \(\|A_D^T D^-1 D_A\| \to \infty\). This implies that the matrix and right-hand side of (7.1) goes to infinity. In the situation where \(A_D^T D^-1 A_D\) is diagonal, then the KKT system can be rescaled so that the equations to be solved are bounded. If \(\tilde{D}_x\) and \(\tilde{D}_w\) denote diagonal matrices such that \(\tilde{D}^2_x = (A_D^T D^-1 A_D)^{-1}\) and \(\tilde{D}^2_w = (L_x^T L_x + D_w)^{-1}\), then \(\|\tilde{D}_x\|\) and \(\|\tilde{D}_w\|\) are bounded as \((x, s) \to (x^*, s^*)\). The equations (7.1) may be written in the form
\[
\begin{pmatrix}
\tilde{D}_x H^p(x, y) \tilde{D}_x + \tilde{D}^2_x E_p \tilde{D}_x^T E_p^T + I_p^x \\
\tilde{D}_w J_p(x) \tilde{D}_x \\
\tilde{D}^2_w D_Y + L_v^T L_p
\end{pmatrix}
\begin{pmatrix}
\Delta \tilde{x}_p \\
\Delta \tilde{y} \\
\Delta \tilde{\ell}^s 
\end{pmatrix} =
\begin{pmatrix}
\tilde{D}_x E_p (\nabla f(x) - J(x)^T y - A^T \pi^V - \pi^\ell) \\
\tilde{D}^2_w D_Y (y - \pi^w) + D_w (y - \pi^w)
\end{pmatrix},
\]
with \(\Delta x_p = \tilde{D}_x \Delta \tilde{x}_p\) and \(\Delta y = \tilde{D}_w \Delta \tilde{y}\). In this case, the scaled KKT matrix remains bounded if \(H(x, y)\) is bounded. Similarly, the right-hand side remains bounded if \(\|D_w D_Y (y - \pi^w)\|\) is bounded.
The associated line-search merit function (4.2) can be written as

$$f(x) - (c(x) - s)^T y^\varepsilon + \frac{1}{2\mu^v} ||c(x) - s||^2 + \frac{1}{2\mu^p} ||c(x) - s + \mu^v(y - y^\varepsilon)||^2$$

$$- (Ax - b)^T v^\varepsilon + \frac{1}{2\mu^v} ||Ax - b||^2 + \frac{1}{2\mu^p} ||Ax - b + \mu^v(v - v^\varepsilon)||^2$$

$$- \sum_{j=1}^{m_L} \left\{ \mu^v [z_f^\varepsilon]_j \ln \left( \left| z_1 \right|_j [E_L x - \ell^x + \mu^v e]_j \right) - \left| z_1 \right|_j \cdot (E_L x - \ell^x + \mu^v e)_j \right\}$$

$$- \sum_{j=1}^{m_L} \left\{ \mu^v [z_f^\varepsilon]_j \ln \left( \left| z_2 \right|_j [u^x - E_U x + \mu^v e]_j \right) - \left| z_2 \right|_j \cdot (u^x - E_U x + \mu^v e)_j \right\}$$

$$- \sum_{i=1}^{m_L} \left\{ \mu^v [w^\varepsilon]_i \ln \left( \left| w_1 \right|_i [L_s e - \ell^s + \mu^v e]_i \right) - \left| w_1 \right|_i \cdot (L_s e - \ell^s + \mu^v e)_i \right\}$$

$$- \sum_{i=1}^{m_L} \left\{ \mu^v [w^\varepsilon]_i \ln \left( \left| w_2 \right|_i [u^s - L_U s + \mu^v e]_i \right) - \left| w_2 \right|_i \cdot (u^s - L_U s + \mu^v e)_i \right\}.$$

8. The primal-dual trust-region direction

Given a vector of primal-dual variables $p = (x, x_1, x_2, s, s_1, s_2, y, v, z_1, z_2, w_1, w_2)$, each iteration of a trust-region method for solving (NLP) involves finding a vector $\Delta p$ of the form $\Delta p = N\delta$, where $N$ is a basis for the null-space of the matrix $C$ of (5.1), and $d$ is an approximate solution of the subproblem

$$\text{minimize} \quad g^T_N d + \frac{1}{2} d^T B_N(p)d \quad \text{subject to} \quad \|d\|_T \leq \delta, \quad (8.1)$$

where $g_N$ and $B_N$ are the reduced gradient and reduced Hessian $g_N = \nabla M$ and $B_N(p) = N^T B(p)N$, $\|d\|_T = (d^T T d)^{1/2}$, $\delta$ is the trust-region radius, and $T$ is positive-definite. The subproblem (8.1) may be written as

$$\text{minimize} \quad g^T_N T^{-1/2} \Delta v_M + \frac{1}{2} \Delta v_M^T T^{-1/2} B_N(p)T^{-1/2} \Delta v_M \quad \text{subject to} \quad \|\Delta v_M\|_2 \leq \delta, \quad (8.2)$$

where $\Delta v_M = T^{1/2}d$. The application of the method of Moré and Sorensen [8] to solve the subproblem (8.2) requires the solution of the so-called secular equations, which have the form

$$(B_N + \sigma T)\Delta v_M = -\tilde{g}_N, \quad (8.3)$$
with $\sigma$ a nonnegative scalar, $\tilde{B}_N = T^{-1/2}B_N(p)T^{-1/2}$, and $g_N = T^{-1/2}g_N$. In this note we consider the solution of the related equations

$$(B_N + \sigma T)d = -g_N,$$  
(8.4)
and recover the solution of the secular equations (8.3) from the computed vector $d$.

The identity (5.6) allows the solution of the approximate Newton equations $B_N(p)d = -g_N$ (5.4) to be written in terms of the change in the variables $(x, s, s_1, s_2, y, v, z_1, z_2, w_1, w_2)$. In particular, we have

$$\begin{pmatrix}
\tilde{H}_F & -2J_F^T D^{-1}_y J_F & J_F^T & A_F^T & E_{L_F}^T & -E_{U_F}^T & 0 & 0 \\
-2L_F D^{-1}_y J_F & 2L_F (D^{-1}_C + D^{-1}_A) L_F & -L_F & 0 & 0 & 0 & 0 & 0 \\
J_F & -L_F & D_y & 0 & 0 & 0 & 0 & 0 \\
A_F & 0 & 0 & D_A & 0 & 0 & 0 & 0 \\
E_{L_F} & 0 & 0 & 0 & D_1^y & 0 & 0 & 0 \\
-E_{U_F} & 0 & 0 & 0 & 0 & D_2^y & 0 & 0 \\
0 & L_{L_F} & 0 & 0 & 0 & 0 & D_1^w & 0 \\
0 & -L_{U_F} & 0 & 0 & 0 & 0 & 0 & D_2^w
\end{pmatrix}
\begin{pmatrix}
\Delta x_F \\
\Delta y \\
\Delta z_1 \\
\Delta z_2 \\
\Delta w_1 \\
\Delta w_2
\end{pmatrix}
= \begin{pmatrix}
g_F - A_F^T (2\pi^v - v) - J_F^T (2\pi^v - y) - E_{L_F}^T (2\pi^z - z_1) + E_{U_F}^T (2\pi^z - z_2) \\
2\pi^v - y_F - L_F (2\pi^v - w_1) + L_{U_F}^T (2\pi^z - w_2) \\
-D_y (\pi^v - y) \\
-D_A (\pi^v - v) \\
-D_1^y (\pi^z - z_1) \\
-D_2^y (\pi^z - z_2) \\
-D_1^w (\pi^w - w_1) \\
-D_2^w (\pi^w - w_2)
\end{pmatrix},$$

where

$$\tilde{H}_F = H_F + J_F^T D^{-1}_y J_F + A_F^T D_A^{-1} A_F + E_F D_E^T E_F^T,$$

with

$$D_y = \mu^y I_m, \quad \pi^v = y^e - \frac{1}{\mu^v} (c - s), \quad D_A = \mu^A I_s, \quad \pi^v = v^e - \frac{1}{\mu^A} (Ax - b),$$

$$D_1^w = S_1^w W_1^{-1}, \quad \pi_1^w = \mu^w (S_1^w)^{-1} w_1^e, \quad D_1^z = X_1^w Z_1^{-1}, \quad \pi_1^z = \mu^w (X_1^w)^{-1} z_1^e,$$

$$D_2^w = S_2^w W_2^{-1}, \quad \pi_2^w = \mu^w (S_2^w)^{-1} w_2^e, \quad D_2^z = X_2^w Z_2^{-1}, \quad \pi_2^z = \mu^w (X_2^w)^{-1} z_2^e,$$

$$\pi^w = L_{L_F}^T \pi_1^w - L_{U_F}^T \pi_2^w, \quad \pi^z = E_{L_F} \pi_1^z - E_{U_F} \pi_2^z.$$
In the trust-region case, we make no assumption that $B$ is positive definite, i.e., $H_F = E_F H(x, y) E_F^T$ with $H(x, y)$ the Hessian of the Lagrangian function.

The first step in the formulation of the trust-region equations (8.4) and their solution is to write the reduced gradient and Hessian of the merit function in terms of the vectors $\vec{x}$ and $\vec{y}$ that combine the primal variables $(x, s)$ and dual variables $(y, v, z_1, z_2, w_1, w_2)$. Let $\vec{g}$, $\vec{H}$, $\vec{J}$ and $\vec{D}$ denote the quantities

$$\vec{g} = \begin{pmatrix} g_F \\ 0 \end{pmatrix}, \quad \vec{H} = \begin{pmatrix} H_F & 0 \\ 0 & 0 \end{pmatrix}, \quad \vec{J} = \begin{pmatrix} J_F & -L_F^T \\ A_F & 0 \\ E_{LF} & 0 \\ -E_{UF} & 0 \\ 0 & L_{LF} \\ 0 & -L_{UF} \end{pmatrix}, \quad \text{and} \quad \vec{D} = \begin{pmatrix} D_Y & 0 & 0 & 0 & 0 & 0 \\ 0 & D_A & 0 & 0 & 0 & 0 \\ 0 & 0 & D_1^y & 0 & 0 & 0 \\ 0 & 0 & 0 & D_2^y & 0 & 0 \\ 0 & 0 & 0 & 0 & D_1^w & 0 \\ 0 & 0 & 0 & 0 & 0 & D_2^w \end{pmatrix}. $$

Similarly, let $\vec{T}_x = \text{diag}(T^x, T^s)$ and $\vec{T}_y = \text{diag}(T^y, T^v, T_1^z, T_2^z, T_1^w, T_2^w)$. The equations $(B_N + \sigma T) \Delta p = -g_N$ may be written in the form

$$\begin{pmatrix} \vec{H} + 2\vec{J}^T \vec{D}^{-1} \vec{J} + \sigma \vec{T}_x \\ \vec{J} \end{pmatrix} \begin{pmatrix} \Delta \vec{x} \\ \Delta \vec{y} \end{pmatrix} = -\begin{pmatrix} \vec{g} - \vec{J}^T \vec{y} \\ -\vec{D}(\vec{y} - \vec{y}) \end{pmatrix},$$

where

$$\vec{y} = \begin{pmatrix} y \\ v \\ z_1 \\ z_2 \\ w_1 \\ w_2 \end{pmatrix}, \quad \vec{\pi} = \begin{pmatrix} \pi^v \\ \pi^v \\ \pi^v \\ \pi^v \\ \pi^w \\ \pi^w \end{pmatrix}, \quad \Delta \vec{x} = \begin{pmatrix} \Delta x_F \\ \Delta s \end{pmatrix}, \quad \text{and} \quad \Delta \vec{y} = \begin{pmatrix} \Delta y \\ \Delta z \\ \Delta w \end{pmatrix}. $$

Applying the nonsingular matrix $\begin{pmatrix} I & -2\vec{J}^T \vec{D}^{-1} \end{pmatrix}$ to both sides of (8.5) gives the equivalent system

$$\begin{pmatrix} \vec{H} + \sigma \vec{T}_x \\ \vec{J} \end{pmatrix} \begin{pmatrix} \Delta \vec{x} \\ \Delta \vec{y} \end{pmatrix} = -\begin{pmatrix} \vec{g} - \vec{J}^T \vec{y} \\ -\vec{D}(\vec{y} - \vec{y}) \end{pmatrix}. $$

As in Gertz and Gill [3], we set $\vec{T}_x = I$ and $\vec{T}_y = \vec{D}$. With this choice, the associated vectors $\Delta \vec{x}$ and $\Delta \vec{y}$ satisfy the equations

$$\begin{pmatrix} \vec{H} + \sigma I \\ \vec{J} \end{pmatrix} \begin{pmatrix} \Delta \vec{x} \\ (1 + 2\sigma) \Delta \vec{y} \end{pmatrix} = -\begin{pmatrix} \vec{g} - \vec{J}^T \vec{y} \\ \vec{D}(\vec{y} - \vec{y}) \end{pmatrix}. $$
where $\bar{\sigma} = (1 + \sigma)/(1 + 2\sigma)$. In terms of the original variables, the unsymmetric equations (8.6) are

\[
\begin{pmatrix}
H_F + \sigma I_F^x & 0 & -J_F^T & -A_F^T & -E_{1,F}^T & E_{1,F}^T & 0 & 0 \\
0 & \sigma I_F^x & \sigma D_F & 0 & 0 & 0 & -L_{1,F}^T & L_{1,F}^T \\
J_F & -L_F^T & \sigma D_Y & 0 & 0 & 0 & 0 & 0 \\
A_F & 0 & 0 & \sigma D_y & 0 & 0 & 0 & 0 \\
E_{1,F} & 0 & 0 & 0 & \sigma D_y^z & 0 & 0 & 0 \\
-E_{1,F} & 0 & 0 & 0 & 0 & \sigma D_y^z & 0 & 0 \\
0 & L_{1,F} & 0 & 0 & 0 & 0 & \sigma D_y^z & 0 \\
0 & -L_{1,F} & 0 & 0 & 0 & 0 & 0 & \sigma D_y^z \\
\end{pmatrix}
\begin{pmatrix}
\Delta x_F \\
\Delta y_F \\
\Delta s_F \\
\end{pmatrix}
= \begin{pmatrix}
(1 + 2\sigma)\Delta y \\
(1 + 2\sigma)\Delta v \\
(1 + 2\sigma)\Delta z_1 \\
(1 + 2\sigma)\Delta z_2 \\
(1 + 2\sigma)\Delta w_1 \\
(1 + 2\sigma)\Delta w_2 \\
\end{pmatrix},
\tag{8.7}
\end{equation}

where $\bar{\sigma} = (1 + \sigma)/(1 + 2\sigma)$. Collecting terms and reordering the equations and unknowns, we obtain

\[
\begin{pmatrix}
\bar{\sigma} D_A & 0 & 0 & 0 & 0 & 0 & A_F & 0 \\
0 & \bar{\sigma} D_Y & 0 & 0 & 0 & 0 & E_{1,F} & 0 \\
0 & 0 & \bar{\sigma} D_y^z & 0 & 0 & 0 & -E_{1,F} & 0 \\
0 & 0 & 0 & \bar{\sigma} D_y^z & 0 & 0 & L_{1,F} & 0 \\
0 & 0 & 0 & 0 & \bar{\sigma} D_y^z & 0 & -L_{1,F} & L_{1,F} \\
0 & 0 & 0 & 0 & -L_{1,F} & \sigma I_F^x & 0 & L_F \\
0 & 0 & 0 & -L_{1,F} & L_{1,F}^T & \sigma I_F^x & 0 & L_F \\
-A_F^T & -E_{1,F}^T & E_{1,F}^T & 0 & 0 & 0 & H_F + \sigma I_F^x & -J_F^T \\
0 & 0 & 0 & 0 & 0 & 0 & L_F & \bar{\sigma} D_Y \\
\end{pmatrix}
\begin{pmatrix}
\Delta \bar{v} \\
\Delta z_1 \\
\Delta z_2 \\
\Delta w_1 \\
\Delta w_2 \\
\Delta s_F \\
\Delta x_F \\
\Delta y \\
\end{pmatrix}
= \begin{pmatrix}
D_A(v - \pi^v) \\
D_Y(z_1 - \pi^z_1) \\
D_y^z(z_2 - \pi^z_2) \\
D_y^z(w_1 - \pi^z_1) \\
D_y^z(w_2 - \pi^z_2) \\
L_F(y - w) \\
E_F(g - J_F^T y - A^Tv - z) \\
\end{pmatrix},
\tag{8.8}
\end{equation}

where $D_A = \sigma D_A$, $D_Y^w = \bar{\sigma} D_Y^w$, $D_y^y = \bar{\sigma} D_y^y$, $D_Y^v = \bar{\sigma} D_Y^v$, $D_y^z = \bar{\sigma} D_y^z$, $D_y = \bar{\sigma} D_Y$, $\Delta \bar{v} = (1 + 2\sigma)\Delta y$, $\Delta \bar{y} = (1 + 2\sigma)\Delta v$, $\Delta z_1 = (1 + 2\sigma)\Delta z_1$, $\Delta z_2 = (1 + 2\sigma)\Delta z_2$, $\Delta w_1 = (1 + 2\sigma)\Delta w_1$, and $\Delta w_2 = (1 + 2\sigma)\Delta w_2$. We define

\[
D_W = (L_F^T(D_1^w)^{-1}L_L + L_U^T(D_2^w)^{-1}L_U)^{\dagger} = \bar{\sigma}(L_F^T(D_1^w)^{-1}L_L + L_U^T(D_2^w)^{-1}L_U)^{\dagger} = \bar{\sigma}D_W,
\]
8. The primal-dual trust-region direction 25

with \( D_w = (L_{LF}^T(D_1^w)^{-1}L_{LF} + L_{UF}^T(D_2^w)^{-1}L_{UF})^\dagger \). Similarly, define

\[
\tilde{D}_w = (D_w^\dagger + \sigma I_p)^\dagger.
\]

Premultiplying the equations (8.8) by the matrix

\[
\begin{pmatrix}
I_A & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \sigma I_A & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \sigma I_A & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \sigma I_A & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \sigma I_A & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \sigma I_A & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \sigma I_A & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma I_A & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma I_A & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma I_A
\end{pmatrix}
\begin{pmatrix}
I_x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{\sigma} L_{LF}^T(D_1^w)^{-1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{\sigma} E_{LF}^T(D_1^w)^{-1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{\sigma} L_{UW}^T(D_2^w)^{-1} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{\sigma} E_{UW}^T(D_2^w)^{-1} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{\sigma} L_{FU}^T & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sigma} E_{FU}^T & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sigma} L_{UW}^T & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sigma} E_{UW}^T & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sigma} L_{FU}^T & 0
\end{pmatrix}
\begin{pmatrix}
\Delta \tilde{v} \\
\Delta \tilde{z}_1 \\
\Delta \tilde{z}_2 \\
\Delta \bar{w}_1 \\
\Delta \bar{w}_2 \\
\Delta s_{p} \\
\Delta x_{p} \\
\Delta \tilde{y}
\end{pmatrix}
\]

\[
= -
\begin{pmatrix}
D_A(v - \pi_v) \\
D_1^x(z_1 - \pi_1^x) \\
D_2^x(z_2 - \pi_2^x) \\
D_1^w(w_1 - \pi_1^w) \\
D_2^w(w_2 - \pi_2^w) \\
L_F(y - w + \frac{1}{\sigma}[w - \pi_w]) \\
E_F\left(g - J_T^yw - A^Tv - z + \frac{1}{\sigma}[A^T(v - \pi_v) + z - \pi_z]\right) \\
D_v(y - \pi_v) + \tilde{D}_w(\sigma(y - w) + w - \pi_w)
\end{pmatrix}
\]
The full vector $\Delta x$ substitution, $\Delta x$, of the preceding section implies that the solution of the path-following equations $F_x^T \Delta x = -F(x)$. Using block back-substitution, $\Delta x$, and $\Delta y$ may be computed by solving the equations

$$
\begin{pmatrix}
\hat{H}_v + \sigma I & -J_v^T \\
J_v & \sigma(D_v + \hat{D}_w)
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta y
\end{pmatrix}
= -\begin{pmatrix}
E_v \left( y - A^T v - z + \frac{1}{\sigma} \left[ A^T(v - \bar{v}) + z - \pi^\sigma \right] \right) \\
D_v (y - \bar{v}) + \hat{D}_w (\bar{y} - w) + \bar{w}
\end{pmatrix}.
$$

The full vector $\Delta x$ is then computed as $\Delta x = E_v^T \Delta x$. Using the identity $\Delta s = L_v^T \Delta s_v$ in the sixth block of equations gives

$$
\Delta s = -\sigma \hat{D}_w \left( y + (1 + 2\sigma) \Delta y - w + \frac{1}{\sigma} \left[ w - \pi^\sigma \right] \right).
$$

There are several ways of computing $\Delta w_1$ and $\Delta w_2$. Instead of using the block upper-triangular system above, we use the last two blocks of equations of (8.7) to give

$$
\Delta w_1 = -\frac{1}{1 + \sigma} (S_v^T)^{-1} (w_1 \cdot (L_v (s + \Delta s) - \ell^\ell + \mu^e e) - \mu^w w_1^e)
$$

and

$$
\Delta w_2 = -\frac{1}{1 + \sigma} (S_v^T)^{-1} (w_2 \cdot (u^e - L_v (s + \Delta s) + \mu^e e) - \mu^w w_2^e).
$$

Similarly, using (8.7) to solve for $\Delta z_1$ and $\Delta z_2$ yields

$$
\Delta z_1 = -\frac{1}{1 + \sigma} (X_v^T)^{-1} (z_1 \cdot (E_v (x + \Delta x) - \ell^\ell + \mu^e e) - \mu^w z_1^e)
$$

and

$$
\Delta z_2 = -\frac{1}{1 + \sigma} (X_v^T)^{-1} (z_2 \cdot (u^e - E_v (x + \Delta x) + \mu^e e) - \mu^w z_2^e).
$$

Similarly, using the first block of equations (8.8) to solve for $\Delta v$ gives $\Delta v = -(v - \bar{v})/(1 + \sigma)$, with $\bar{v} = v^e - \frac{1}{\mu^w} (A(x + \Delta x) - b)$. Finally, the vectors $\Delta w_x$ and $\Delta z_x$ are recovered as $\Delta w_x = [y + \Delta y - w]_x$ and $\Delta z_x = [g + H \Delta x - J^T (y + \Delta y) - z]_x$.

9. Summary: equations for the trust-region direction

The results of the preceding section implies that the solution of the path-following equations $F'(v_x) \Delta v_x = -F(v_x)$ with $F$ and $F'$ given by (3.2) and (3.3) may be computed as follows. Let $x$ and $s$ be given primal variables and slack variables such that
\( E_x x = b_x, \ L_x s = h_x \) with \( \ell^x - \mu^u < E_x x, \ E_v x < u^x + \mu^u, \ \ell^x - \mu^u < L_v s, \ L_v s < u^x + \mu^u. \) Similarly, let \( z_1, z_2, w_1, w_2 \) and \( y \) denotes dual variables such that \( w_1 > 0, w_2 > 0, z_1 > 0, \) and \( z_2 > 0. \) Consider the diagonal matrices \( X_1^u = \text{diag}(E_x x - \ell^x + \mu^u e), \ X_2^u = \text{diag}(u^x - E_v x + \mu^u e), \ Z_1 = \text{diag}(z_1), \ Z_2 = \text{diag}(z_2), \ W_1 = \text{diag}(w_1), \ W_2 = \text{diag}(w_2), \ S_1^u = \text{diag}(L_v s - \ell^x + \mu^u e) \) and \( S_2^u = \text{diag}(u^x - L_v s + \mu^u e). \) Given the quantities

\[
\begin{align*}
D_v &= \mu^u I_m, \\
D_A &= \mu^u I_n, \\
(D_1^u)^{-1} &= (X_1^u)^{-1} Z_1, \\
(D_2^u)^{-1} &= (X_2^u)^{-1} Z_2, \\
D_z &= (E_v^T (D_1^u)^{-1} E_v E_v^T (D_2^u)^{-1} E_v)^T, \\
\pi_1^v &= \mu^u (X_1^u)^{-1} z_1^e, \\
\pi_2^v &= \mu^u (X_2^u)^{-1} z_2^e, \\
\pi^z &= E_v^T \pi_1^v - E_v^T \pi_2^v,
\end{align*}
\]

solves the KKT system

\[
\begin{pmatrix}
H_x(x, y) + \sigma I_p + \frac{1}{\sigma} A_x^T D_A^{-1} A_v + \frac{1}{\sigma} E_v D_v^T E_v^T \quad -J_y(x)^T \\
\quad J_y(x) \\
\end{pmatrix}
\begin{pmatrix}
\Delta x_r \\
\Delta y
\end{pmatrix} = - \begin{pmatrix}
E_v (g - J^T y - A^T v - z + \frac{1}{\sigma} [A^T (v - \pi^v) + z - \pi^z]) \\
D_v (y - \pi^v) + \tilde{D}_w (\pi - w) + w - \pi^w
\end{pmatrix}.
\]
Then

\[
\Delta x = E^T \Delta x, \quad \tilde{x} = x + \Delta x, \\
\Delta y = \Delta \tilde{y}/(1 + 2\sigma), \quad \tilde{y} = y + \Delta y, \\
\tilde{s} = s + \Delta s, \\
\tilde{v} = v + \Delta v,
\]

\[
\Delta z_1 = -\frac{1}{1 + \sigma} (X^\mu)^{-1} (z_1 \cdot (E_v \tilde{x} - \ell^\mu + \mu^\delta \varepsilon) - \mu^\delta \varepsilon),
\]

\[
\Delta z_2 = -\frac{1}{1 + \sigma} (X^\mu)^{-1} (z_2 \cdot (u^\mu - E_v \tilde{x} + \mu^\mu \mu - \mu^\mu \mu) - \mu^\mu \mu),
\]

\[
\Delta s = -\sigma \tilde{D}_w \left( y + (1 + 2\sigma) \Delta y - w + \frac{1}{\sigma} [w - \pi^w] \right),
\]

\[
\Delta w_1 = -\frac{1}{1 + \sigma} (S^\mu)^{-1} (w_1 \cdot (L \tilde{s} - \ell^\mu + \mu^\delta \varepsilon) - \mu^\delta \varepsilon w_1),
\]

\[
\Delta w_2 = -\frac{1}{1 + \sigma} (S^\mu)^{-1} (w_2 \cdot (u^\delta - L \tilde{s} + \mu^\mu \mu - \mu^\mu \mu w_2),
\]

\[
\hat{v}^v = v^e - \frac{1}{\mu^\delta} (A \tilde{x} - b),
\]

\[
\Delta v = -\frac{1}{1 + \sigma} (v - \hat{v}^v),
\]

\[
w = L^T w, \quad \hat{v} = v + \Delta v,
\]

\[
z = E^T \tilde{x} + E^T \tilde{z}_1 - E^T \tilde{z}_2,
\]

\[
\Delta x = [\tilde{y} - w]_x, \\
\Delta z = [g + H \Delta x - J^T \tilde{y} - z]_x.
\]

10. **Solution of the trust-region equations with an arbitrary right-hand-side**

Moré and Sorensen define a routine \texttt{znull()} that uses the Cholesky factors of \( B_N + \sigma I \) and the condition estimator proposed by Cline, Moler, Stewart and Wilkinson [2]. As the method of Gill, Kungurtsev and Robinson does not compute an explicit factorization of \( B_N + \sigma I \), we define \texttt{znull()} using the condition estimator \texttt{DLACON} supplied with \texttt{LAPACK} [1]. This routine generates an approximate null vector using Higham’s [7] modification of Hager’s algorithm [6]. This routine uses matrix-vector products with \( (B_N + \sigma I)^{-1} \), rather than a matrix factorization, to estimate \( \| (B_N + \sigma I)^{-1} \|_1 \). By-products of the computation of \( \| (B_N + \sigma I)^{-1} \|_1 \) are vectors \( v \) and \( w \) such that \( w = (B_N + \sigma I)^{-1} v \), \( \| v \|_1 = 1 \) and

\[
\| (B_N + \sigma I)^{-1} v \|_1 = \| w \|_1 \approx \| (B_N + \sigma I)^{-1} \|_1 = \max \| (B_N + \sigma I)^{-1} u \|_1.
\]

Thus, unless \( \| w \| = 0 \), the vector \( y = w/\| w \| \) is a unit approximate null vector from which we determine an appropriate \( z \) such that \( \| \Delta v_m + z \|_x = \delta \).
The reduced trust-region equations with a general right-hand side are given by

\[
\begin{pmatrix}
\bar{\sigma}D_A & 0 & 0 & 0 & 0 & A_F & 0 \\
0 & \bar{\sigma}D_1^2 & 0 & 0 & 0 & E_{LF} & 0 \\
0 & 0 & \bar{\sigma}D_2^2 & 0 & 0 & -E_{UF} & 0 \\
0 & 0 & 0 & \bar{\sigma}D_1^y & 0 & L_{LF} & 0 \\
0 & 0 & 0 & 0 & \bar{\sigma}D_2^y & -L_{UF} & 0 \\
0 & 0 & 0 & 0 & 0 & \bar{\sigma}I_p & 0 \\
0 & 0 & 0 & 0 & 0 & \bar{\sigma}I_p & 0 \\
-\bar{A}_F^T & -E_{LF}^T & E_{UF}^T & 0 & 0 & 0 & H_F + \sigma I_p^x - J_F^T \\
0 & 0 & 0 & 0 & -L_{LF}^T & L_{UF}^T & \sigma I_p \\
\end{pmatrix}
\begin{pmatrix}
\tilde{q}_1 \\
\tilde{q}_2 \\
\tilde{q}_3 \\
\tilde{q}_4 \\
\tilde{q}_5 \\
\tilde{q}_6 \\
\tilde{q}_7 \\
\tilde{q}_8 \\
\end{pmatrix}
= 
\begin{pmatrix}
r_1 \\
r_2 \\
r_3 \\
r_4 \\
r_5 \\
r_6 \\
r_7 \\
r_8 \\
\end{pmatrix}
\]

Premultiplying these equations by

\[
\begin{pmatrix}
I_A & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & I_{LF}^x & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & I_{LF}^y & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & I_{LF} & 0 & 0 & 0 & 0 \\
\frac{1}{\sigma}A_F^{-1}D_A^{-1} & \frac{1}{\sigma}E_{LF}^T(D_1)^{-1} & -\frac{1}{\sigma}E_{UF}^T(D_2)^{-1} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}
\begin{pmatrix}
\tilde{q}_1 \\
\tilde{q}_2 \\
\tilde{q}_3 \\
\tilde{q}_4 \\
\tilde{q}_5 \\
\tilde{q}_6 \\
\tilde{q}_7 \\
\tilde{q}_8 \\
\end{pmatrix}
\]

gives the block upper-triangular system

\[
\begin{pmatrix}
\bar{\sigma}D_A & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \bar{\sigma}D_1^2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \bar{\sigma}D_2^2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \bar{\sigma}D_1^y & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \bar{\sigma}D_2^y & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \bar{\sigma}I_p & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \bar{\sigma}I_p & 0 & 0 \\
-\bar{A}_F^T & -E_{LF}^T & E_{UF}^T & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -L_{LF}^T & L_{UF}^T & \sigma I_p & 0 \\
\end{pmatrix}
\begin{pmatrix}
\tilde{q}_1 \\
\tilde{q}_2 \\
\tilde{q}_3 \\
\tilde{q}_4 \\
\tilde{q}_5 \\
\tilde{q}_6 \\
\tilde{q}_7 \\
\tilde{q}_8 \\
\end{pmatrix}
= 
\begin{pmatrix}
r_1 \\
r_2 \\
r_3 \\
r_4 \\
r_5 \\
r_6 \\
r_7 \\
r_8 \\
\end{pmatrix}
\]

\[
\begin{pmatrix}
\tilde{q}_1 \\
\tilde{q}_2 \\
\tilde{q}_3 \\
\tilde{q}_4 \\
\tilde{q}_5 \\
\tilde{q}_6 \\
\tilde{q}_7 \\
\tilde{q}_8 \\
\end{pmatrix}
= 
\begin{pmatrix}
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
\end{pmatrix}
\]
with \( \tilde{H}_F = E_F \left( H + \frac{1}{\sigma} (J^T D_\gamma^{-1} J + A^T D_\alpha^{-1} A + D_\delta^T) \right) E_F^T \). Using block back-substitution, \( \tilde{q}_7 \) and \( \tilde{q}_8 \) can be computed by solving the equations

\[
\begin{pmatrix}
\tilde{H}_F & -J_F^T \\
J_F & \tilde{\sigma}(D_V + \bar{D}_W)
\end{pmatrix}
\begin{pmatrix}
\tilde{q}_7 \\
\tilde{q}_8
\end{pmatrix} = \begin{pmatrix}
\frac{1}{\sigma} E_F \left( A^T D_\alpha^{-1} r_1 + E_F^T (D_1^T)^{-1} r_2 - E_F^T (D_2^T)^{-1} r_3 + \bar{\sigma} r_7 \right) \\
\bar{D}_w \left( L_F^T (D_1^w)^{-1} r_4 - L_F^T (D_2^w)^{-1} r_5 + \bar{\sigma} r_6 \right) + r_8
\end{pmatrix},
\]

with the remaining vectors computed as

\[
\begin{align*}
\tilde{q}_6 &= \bar{D}_w \left( L_F^T (D_1^w)^{-1} r_4 - L_F^T (D_2^w)^{-1} r_5 + \bar{\sigma} (r_6 - \tilde{q}_8) \right) \\
\tilde{q}_5 &= \frac{1}{\sigma} (D_2^w)^{-1} (r_5 - L_F^T \tilde{q}_6) \\
\tilde{q}_4 &= \frac{1}{\sigma} (D_1^w)^{-1} (r_4 - L_F^T \tilde{q}_6) \\
\tilde{q}_3 &= \frac{1}{\sigma} (D_2^z)^{-1} (r_3 - E_F E_F^T \tilde{q}_7) \\
\tilde{q}_2 &= \frac{1}{\sigma} (D_1^z)^{-1} (r_2 - E_F \tilde{q}_7) \\
\tilde{q}_1 &= \frac{1}{\sigma} (D_\alpha)^{-1} (r_1 - A_F \tilde{q}_7).
\end{align*}
\]

References


