Cahier du GERAD G-2012-42

BOUNDS ON EIGENVALUES OF MATRICES ARISING FROM INTERIOR-POINT METHODS

CHEN GREIF*, ERIN MOULDING†, AND DOMINIQUE ORBAN‡

Abstract. Interior-point methods feature prominently among numerical methods for inequality-constrained optimization problems, and involve the need to solve a sequence of linear systems that typically become increasingly ill-conditioned with the iterations. To solve these systems, whose original form has a nonsymmetric 3×3 block structure, it is common practice to perform block elimination and either solve the resulting reduced saddle-point system, or further reduce the system to the Schur complement and apply a symmetric positive definite solver. In this paper we use energy estimates to obtain bounds on the eigenvalues of the matrices, and conclude that the original unreduced matrix has more favorable eigenvalue bounds than the alternative reduced versions. We also examine regularized variants of those matrices that do not require typical regularity assumptions.

Keywords. primal-dual interior-point methods, convex quadratic optimization, indefinite linear systems, eigenvalues, condition number, inertia, eigenvalue bounds, regularization

1. Introduction. Given a symmetric and positive semidefinite Hessian matrix $H \in \mathbb{R}^{n \times n}$, vectors $c \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$, and a Jacobian matrix $J \in \mathbb{R}^{m \times n}$, where $m \leq n$, consider the primal-dual pair of quadratic programs (QP) in standard form

$$\underset{x}{\text{minimize}} \;\; c^T x + \tfrac{1}{2} x^T H x \quad \text{subject to} \qquad\qquad J x = b, \quad x \geq 0, \tag{1.1a}$$

$$\underset{x,y,z}{\text{maximize}} \;\; b^T y - \tfrac{1}{2} x^T H x \quad \text{subject to } J^T y + z - Hx = c, \quad z \geq 0, \tag{1.1b}$$

where inequalities are understood elementwise, and y and z are the vectors of Lagrange multipliers associated with the equality and nonnegativity constraints of (1.1a), respectively. The case $H = 0$ corresponds to the linear programming (LP) problem in standard form. Numerical methods for solving (1.1) include the class of widely successful primal-dual interior-point methods. Their distinctive feature is that they approximately follow a smooth path lying inside the primal-dual feasible set all the way to an optimal solution.

This paper focuses on the linear systems that form the core of the iterations of primal-dual interior-point methods. Specifically, the matrices associated with those linear systems have a special block form, and techniques that rely on partial elimination of the unknowns are fairly popular, as the underlying sparsity pattern naturally lends itself to such reductions. However, we claim that in terms of spectral distribution and conditioning, it may be beneficial to avoid performing such elimination steps before applying a linear solver. To make our point, we use energy estimates in the spirit of Rusten and Winther (1992), and obtain upper and lower bounds on the eigenvalues of the various matrices that we consider. We also consider *regularized* variants of those matrices that arise when an interior-point method is applied to a modified optimization problem that includes regularization terms.

*Department of Computer Science, The University of British Columbia, Vancouver, BC, Canada. E-mail: greif@cs.ubc.ca. Research partially supported by an NSERC Discovery Grant.

†Department of Mathematics, The University of British Columbia, Vancouver, BC, Canada. E-mail: moulding@math.ubc.ca. Research partially supported by an NSERC CGS M Scholarship.

‡GERAD and Department of Mathematics and Industrial Engineering, École Polytechnique, Montréal, QC, Canada. E-mail: dominique.orban@gerad.ca. Research partially supported by an NSERC Discovery Grant.

Our primary goal is to provide a theoretical foundation to the study of spectral properties of the matrices involved. When the problem is very large, the spectral structure of the matrices plays a central role in the performance of the interior-point solver, especially when the underlying linear systems are solved iteratively. We stress however that we do not perform analysis for iterative solvers or offer specialized preconditioning approaches in the present paper.

In §2 we provide a short overview of primal-dual interior-point methods and present the linear systems that arise throughout the iterations. In §3 we investigate the linear systems in all formulations, specify conditions for nonsingularity, provide bounds on the eigenvalues, and study the inertia and condition numbers. In §4 we analyze a regularized optimization problem in the same way. In §5 we provide numerical validation of our analytical claims and in §6 we cover several alternative system formulations. Concluding remarks appear in §7.

2. Background and Preliminaries. In this section we provide a brief overview of primal-dual interior-point methods and the linear systems that arise throughout the iterations. Our purpose is to set the stage for the analysis of the subsequent sections.

For an index set $\mathcal{N} \subseteq \{1, \dots, n\}$ and a vector $v \in \mathbb{R}^n$ we denote by $v_{\mathcal{N}}$ the subvector of v indexed by $\mathcal{N}$. Similarly, if A is a matrix with n columns, $A_{\mathcal{N}}$ is the submatrix of the columns of A corresponding to indices in $\mathcal{N}$. If A is square, $A_{\mathcal{N}\mathcal{N}}$ represents the matrix with both rows and columns corresponding to indices in $\mathcal{N}$.

Throughout the paper we separate vector components by commas. To avoid ambiguities, inner products are denoted by a transpose operation. Thus, (x, y) means a vector whose components are x and y (each of which may be a vector) whereas $x^T y$ means an inner product of the vectors x and y.

2.1. Primal-Dual Interior-Point Methods. If x is feasible for (1.1a), we let

$$\mathcal{A}(x) := \{i = 1, \dots, n \mid x_i = 0\} \quad \text{and} \quad \mathcal{I}(x) := \{1, \dots, n\} \setminus \mathcal{A}(x)$$

be the index sets of active and inactive bounds, respectively. For simplicity and when there is no ambiguity, we will write $\mathcal{A}$ and $\mathcal{I}$ instead of $\mathcal{A}(x)$ and $\mathcal{I}(x)$. All solutions (x, y, z) of (1.1) must satisfy the complementarity condition

$$x_i z_i = 0 \quad \text{for all} \quad i = 1, \dots, n,$$

which may also be written $z_{\mathcal{I}(x)} = 0$. A solution (x, y, z) of (1.1) is *strictly complementary* if $z_i > 0$ for all $i \in \mathcal{A}(x)$, which may also be written $z_{\mathcal{A}} > 0$.

It is generally assumed that Slater's constraint qualification condition holds, i.e., that there exists a primal-dual triple (x, y, z) such that $Jx = b$, $J^T y + z - Hx = c$ and $(x, z) > 0$. Failing this assumption, there may exist no solution to (1.1). Primal-dual interior-point methods generate a sequence of iterates (x_k, y_k, z_k) that remain strictly feasible with respect to the bound constraints, i.e., $(x_k, z_k) > 0$, but not necessarily to the equality constraints, with the intention of satisfying, in the limit, the common necessary and sufficient first-order optimality conditions of (1.1a) and (1.1b). Those iterates (approximately) satisfy the perturbed optimality conditions

$$\begin{bmatrix} c + Hx - J^T y - z \\ Jx - b \\ \tau e - XZe \end{bmatrix} = 0, \qquad (x, z) > 0. \tag{2.1}$$

Here the triple (x, y, z) represents a generic current iterate. We drop the subscript k for brevity and we use the standard notation X and Z to denote diagonal matrices whose diagonal elements are the vectors x and z, respectively, and e to represent the vector of ones of appropriate size. The system depends on $\tau > 0$, the *barrier parameter*, which governs the progress of the interior-point method and converges to zero. This parameter is typically set as the product $\tau := \sigma\mu$, where

$$\mu := \frac{x^T z}{n}$$

is the *duality measure* and $\sigma \in [0, 1]$ is the *centering parameter* used to achieve a desirable reduction in the duality measure at each iteration. Tied to these parameters is the notion of *central path*, which is the set of triples $(x, y, z) = (x_\tau, y_\tau, z_\tau)$ of exact solutions to (2.1) associated with $\tau > 0$. Practical methods typically seek a compromise between improving centrality by taking a relatively large value of σ, which allows for taking a longer step in the next iteration, and reducing the duality measure μ by taking a small value of σ, which results in a direction closer to the pure Newton direction, often called the affine-scaling direction.

The matrix associated with the interior-point iterations is the Jacobian of the system of equations (2.1). This matrix is of size $(2n + m) \times (2n + m)$ and can be most naturally written in a block 3×3 form. The linear system is given by

$$\begin{bmatrix} H & J^T & -I \\ J & 0 & 0 \\ -Z & 0 & -X \end{bmatrix} \begin{bmatrix} \Delta x \\ -\Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} -c - Hx + J^T y + z \\ b - Jx \\ XZe - \tau e \end{bmatrix} := \begin{bmatrix} r_c \\ r_b \\ r_\tau \end{bmatrix}. \tag{2.2}$$

Most blocks of the matrix stay fixed throughout the interior-point iterations. The only ones that change, X and Z, are diagonal with strictly positive diagonal elements during the iteration, although in the limit some diagonal elements will typically become zero. We denote the matrix of (2.2) by K_3, the subscript '3' standing for the size of the block system.

We note that there are several ways to arrange K_3 of (2.2); we choose this formulation in anticipation of symmetrizing the matrix in the form of a suggestion of Saunders—see (Forsgren, 2002). The barrier parameter appears explicitly only in the right-hand side, but it also influences the matrix itself since the iterates x and z, if they converge, do so at an asymptotic rate that is a function of the duality measure μ. The solution of (2.2) for $(\Delta x, \Delta y, \Delta z)$ defines the next iterate

$$(x^+, y^+, z^+) = (x, y, z) + \alpha(\Delta x, \Delta y, \Delta z),$$

where $\alpha \in (0, 1]$ is a step length chosen to ensure that $(x^+, z^+) > 0$.

2.2. Block Elimination Approaches. Given the block-structure of K_3 shown in (2.2), a few possibilities for solving the system naturally emerge due to the special sparsity structure, particularly the diagonality and positive definiteness of X and Z. An obvious approach is that of directly solving the linear system (2.2). The matrix K_3 is mildly nonsymmetric but easily symmetrizable (see §3.3.2), and so it is possible to apply symmetric solvers.

A second approach is that of exploiting the nonsingularity and diagonality of X to perform one step of block Gaussian elimination and obtain

$$\begin{bmatrix} H + X^{-1}Z & J^T \\ J & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} r_c - X^{-1} r_\tau \\ r_b \end{bmatrix}. \tag{2.3}$$

The matrix of (2.3), denoted K_2, is a typical symmetric indefinite saddle-point matrix and has size $(n + m) \times (n + m)$. Significant progress has been made on numerical solution methods for systems of this form in the past few decades (Benzi et al., 2005), but here a specific challenge is that $X^{-1}Z$ may cause ill conditioning as X and Z have diagonal entries iterating toward zero. The stability of the symmetric indefinite factorization of K_2 has been studied in Forsgren et al. (1996).

A third common approach is to take an additional step of block Gaussian elimination before applying a linear solver. This amounts to forming the Schur complement equations

$$\left[J(H+X^{-1}Z)^{-1}J^T\right]\Delta y=\left[r_b-J(H+X^{-1}Z)^{-1}(r_c-X^{-1}r_\tau)\right]. \tag{2.4}$$

The matrix of (2.4) is denoted K_1 and is positive definite. This approach is popular with practitioners, since symmetric positive definite solvers are often preferred over indefinite solvers, and the Schur complement equations are smaller, of size $m\times m$. However, in cases other than linear programming, forming system (2.4) comes at the potentially costly price of having to first invert or factorize $H+X^{-1}Z$ and having to deal with a significant loss of sparsity.

2.3. Related Work. While the algorithms of modern interior-point solvers are mostly settled, the choice of linear system formulation differs across software packages. Many modern solvers reduce to the Schur complement equations form, e.g., PCx for linear programming (Czyzyk et al., 1999). Others reduce to the saddle-point form, e.g., OOQP for quadratic programming (Gertz and Wright, 2003) and IPOPT and KNITRO for general nonlinear programming (Byrd et al., 1999, 2000; Wächter and Biegler, 2006). Another example is HOPDM for linear programming and convex quadratic programming, which automatically chooses either the Schur complement equations or saddle-point form (Altman and Gondzio, 1999). We are not aware of existing solvers that solve the unreduced system (2.2) for any of these problems.

Properties of the Schur complement equations form (2.4) are straightforward, and there is relatively little analysis of this form. The saddle-point formulation (2.3) has properties that directly follow from the general properties of such matrices—see (Rusten and Winther, 1992; Benzi et al., 2005; Silvester and Wathen, 1994; Gould and Simoncini, 2009) for some relevant general results, and (Friedlander and Orban, 2012) for results specialized to optimization. The ill-conditioning of some reduced matrices is well-known (Fiacco and McCormick, 1990; Wright, 1992, 2005; Forsgren, 2002; Forsgren et al., 2002), but it has been referred to, with some assumptions on solution methods, as "benign" (Wright, 2005; Forsgren, 2002), "usually harmless" (Forsgren et al., 2002), and "highly structured" (Forsgren et al., 2002). The matrices for classical barrier methods, corresponding to the choice $Z=\tau X^{-1}$ are also ill-conditioned (Wright, 1994; Forsgren et al., 2002).

There exist few results on the unreduced 3×3 formulation. Korzak (1999) covers some spectral properties of various formulations for the special case of linear programming. Armand and Benoist (2011) prove uniform boundedness of the inverse under several assumptions, intended to be used in further theoretical analysis. A private communication of Saunders is cited by Forsgren (2002) who notes the symmetrizability and potential appeal of the 3×3 system, equivalent to the symmetrized matrix used in this paper. Forsgren (2002) and Forsgren et al. (2002) note that the matrix of this system remains well-conditioned though ill-conditioning remains an issue when forming the right-hand side and unscaling the variables, due to multiplication by a diagonal matrix with large elements; these papers mention also a different (ill-conditioned) symmetric formulation.

2.4. Notation and Further Definitions. We denote the eigenvalues of the Hessian H by $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$, and the singular values of J by $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_m \geq 0$. The spectral condition number of a matrix B, defined as $\sigma_{\max}(B)/\sigma_{\min}(B)$, will be denoted by $\kappa(B)$, where $\sigma_{\max}$ and $\sigma_{\min}$ denote the largest and smallest singular values, respectively.

For two related positive scalar sequences $\{\alpha_k\}$ and $\{\beta_k\}$, we write $\alpha_k = O(\beta_k)$ if there exists a constant $\varphi > 0$ such that $\alpha_k \leq \varphi\beta_k$ for all sufficiently large k. Equivalently, we write $\beta_k = \Omega(\alpha_k)$. We write $\alpha_k = \Theta(\beta_k)$ if $\alpha_k = O(\beta_k)$ and $\beta_k = O(\alpha_k)$. If $\alpha_k = O(\beta_k)$ and $\alpha_k = \Omega(\gamma_k)$ for some third positive sequence $\{\gamma_k\}$, we use the shorthand notation $\gamma_k \lesssim \alpha_k \lesssim \beta_k$.

Our main working assumption is that (1.1a) is convex. We formalize this assumption now and refer to it throughout the paper.

Assumption 2.1 (Convexity). *The Hessian H is symmetric and positive semidefinite.*

Throughout our analysis, the following two definitions will be useful.

Definition 2.1 (Inertia). *For a given square matrix M with real eigenvalues, the inertia of M is the triple (n^+, n^-, n^0) where n^+, n^- and n^0 are the numbers of positive, negative, and zero eigenvalues of M, respectively.*

We note that if C is an arbitrary nonsingular matrix and if $N = CMC^T$, Sylvester's law of inertia asserts that N and M have the same inertia.

The following definition states a standard qualification condition required to ensure certain nonsingularity properties.

Definition 2.2 (LICQ). *The Linear Independence Constraint Qualification condition is satisfied at x, feasible for (1.1a), if $\begin{bmatrix} J^T & -I_{\mathcal{A}(x)} \end{bmatrix}$ has full column rank.*

Note that the LICQ imposes an upper bound on the size of the active set: $|\mathcal{A}(x)| \leq n - m$. If the LICQ and strict complementarity are satisfied, we say that the problem is nondegenerate. As we shall see, these conditions guarantee that the matrices of interest are nonsingular. These are common assumptions in optimization.

Throughout the paper, we illustrate our bounds on the following generic, but typical, example situation.

Example 2.1. *We consider a generic interior-point method guaranteeing the following asymptotic estimates:*

$$x_i = \Theta(\mu) \quad (i \in \mathcal{A}), \qquad x_i = \Theta(1) \quad (i \in \mathcal{I}), \tag{2.5a}$$
$$z_i = \Theta(\mu) \quad (i \in \mathcal{I}), \qquad z_i = \Theta(1) \quad (i \in \mathcal{A}). \tag{2.5b}$$

We assume that $\mathcal{A} \neq \varnothing$ and $\mathcal{I} \neq \varnothing$.

Most problems are such that $\mathcal{A} \neq \varnothing$ and $\mathcal{I} \neq \varnothing$ and most interior-point methods applied to a nondegenerate problem match the situation of Example 2.1. In particular, when (x, y, z) are exact solutions of (2.1), we have $X^{-1}Z = \mu X^{-2} = \Theta(1/\mu)$, but this estimate also holds sufficiently close to the central path. Indeed most interior-point algorithms for convex quadratic programming confine the iterates to a neighborhood of the central path defined, among others, by the condition that $\gamma_1\mu \leq x_i z_i \leq \gamma_2\mu$ for all $i = 1, \ldots, n$, for some positive constants γ_1 and γ_2. That (2.5) holds under strict complementarity is then a simple consequence. See, e.g., (Wright, 1997).

In our implementation we use the predictor-corrector scheme due to Mehrotra (1992), which is based on first taking the pure Newton direction, i.e., with $\sigma = 0$, and then following a step aiming toward the central path as a correction for the linearization error in XZ. The algorithm thus solves two linear systems with the same matrix but with different right-hand sides. Although this algorithm does not confine the iterates to a neighborhood of the sort mentioned above, we will assume that (2.5) holds.

Our results essentially consider two distinct situations to analyze the properties of K_1, K_2 and K_3. The first concerns values of the matrix throughout the iterations while the second concerns the value in the limit, at a point satisfying complementarity, such as a solution of (1.1). We formalize those situations as assumptions for later reference and note that the two assumptions are mutually exclusive.

Assumption 2.2 (Positivity Assumption). *The triple* (x, y, z) *satisfies* $(x, z) > 0$.

Assumption 2.3 (Complementarity Assumption). *The triple* (x, y, z) *satisfies* $(x, z) \geq 0$ *and* $x_i z_i = 0$ *for all* $i = 1, \ldots, n$.

3. Analysis of the Traditional Systems.

In this section we provide eigenvalue analysis for systems (2.4), (2.3), and (2.2), in this order.

3.1. The Schur Complement System.

In (2.4), K_1 is positive definite even if H is only semidefinite, provided that J has full row rank. The positive definiteness makes the approach of reducing the original system (2.2) to (2.4) potentially attractive. On the other hand, reducing the system this way requires the inversion of $H + X^{-1}Z$ and may cause other types of computational difficulties, such as potential loss of sparsity. The result below follows directly from an elementary analysis of the eigenvalue equation for K_1.

Lemma 3.1. *Let* K_1 *be defined as in* (2.4) *and suppose Assumptions 2.1 and 2.2 hold and* J *has full row rank. The eigenvalues of* K_1 *are contained in the interval*

$$\left[\frac{\sigma_m^2}{\lambda_{\max}(H + X^{-1}Z)}, \ \frac{\sigma_1^2}{\lambda_{\min}(H + X^{-1}Z)}\right].$$

As a consequence, we have an upper bound on the spectral condition number of K_1*:*

$$\kappa(K_1) \leq \kappa(J)^2 \ \kappa(H + X^{-1}Z).$$

In the typical situation of Example 2.1, it is clear that K_1 approaches singularity. We have the asymptotic estimates

$$\lambda_{\min}(H + X^{-1}Z) = \Omega(\lambda_n + \mu) \quad \text{and} \quad \lambda_{\max}(H + X^{-1}Z) = \Theta(1/\mu). \tag{3.1}$$

As $\mu \to 0$, Lemma 3.1 thus yields the asymptotic estimate

$$\kappa(K_1) = \kappa(J)^2 \, O\Big(1/(\mu(\lambda_n + \mu))\Big). \tag{3.2}$$

Note finally that at least in the case of linear programming, the bounds of Lemma 3.1 and (3.2) are tight and (3.2) becomes $\kappa(K_1) = \kappa(J)^2 \, \Theta(1/\mu^2)$. On the other hand, when $\lambda_n > 0$, (3.2) becomes $\kappa(K_1) = \kappa(J)^2 \, O(1/\mu)$. This last estimate is in line with those of Wright (1994), who assumes that a second-order sufficiency condition holds.

3.2. The 2×2 Block System.

It is easy to see that K_2 is nonsingular during the iterations if and only if J has full row rank. For this type of linear system much is known in the literature, and we state now observations that can be concluded from existing results, e.g., (Benzi et al., 2005, §3) and (Rusten and Winther, 1992, §2).

LEMMA 3.2. *If J has full row rank and Assumptions 2.1 and 2.2 are satisfied, the inertia of K_2 is $(n, m, 0)$.*

LEMMA 3.3. *If H is positive semi-definite, J has full row rank and Assumptions 2.1 and 2.2 are satisfied, then*

$$\lambda^+ \geq \lambda_{\min}(H + X^{-1}Z),$$
$$\lambda^+ \leq \tfrac{1}{2}\left(\lambda_{\max}(H + X^{-1}Z) + \sqrt{\lambda_{\max}(H + X^{-1}Z)^2 + 4\sigma_1^2}\right)$$

for any positive eigenvalue λ^+ of K_2, and

$$\lambda^- \geq \tfrac{1}{2}\left(\lambda_{\min}(H + X^{-1}Z) - \sqrt{\lambda_{\min}(H + X^{-1}Z)^2 + 4\sigma_1^2}\right),$$
$$\lambda^- \leq \tfrac{1}{2}\left(\lambda_{\max}(H + X^{-1}Z) - \sqrt{\lambda_{\max}(H + X^{-1}Z)^2 + 4\sigma_m^2}\right)$$

for any negative eigenvalue λ^- of K_2.

From Lemma 3.3, which follows from (Rusten and Winther, 1992, Lemma 2.1), we see that the lower bound on the negative eigenvalues of K_2 is finite and bounded away from zero unless $J = 0$. It is the other three bounds that are responsible for the ill-conditioning of K_2. Using again the situation of Example 2.1 where the extremal eigenvalues of $H + X^{-1}Z$ are approximated by (3.1), we obtain the asymptotic estimates

$$\lambda_n + \mu \lesssim \lambda^+ \lesssim 1/\mu \qquad \text{and} \qquad \tfrac{1}{2}\left(\lambda_n - \sqrt{\lambda_n^2 + 4\sigma_1^2}\right) \lesssim \lambda^- \lesssim -\mu\,\sigma_m^2,$$

where the upper bound on the negative eigenvalues is obtained via the Taylor expansion $\sqrt{1+x} \approx 1 + \frac{1}{2}x$ for $x \ll 1$. These estimates yield the asymptotic condition number

$$\kappa(K_2) = O\left(\frac{1/\mu}{\min(\mu\sigma_m^2, \lambda_n + \mu)}\right) = O\left(\frac{1}{\mu^2}\right).$$

Several authors, including Fourer and Mehrotra (1993) and Korzak (1999) suggest scalings of K_2 that alleviate this ill-conditioning.

The above asymptotic estimates must be considered cautiously, as they do not always fully capture the actual value of the condition number. Nevertheless, they illustrate the ill-conditioning of the 1×1 and 2×2 formulations.

3.3. The 3×3 Block System. In this section we perform eigenvalue analysis for the matrix K_3. A challenge here is that the 3×3 block form gives rise to rather complicated cubic inequalities in some cases. As we show, simplifying assumptions using the limiting behavior of elements in X and Z lead to effective bounds, although the case of an upper bound on the negative eigenvalues proves to be significantly harder to deal with, and in fact we do not have satisfying results for it. The analysis for the 3×3 system in this section and §4 forms the core of our new results.

3.3.1. Nonsingularity of K_3. We first recall that K_3 is nonsingular throughout the interior-point iterations.

THEOREM 3.4. *Suppose Assumptions 2.1 and 2.2 are satisfied. The matrix K_3 in (2.2) is nonsingular if and only if J has full rank.*

Proof. Suppose (u, v, w) lies in the nullspace of K_3, that is,

$$\begin{bmatrix} H & J^T & -I \\ J & 0 & 0 \\ -Z & 0 & -X \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{3.3}$$

The third block row of (3.3) yields $w = -X^{-1}Zu$. Taking the inner product of the first block row with u, substituting for w and using the second block row yields

$$u^T(H + X^{-1}Z)u + (Ju)^T v = u^T(H + X^{-1}Z)u = 0. \tag{3.4}$$

Since X and Z are diagonal and strictly positive throughout the iteration, $H + X^{-1}Z$ is positive definite, and hence (3.4) gives $u = 0$, which implies $w = 0$. It readily follows from the first block row of (3.3) that $J^T v = 0$. If J has full rank, this implies $v = 0$.

If J does not have full row rank, any (u, v, w) such that $v \in \text{Null}(J^T)$, $v \neq 0$, $u = 0$, $w = 0$ is a nontrivial null vector. □

REMARK 1. *Inspection of the proof of Theorem 3.4 reveals that our initial assumption of positive semidefinite H may be weakened. Indeed, it is sufficient to assume that H is positive semidefinite on the nullspace of J only, since then (3.4) still implies that $u = 0$. In this case however, (1.1a) is no longer a convex QP and the duality relationship with (1.1b) no longer holds. Nevertheless, such a restricted definiteness assumption is classic in nonconvex optimization—see, e.g., Gould (1985).*

We now consider what happens to K_3 in the limit of the interior-point method.

THEOREM 3.5. *Suppose Assumptions 2.1 and 2.3 hold at (x, y, z). Then K_3 is nonsingular if and only if the solution (x, y, z) is strictly complementary, $\text{Null}(H) \cap \text{Null}(J) \cap \text{Null}(Z) = \{0\}$, and the LICQ is satisfied.*

Proof. If (x, y, z) is not strictly complementary, there is a zero row in the third block row of (2.2) and K_3 is singular. Therefore, strict complementarity is necessary.

If $\text{Null}(H) \cap \text{Null}(J) \cap \text{Null}(Z) \neq \{0\}$, take $0 \neq u \in \text{Null}(H) \cap \text{Null}(J) \cap \text{Null}(Z)$, $v = 0$ and $w = 0$. Since $Hu = Ju = -Zu = 0$, it follows from (3.3) that (u, v, w) is a nontrivial null vector of K_3. Thus this condition is necessary.

Now, assume strict complementarity and $\text{Null}(H) \cap \text{Null}(J) \cap \text{Null}(Z) = \{0\}$, and suppose (u, v, w) is in the nullspace of K_3. Since $z_{\mathcal{I}} = 0$ at the solution (see §2.1) by complementarity, we have $Z = \text{diag}\,(z_{\mathcal{A}}, z_{\mathcal{I}}) = \text{diag}\,(z_{\mathcal{A}}, 0)$, with $z_{\mathcal{A}} > 0$, so that $\text{Null}(Z) = \text{span}\{e_i | i \in \mathcal{I}\}$. The third block row of (3.3) and strict complementarity necessarily yield $u_{\mathcal{A}} = 0$ and $w_{\mathcal{I}} = 0$, and so $u = (0, u_{\mathcal{I}})$, which implies that u lies entirely in the nullspace of Z. Therefore, $u^T w = 0$. Taking the inner product of the first block row of (3.3) with u and substituting $Ju = 0$ from the second block row gives $u^T Hu = 0$. We must thus have $u \in \text{Null}(H) \cap \text{Null}(J) \cap \text{Null}(Z)$, which implies that $u = 0$. Eliminating u and $w_{\mathcal{I}}$ from (3.3), we have

$$\begin{bmatrix} J^T & -I_{\mathcal{A}} \end{bmatrix} \begin{bmatrix} v \\ w_{\mathcal{A}} \end{bmatrix} = 0,$$

which has only the trivial solution if and only if the LICQ holds. □

3.3.2. Inertia. Having established conditions for the nonsingularity of K_3, we now consider its eigenvalues. To that end, it is useful to consider a symmetrized version of the matrix, making it possible to work with real arithmetic.

Indeed, it is easy to show that K_3 is symmetrizable and has real eigenvalues. Consider the diagonal matrix

$$D = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & Z^{\frac{1}{2}} \end{bmatrix}. \tag{3.5}$$

Using the similarity transformation associated with D, we obtain the symmetric matrix

$$\hat{K}_3 := D^{-1} K_3 D = \begin{bmatrix} H & J^T & -Z^{\frac{1}{2}} \\ J & 0 & 0 \\ -Z^{\frac{1}{2}} & 0 & -X \end{bmatrix}. \tag{3.6}$$

We find results on the inertia of $\hat{K}_3$ both during the iterations and in the limit. We first need the following result.

Lemma 3.6 (Gould (1985), Lemma 3.1). *Let S be an arbitrary $q \times q$ matrix and let I_ℓ be the $\ell \times \ell$ identity matrix. The matrix*

$$B := \begin{bmatrix} 0 & 0 & I_\ell \\ 0 & S & 0 \\ I_\ell & 0 & 0 \end{bmatrix}$$

has ℓ eigenvalues -1, ℓ eigenvalues 1 and q eigenvalues equal to those of S.

Theorem 3.7. *Suppose Assumptions 2.1, $\operatorname{Null}(H) \cap \operatorname{Null}(J) = \{0\}$ and either*

1. *Assumption 2.2 is satisfied and J has full rank, or*
2. *Assumption 2.3 holds at (x, y, z), where strict complementarity and the LICQ are satisfied.*

Then the inertia of $\hat{K}_3$ is $(n, n+m, 0)$.

Proof. Consider first the case where Assumption 2.2 is satisfied and J has full rank. Let the matrix N be an orthonormal nullspace basis matrix for J. Since $\operatorname{Null}(H) \cap \operatorname{Null}(J) = \{0\}$, $N^T H N$ is positive definite. The following construction follows the proof of (Gould, 1985, Lemma 3.2). Complete the basis N so that $\begin{bmatrix} Y & N \end{bmatrix}$ is orthogonal. Then $J \begin{bmatrix} Y & N \end{bmatrix} = \begin{bmatrix} L & 0 \end{bmatrix}$, where L is nonsingular. Define

$$M_1 := \begin{bmatrix} Y & N & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix}.$$

Then we have

$$M_1^T \hat{K}_3 M_1 = \begin{bmatrix} Y^T H Y & Y^T H N & L^T & -Y^T Z^{\frac{1}{2}} \\ N^T H Y & N^T H N & 0 & -N^T Z^{\frac{1}{2}} \\ L & 0 & 0 & 0 \\ -Z^{\frac{1}{2}} Y & -Z^{\frac{1}{2}} N & 0 & -X \end{bmatrix}.$$

Let

$$M_2 := \begin{bmatrix} I_{n-m} & 0 & 0 & 0 \\ 0 & I_m & 0 & 0 \\ -\frac{1}{2}L^{-T}Y^T HY & -L^{-T}Y^T HN & L^{-T} & 0 \\ 0 & 0 & 0 & I \end{bmatrix}.$$

Then

$$M_2^T M_1^T \hat{K}_3 M_1 M_2 = \begin{bmatrix} 0 & 0 & I_m & -Y^T Z^{\frac{1}{2}} \\ 0 & N^T HN & 0 & -N^T Z^{\frac{1}{2}} \\ I_m & 0 & 0 & 0 \\ -Z^{\frac{1}{2}}Y & -Z^{\frac{1}{2}}N & 0 & -X \end{bmatrix}.$$

Finally, let

$$M_3 := \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & R \\ 0 & 0 & I & S \\ 0 & 0 & 0 & I \end{bmatrix},$$

where $R := (N^T HN)^{-1} N^T Z^{\frac{1}{2}}$ and $S := Y^T Z^{\frac{1}{2}}$. The product $M := M_1 M_2 M_3$ is nonsingular, and since $\hat{K}_3$ is nonsingular by Theorem 3.4, the matrix

$$M^T \hat{K}_3 M = \begin{bmatrix} 0 & 0 & I_m & \\ 0 & N^T HN & 0 & \\ I_m & 0 & 0 & \\ & & & G \end{bmatrix}, \tag{3.7}$$

with $G = -X - Z^{\frac{1}{2}} N (N^T HN)^{-1} N^T Z^{\frac{1}{2}}$, is also nonsingular. From the previously noted Lemma 3.6, Sylvester's law of inertia, and the fact that G is negative definite, we obtain that $\hat{K}_3$ has $n + m$ negative and n positive eigenvalues.

Turning now to the case where Assumption 2.3 holds, we see that in the proof above, (3.7) still holds. Clearly, G is at least negative semidefinite, and since M is nonsingular and $\hat{K}_3$ is nonsingular by Theorem 3.5, G must be negative definite. □

Remark 2. *In Theorem 3.7, the assumption of LICQ can be weakened to assume only that J is full rank, and D can be proven negative definite via strict complementarity.*

The fact that the inertia of $\hat{K}_3$ remains the same in the limit as during the iterations is in contrast with the situation for the 1×1 and 2×2 formulations. In addition, Theorem 3.7 also holds if H is indefinite yet positive definite on the nullspace of J. This corresponds more closely to a typical case of nonconvex optimization in a neighborhood of an isolated minimizer, where H is positive definite over a subset of Null(J) only—a cone defined by the gradients of the active bounds.

3.3.3. Eigenvalue Bounds. We now turn to finding bounds on the eigenvalues of the symmetric indefinite matrix $\hat{K}_3$, which given the similarity transformation (3.5), apply also to K_3. Our technique largely relies on energy estimates in the spirit of Rusten and Winther (1992).

The eigenvalue problem for $\hat{K}_3$ is formulated as

$$\begin{bmatrix} H & J^T & -Z^{\frac{1}{2}} \\ J & 0 & 0 \\ -Z^{\frac{1}{2}} & 0 & -X \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \theta \begin{bmatrix} u \\ v \\ w \end{bmatrix}. \tag{3.8}$$

For any vectors $u \in \mathbb{R}^n$ and $v \in \text{Null}(J)^\perp \subset \mathbb{R}^n$, recall that

$$\lambda_n \|u\|^2 \leq u^T H u \leq \lambda_1 \|u\|^2, \tag{3.9}$$
$$\sigma_m \|v\| \ \leq \|Jv\| \ \ \leq \sigma_1 \|v\|. \tag{3.10}$$

Note that the right inequality in (3.10) is satisfied for all $v \in \mathbb{R}^n$.

THEOREM 3.8. *Suppose Assumption 2.1 holds. For as long as Assumption 2.2 holds and J has full rank, the positive eigenvalues of $\hat{K}_3$, and thus also K_3, are bounded in*

$$I_+ := \left[\min_j \tfrac{1}{2} \left(\lambda_n - x_j + \sqrt{(\lambda_n + x_j)^2 + 4z_j} \right), \ \tfrac{1}{2} \left(\lambda_1 + \sqrt{\lambda_1^2 + 4(\sigma_1^2 + z_{\max})} \right) \right],$$

where $z_{\max} := \max_i z_i$. When Assumption 2.3 holds, the lower bound reduces to $\lambda_n \geq 0$.

Proof. We proceed one bound at a time, using energy estimates and establishing the desired inequalities.

Upper bound on positive eigenvalues. Since we consider $\theta > 0$, $\theta I + X$ is nonsingular. We solve for w in the third row of (3.8) to get

$$w = -(\theta I + X)^{-1} Z^{\frac{1}{2}} u,$$

which we substitute into the first row of (3.8) to obtain

$$Hu + J^T v + Z^{\frac{1}{2}} (\theta I + X)^{-1} Z^{\frac{1}{2}} u = \theta u.$$

Taking the inner product with u, and noting that the matrices $Z^{\frac{1}{2}}$ and $(\theta I + X)^{-1}$ are diagonal and therefore commute, gives the following equation for θ:

$$u^T H u + u^T J^T v + u^T (\theta I + X)^{-1} Z u = \theta \|u\|^2. \tag{3.11}$$

Solving for v in the second row of (3.8) gives $v = \frac{1}{\theta} J u$, which we substitute into (3.11) to get

$$u^T H u + \tfrac{1}{\theta} \|Ju\|^2 + u^T (\theta I + X)^{-1} Z u = \theta \|u\|^2. \tag{3.12}$$

We use (3.9) and (3.10) to bound the first and second terms in (3.12):

$$\lambda_1 \|u\|^2 + \tfrac{1}{\theta} \sigma_1^2 \|u\|^2 + u^T (\theta I + X)^{-1} Z u \geq \theta \|u\|^2.$$

Since $(\theta I + X)^{-1} Z$ is diagonal, we may bound the last term in the left-hand side above as follows:

$$\left(\lambda_1 + \frac{1}{\theta} \sigma_1^2 + \max_i \frac{z_i}{\theta + x_i} \right) \|u\|^2 \geq \theta \|u\|^2.$$

The term $\max_i \frac{z_i}{\theta + x_i}$ can be bounded by $\frac{z_{\max}}{\theta}$, where $z_{\max}$ indicates the maximum value of z_i at the current iterate. This bound becomes tight as the iterations proceed, since generally at least one $x_i \to 0$. On the other hand, if this does not occur, then all $z_i = 0$ at the limit and therefore $z_{\max} = 0$. In both situations, we have

$$\left(\lambda_1 + \frac{1}{\theta} \sigma_1^2 + \frac{z_{\max}}{\theta} \right) \|u\|^2 \geq \theta \|u\|^2.$$

Multiplying by θ and rearranging gives

$$\left(\theta^2 - \lambda_1\theta - (\sigma_1^2 + z_{\max})\right)\|u\|^2 \leq 0.$$

We must have $u \neq 0$, because if $u = 0$ then the second row of (3.6) implies $\theta v = 0$. Since K_3 is nonsingular, $\theta \neq 0$, so this gives $v = 0$. The first row then yields $Z^{\frac{1}{2}}w = 0$ and $w = 0$, which is a contradiction since an eigenvector must be nontrivial. We can therefore divide by $\|u\|^2$ and bound by the root of the quadratic to obtain the desired upper bound.

Lower bound on positive eigenvalues. Taking the inner product of v with the second row of (3.8),

$$v^T Ju = \theta\|v\|^2,$$

which we substitute into (3.11) to give

$$u^T Hu + \theta\|v\|^2 + u^T(\theta I + X)^{-1}Zu = \theta\|u\|^2.$$

Using (3.9), we have

$$\lambda_n\|u\|^2 + \theta\|v\|^2 + u^T(\theta I + X)^{-1}Zu \leq \theta\|u\|^2.$$

Bounding the last term on the left with a minimum, this becomes

$$\lambda_n\|u\|^2 + \theta\|v\|^2 + \min_i \frac{z_i}{\theta + x_i}\|u\|^2 \leq \theta\|u\|^2. \tag{3.13}$$

We denote the index where the minimum occurs by j, then multiply by $\theta + x_j$ and rearrange into

$$(\theta^2 + (x_j - \lambda_n)\theta - (\lambda_n x_j + z_j))\|u\|^2 \geq \theta(\theta + x_j)\|v\|^2 \geq 0.$$

Since again $u \neq 0$, we then bound by the positive root of the quadratic. Taking the minimum over $j \in \mathcal{I}$ gives the bound.

To find a uniform lower bound we begin from (3.13) and use the fact that $\min_i z_i/(\theta + x_i) \geq 0$. Upon rearranging we have

$$(\theta - \lambda_n)\|u\|^2 \geq \theta\|v\|^2 \geq 0,$$

which yields the desired bound because u is nonzero. □

Note that the lower bound on positive eigenvalues in Theorem 3.8 is strictly larger than λ_n as long as Assumption 2.2 holds. In a sufficiently small neighborhood of an isolated minimizer, the minimum term in the lower bound will be attained for some $j \in \mathcal{I}$. This lower bound is strictly positive as long as $(x, z) > 0$ but in the limit, by definition of $\mathcal{I}$, it reduces to λ_n. If $\lambda_n = 0$, this limiting bound is overly pessimistic and is not tight if the LICQ is satisfied since K_3 is nonsingular.

Next, we consider bounds on the negative eigenvalues. We are only able to find an effective lower bound.

Theorem 3.9. *Suppose Assumption 2.1 holds. Suppose the matrix $\theta I + X$ is nonsingular for all $\theta < 0$ in the spectrum of $\hat{K}_3$. The negative eigenvalues of $\hat{K}_3$, and thus also K_3, are bounded in $I_- = [\zeta, 0)$, where*

$$\zeta := \min \left\{ \tfrac{1}{2} \left(\lambda_n - \sqrt{\lambda_n^2 + 4\sigma_1^2} \right), \min_{\{j \mid \theta + x_j < 0\}} \theta_j^* \right\}$$

and θ_j^* is the smallest negative root of the cubic equation

$$\theta^3 + (x_j - \lambda_n)\theta^2 - (\sigma_1^2 + z_j + x_j\lambda_n)\theta - \sigma_1^2 x_j = 0.$$

Proof. Proceeding as in the proof of Theorem 3.8, we start from (3.12) with the bounds in (3.9) and (3.10) to get

$$\lambda_n \|u\|^2 + \tfrac{1}{\theta}\sigma_1^2\|u\|^2 + u^T(\theta I + X)^{-1} Z u \le \theta \|u\|^2.$$

Bounding the last term of the left-hand side by the minimum, we obtain

$$\left(\lambda_n + \frac{1}{\theta}\sigma_1^2 + \min_i \frac{z_i}{\theta + x_i} \right) \|u\|^2 \le \theta \|u\|^2. \tag{3.14}$$

We now need to consider two cases. In case one, $\theta + x_i > 0$ for all indices i, which holds during early iterations, as well as in the situation where all the $x_i = \mathcal{O}(1)$ in the limit. In this case we can bound the minimum term from below by zero. In case two, some $\theta + x_i < 0$, and there exists an index j such that $\min_i \frac{z_i}{\theta + x_i} = \frac{z_j}{\theta + x_j}$.

Case one: Starting from (3.14), we bound the minimum term below by zero, giving

$$(\lambda_n + \tfrac{1}{\theta}\sigma_1^2)\|u\|^2 \le \theta\|u\|^2,$$

which we can multiply by θ and rearrange to give

$$(\theta^2 - \lambda_n\theta - \sigma_1^2)\|u\|^2 \le 0.$$

Since $u \neq 0$, we can divide by $\|u\|^2$ and bound by the root to get

$$\theta \ge \tfrac{1}{2}\left(\lambda_n - \sqrt{\lambda_n^2 + 4\sigma_1^2} \right).$$

Case two: We use the index j for the minimum in (3.14) to get

$$\left(\lambda_n + \frac{1}{\theta}\sigma_1^2 + \frac{z_j}{\theta + x_j} \right) \|u\|^2 \le \theta\|u\|^2.$$

Multiplying by $\theta(\theta + x_j) > 0$ and rearranging, we get

$$\left(\theta^3 + (x_j - \lambda_n)\theta^2 - (\sigma_1^2 + z_j + x_j\lambda_n)\theta - \sigma_1^2 x_j \right) \|u\|^2 \ge 0.$$

Since $u \neq 0$, we can divide by $\|u\|^2$ and define θ_j^* to be the smallest root of the cubic. There is exactly one positive real root, since the values of the cubic and its derivative are negative at zero, so there are either two or zero negative roots. If no cubic offers a real negative root, this implies that case one applies. The bound is given by the minimum of these possible bounds from cases one and two. ∎

Let us provide some insight on the negative bounds. First, regarding Theorem 3.9, in practice we cannot know which indices j satisfy $\theta + x_j < 0$, but we can simply compute θ_j^* for all indices j and use this in the comparison. Another observation is that in the proof of the theorem, the possibility that some $\theta < 0$ in the spectrum of $\hat{K}_3$ be an eigenvalue of $-X$ could arise in the course of the iterations or in the limit if there are inactive bounds.

Finally, we remark on the unsatisfactory situation with regard to the upper bound on negative eigenvalues. We have not been able to find a bound strictly smaller than 0. However, the bounds of Theorems 3.8 and 3.9 are pessimistic. Under our regularity and strict complementarity assumptions, as stated in Theorems 3.4 and 3.5, $\hat{K}_3$ is nonsingular and converges to a well-defined limit as $\mu \to 0$. Therefore, its condition number is asymptotically uniformly bounded, which is not reflected by the bounds on the eigenvalues. It is therefore difficult to give asymptotic estimates in this case. From Theorem 3.9, the lower bound on negative eigenvalues remains finite asymptotically, while Theorem 3.8 shows that the asymptotic bounds on the positive eigenvalues essentially reduce to those of K_2. In practice, however, we have observed that the condition number of $\hat{K}_3$ is typically substantially better than that of K_1 or K_2 and slightly, but consistently, better than that of K_3.

In the next section we use regularization and are able to remove our regularity assumptions and reach more definitive conclusions regarding the upper bound on the negative eigenvalues and asymptotic condition number estimates.

4. Analysis of a Regularized System. Despite the advantages that the 3×3 block formulation has over the reduced formulations, numerical difficulties still creep up in situations where J does not have full row rank, the LICQ is not satisfied and/or strict complementarity is not satisfied in the limit.

One way to alleviate some of those difficulties is by introducing regularization parameters (Saunders, 1996). There are various options here, and we focus on a two-parameter regularization approach, aimed at taking eigenvalues of the Hessian and singular values of the Jacobian away from zero. See, e.g., Gondzio (2012) for a similar approach. We introduce parameters $\rho > 0$ and $\delta > 0$, and consider a regularized primal QP problem of the form proposed by Friedlander and Orban (2012):

$$\begin{aligned} \underset{x,r}{\text{minimize}} \quad & c^T x + \tfrac{1}{2} x^T H x + \tfrac{1}{2}\rho \|x - x_k\|^2 + \tfrac{1}{2}\delta \|r + y_k\|^2 \\ \text{subject to} \quad & Jx + \delta r = b, \ x \geq 0. \end{aligned} \tag{4.1}$$

Here x_k and y_k are current primal and dual approximations, respectively. The corresponding dual problem is given as

$$\begin{aligned} \underset{x,y,z,s}{\text{maximize}} \quad & b^T y - \tfrac{1}{2} x^T H x - \tfrac{1}{2}\delta \|y - y_k\|^2 - \tfrac{1}{2}\rho \|s + x_k\|^2 \\ \text{subject to} \quad & -Hx + J^T y + z - \rho s = c, \ z \geq 0. \end{aligned} \tag{4.2}$$

Note that setting $\delta = \rho = 0$ recovers the original primal-dual pair (1.1a)–(1.1b).

Friedlander and Orban (2012) propose an interior-point method for (4.1)–(4.2) that converges under standard conditions with either fixed or decreasing values of the regularization parameters, without assumptions on the rank of J. The linear systems associated with (4.1)–(4.2) involve a modified version of (2.2):

$$K_{3,\text{reg}} := \begin{bmatrix} H + \rho I & J^T & -I \\ J & -\delta I & 0 \\ -Z & 0 & -X \end{bmatrix}. \tag{4.3}$$

Upon reduction, the 2×2 matrix reads

$$K_{2,\mathrm{reg}} := \begin{bmatrix} H + X^{-1}Z + \rho I & J^T \\ J & -\delta I \end{bmatrix}. \tag{4.4}$$

Finally, the Schur complement equations now have the matrix

$$K_{1,\mathrm{reg}} := J(H + X^{-1}Z + \rho I)^{-1} J^T + \delta I. \tag{4.5}$$

We examine each in turn in the following subsections.

4.1. The Regularized Schur Complement Equations. Whenever $\delta > 0$, $K_{1,\mathrm{reg}}$ is unconditionally positive definite. An eigenvalue analysis similar to that of §3.1 yields the following result.

LEMMA 4.1. *Let $K_{1,\mathrm{reg}}$ be defined as in (4.5) and suppose Assumptions 2.1 and 2.2 hold. The eigenvalues of $K_{1,\mathrm{reg}}$ are contained in the interval*

$$\left[\frac{\sigma_m^2}{\lambda_{\max}(H + X^{-1}Z + \rho I)} + \delta,\ \frac{\sigma_1^2}{\lambda_{\min}(H + X^{-1}Z + \rho I)} + \delta \right].$$

As a consequence, we have the following bound on the spectral condition number:

$$\kappa(K_{1,\mathrm{reg}}) \le \frac{\sigma_1^2 + \delta \lambda_{\min}(H + X^{-1}Z + \rho I)}{\sigma_m^2 + \delta \lambda_{\max}(H + X^{-1}Z + \rho I)}\ \kappa(H + X^{-1}Z + \rho I).$$

Similar bounds were derived by Gondzio (2012). Contrary to Lemma 3.1, no clear relation readily emerges from the bound on the condition number given in Lemma 4.1. However, it is possible to see that if both ρ and δ are positive, the condition number is strictly smaller than that of the unregularized matrix. Asymptotically, in the scenario of Example 2.1, the bound reduces to

$$\kappa(K_{1,\mathrm{reg}}) = O(\sigma_1^2/(\rho\delta)).$$

In practice, ρ and δ are allowed to take values as small as about $\sqrt{\varepsilon_{\mathrm{mach}}}$. In this case, there appears to be a definite disadvantage to using the Schur complement equations because the condition number likely exceeds the inverse of machine precision early. Indeed, the implementation of Friedlander and Orban (2012) initializes $\rho = \delta = 1$ and divides both by 10 at each iteration. In IEEE double precision, after just 8 iterations, the smallest allowed value of 10^{-8} is reached but convergence is typically not yet achieved.

4.2. The Regularized 2 × 2 Block System. Results for the regularized K_2 are given in (Friedlander and Orban, 2012, Theorem 5.1), which is itself a specialization of results of Rusten and Winther (1992) and Silvester and Wathen (1994). In the analysis it is not assumed that J has full rank, only that $\rho > 0$ and $\delta > 0$.

First, it is not difficult to show that regularizing does not affect the inertia.

LEMMA 4.2 (Friedlander and Orban (2012)). *If Assumptions 2.1 and 2.2 hold, the inertia of $K_{2,\mathrm{reg}}$ is $(n, m, 0)$.*

The next result parallels Lemma 3.3.

LEMMA 4.3 (Friedlander and Orban (2012)). *Suppose Assumptions 2.1 and 2.2 hold, and let $H_\rho := H + X^{-1}Z + \rho I$. Then,*

$$\lambda^+ \geq \lambda_{\min}(H_\rho),$$
$$\lambda^+ \leq \tfrac{1}{2}\left((\lambda_{\max}(H_\rho) - \delta) + \sqrt{(\lambda_{\max}(H_\rho) - \delta)^2 + 4(\sigma_1^2 + \delta\lambda_{\max}(H_\rho))} \right)$$

for any positive eigenvalue λ^+ of $K_{2,\mathrm{reg}}$, and

$$\lambda^- \geq \tfrac{1}{2}\left((\lambda_{\min}(H_\rho) - \delta) - \sqrt{(\lambda_{\min}(H_\rho) - \delta)^2 + 4(\sigma_1^2 + \delta\lambda_{\min}(H_\rho))} \right),$$
$$\lambda^- \leq \tfrac{1}{2}\left((\lambda_{\max}(H_\rho) - \delta) - \sqrt{(\lambda_{\max}(H_\rho) - \delta)^2 + 4(\sigma_m^2 + \delta\lambda_{\max}(H_\rho))} \right)$$

for any negative eigenvalue λ^- of $K_{2,\mathrm{reg}}$. In addition, $-\delta$ is an eigenvalue of $K_{2,\mathrm{reg}}$ if and only if J is rank deficient.

Again, similar bounds were derived by Gondzio (2012). Note that whether or not J is rank deficient, Lemma 4.3 implies that all negative eigenvalues of $K_{2,\mathrm{reg}}$ are bounded above by $-\delta$. The most noticeable effect of regularization in this case is to buffer the eigenvalues away from zero.

In the scenario of Example 2.1, Lemma 4.3 yields the following asymptotic bounds:

$$-\sigma_1 \lesssim \lambda^- \leq -\delta < 0 < \rho \leq \lambda^+ \lesssim \lambda_{\max}(H_\rho)$$

so that we obtain the following asymptotic condition number estimate:

$$\kappa(K_{2,\mathrm{reg}}) = O(\lambda_{\max}(H_\rho)/\min(\rho,\delta)) = O(1/(\mu \min(\rho,\delta))).$$

The limits of machine precision, given the common bounds on δ and ρ, are thus not achieved until μ reaches $\sqrt{\varepsilon_{\mathrm{mach}}}$, which typically occurs in the last few iterations.

4.3. The Regularized 3×3 Block System. We now consider the unreduced 3×3 block system, and introduce several new analytical results.

4.3.1. Nonsingularity of $\mathbf{K_{3,\mathrm{reg}}}$. We begin with conditions for $K_{3,\mathrm{reg}}$ to be nonsingular throughout the interior-point iteration. The next result covers also the unregularized setting discussed in §3.3.1, namely $\delta = \rho = 0$.

THEOREM 4.4. *Let $\rho \geq 0$ and suppose Assumptions 2.1 and 2.2 hold. The matrix $K_{3,\mathrm{reg}}$ in (4.3) is nonsingular if and only if either (i) $\delta > 0$, or (ii) $\delta = 0$ and J has full rank.*

Proof. Consider the case *(i)* where $\delta > 0$. Looking at the system

$$\begin{bmatrix} H + \rho I & J^T & -I \\ J & -\delta I & 0 \\ -Z & 0 & -X \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \tag{4.6}$$

we attempt to find when a nontrivial solution exists. From the third block row, we have $w = -X^{-1}Zu$. Taking the inner product of the first block row with u and substituting for w yields $u^T(H + \rho I + X^{-1}Z)u + (Ju)^T v = 0$, which simplifies, using $v = \frac{1}{\delta}Ju$ from the second block row, to

$$u^T\left((H + \rho I + X^{-1}Z) + \tfrac{1}{\delta}J^T J\right)u = 0.$$

Since X and Z are diagonal and positive definite, the matrix associated with the above equation is positive definite for any $\rho \geq 0$. Therefore, the only solution is $u = 0$, which implies $v = 0$ and $w = 0$.

Consider now the case *(ii)* where $\delta = 0$. Assume that (u, v, w) lies in the nullspace of $K_{3,\mathrm{reg}}$. The only difference from the proof of the first part is that $Ju = 0$. Following the same steps, we obtain

$$u^T \left(H + \rho I + X^{-1} Z\right) u = 0.$$

Again, $H + \rho I + X^{-1} Z$ is positive definite. Thus the only solution is $u = 0$, which implies that $w = 0$. The first block row gives $J^T v = 0$, implying $v = 0$, due to J having full rank. Therefore the full rank condition for J is a sufficient condition for nonsingularity of $K_{3,\mathrm{reg}}$.

If J does not have full row rank, then (u, v, w) with $0 \neq v \in \mathrm{Null}(J^T)$, $u = 0$, $w = 0$ is a nontrivial nullspace vector. Therefore the full-rank condition on J is both necessary and sufficient. □

We now consider what happens to $K_{3,\mathrm{reg}}$ in the limit of the interior-point iteration. If (x, y, z) is not strictly complementary, there is a zero row in the third block row of (4.3) and $K_{3,\mathrm{reg}}$ is singular. Thus strict complementarity is necessary for nonsingularity in each case. The theorem includes the unregularized case, $\delta = \rho = 0$, which is covered also by Theorem 3.5.

<table><tr><td>

Theorem 4.5. *Suppose Assumptions 2.1 and 2.3 hold at (x, y, z). Necessary and sufficient conditions for the matrix $K_{3,\mathrm{reg}}$ to be nonsingular are that (x, y, z) be strictly complementary,* $\mathrm{Null}(H) \cap \mathrm{Null}(J) \cap \mathrm{Null}(Z) \neq \{0\}$ *if $\rho = 0$, and the LICQ be satisfied if $\delta = 0$.*

</td></tr></table>

Proof. If both $\rho = 0$ and $\delta = 0$, Theorem 3.5 establishes the result.

Consider first the case where $\rho > 0$ and $\delta > 0$. The third block row of (4.6) yields, componentwise, $u_{\mathcal{A}} = 0$ and $w_{\mathcal{I}} = 0$, and therefore, $u^T w = 0$. We take the inner product of the first block row with u, substitute $v = \frac{1}{\delta} Ju$ from the second block row and use $u^T w = 0$ to get

$$u^T \left(H + \rho I + \tfrac{1}{\delta} J^T J\right) u = 0.$$

Since $\rho > 0$, the matrix in the above equation is positive definite. Therefore, $u = 0$. It follows that $v = 0$, and the first block row gives $w = 0$. Therefore $K_{3,\mathrm{reg}}$ is nonsingular, and strict complementarity is sufficient.

Consider now the case where $\rho > 0$ and $\delta = 0$. The only difference from the previous case is that $Ju = 0$. Following the same steps, we obtain $u^T (H + \rho I) u = 0$. Since $H + \rho I$ is positive definite, the only solution is $u = 0$. Examining the remaining equations, we have

$$\begin{bmatrix} J^T & -I_{\mathcal{A}} \end{bmatrix} \begin{bmatrix} v \\ w_{\mathcal{A}} \end{bmatrix} = 0,$$

which has only the trivial solution if and only if the LICQ holds. Thus these two conditions together are sufficient, and the LICQ is necessary.

Consider finally the case where $\rho = 0$ and $\delta > 0$. Following the same steps as in the first two cases, taking an inner product with u, we obtain

$$u^T \left(H + \tfrac{1}{\delta} J^T J\right) u = 0.$$

The last matrix is positive semidefinite, and thus we must have $u \in \text{Null}(H) \cap \text{Null}(J)$. Since $u_{\mathcal{A}} = 0$, we also have $u \in \text{Null}(Z)$. Thus $u = 0$, and the second block row gives $v = 0$. The first block row leaves $w = 0$, so the only solution is the trivial solution. Thus, the conditions $\text{Null}(H) \cap \text{Null}(J) \cap \text{Null}(Z) = \{0\}$ together with strict complementarity are sufficient for nonsingularity of $K_{3,\text{reg}}$.

Assume now that $\text{Null}(H) \cap \text{Null}(J) \cap \text{Null}(Z) \neq \{0\}$. Then for any nonzero $u \in \text{Null}(H) \cap \text{Null}(J) \cap \text{Null}(Z)$, we have a nontrivial nullspace vector of the form $(u, 0, 0)$. Therefore the condition is both necessary and sufficient. ☐

4.3.2. Inertia. Similarly to §3.3.2, we first show that $K_{3,\text{reg}}$ is symmetrizable, and generate the symmetric matrix $\hat{K}_{3,\text{reg}}$. This allows us to consider bounds in real arithmetic for the latter, which will also apply to $K_{3,\text{reg}}$. The matrix $\hat{K}_{3,\text{reg}}$ is readily obtained by the same diagonal symmetrizer D defined and used earlier in (3.5):

$$\begin{bmatrix} H+\rho I & J^T & -Z^{\frac{1}{2}} \\ J & -\delta I & 0 \\ -Z^{\frac{1}{2}} & 0 & -X \end{bmatrix}. \tag{4.7}$$

Proceeding as in the unregularized case, we first state the inertia of $\hat{K}_{3,\text{reg}}$ and provide eigenvalue bounds next.

We find results for the inertia of $\hat{K}_{3,\text{reg}}$, showing that it is the same as in the unregularized case both during the iterations and in the limit.

THEOREM 4.6. *For $\rho \geq 0$, $\delta > 0$, assume $H + \rho I$ is positive definite. Suppose that*

1. *Assumption 2.2 holds, or*
2. *Assumption 2.3 holds at (x, y, z), where strict complementarity is satisfied.*

Then the inertia of $\hat{K}_{3,\text{reg}}$ is $(n, n + m, 0)$.

Proof. Consider first the case where Assumption 2.2 holds. Define

$$M := \begin{bmatrix} I & & \\ A & I & \\ B & C & I \end{bmatrix},$$

with $A := -J(H + \rho I)^{-1}, C := Z^{\frac{1}{2}}(H + \rho I)^{-1}J^T \left(J(H + \rho I)^{-1}J^T + \delta I\right)^{-1}$, and $B := (Z^{\frac{1}{2}} - CJ)(H + \rho I)^{-1}$, and consider the congruence

$$M\hat{K}_{3,\text{reg}}M^T = \begin{bmatrix} H+\rho I & & \\ & U & \\ & & W \end{bmatrix}, \tag{4.8}$$

where $U := -\delta I - J(H + \rho I)^{-1}J^T$ and $W := -BZ^{1/2} - X$. We now show that W is negative definite, and then since $H + \rho I$ is positive definite and U is negative definite, we will have that the inertia of $\hat{K}_{3,\text{reg}}$ is $(n, n + m, 0)$.

Inserting the definitions of B and C and defining $\tilde{H} = H + \rho I$ for simplicity, W is given by

$$-Z^{1/2}\tilde{H}^{-1}Z^{1/2} + Z^{1/2}\tilde{H}^{-1}J^T(J\tilde{H}^{-1}J^T + \delta I)^{-1}J\tilde{H}^{-1}Z^{1/2} - X.$$

Factoring the first two terms, we have

$$-Z^{1/2}\tilde{H}^{-1/2}\left[I - \tilde{H}^{-1/2}J^T(J\tilde{H}^{-1}J^T + \delta I)^{-1}J\tilde{H}^{-1/2}\right]\tilde{H}^{-1/2}Z^{1/2} - X.$$

Now defining $P = \tilde{H}^{-1/2} J^T$, we have

$$-Z^{1/2} \tilde{H}^{-1/2} \left[I - P(P^T P + \delta I)^{-1} P^T \right] \tilde{H}^{-1/2} Z^{1/2} - X.$$

When $\delta = 0$, the matrix $I - P(P^T P + \delta I)^{-1} P^T$ is an orthogonal projector, with eigenvalues 0 and 1, and when $\delta > 0$, the matrix $(P^T P + \delta I)^{-1}$ is a contraction. Thus $I - P(P^T P + \delta I)^{-1} P^T$ has eigenvalues in $(0, 1]$, and is strictly positive definite for $\delta > 0$. Thus W is negative definite, completing the proof.

Consider now the case where Assumption 2.3 holds. The relation (4.8) still holds. By Theorem 4.5, $\hat{K}_{3,\text{reg}}$ is nonsingular, and therefore since M is nonsingular, the block diagonal matrix (4.8) is nonsingular as well. Then $H + \rho I$ is positive definite and U and W are negative definite, and the inertia of $\hat{K}_{3,\text{reg}}$ is $(n, n+m, 0)$. ◻

4.3.3. Eigenvalue Bounds. The eigenvalue problem for $\hat{K}_{3,\text{reg}}$ is formulated as

$$\begin{bmatrix} H + \rho I & J^T & -Z^{\frac{1}{2}} \\ J & -\delta I & 0 \\ -Z^{\frac{1}{2}} & 0 & -X \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \theta \begin{bmatrix} u \\ v \\ w \end{bmatrix}. \tag{4.9}$$

Our first result provides upper and lower bounds on the positive eigenvalues of $\hat{K}_{3,\text{reg}}$. Most notably, we show that upon regularization the lower bound is approximately additively shifted by ρ.

Theorem 4.7. *Suppose Assumption 2.1 is satisfied. As long as Assumption 2.2 is satisfied, the positive eigenvalues of $\hat{K}_{3,reg}$ are bounded in $I_+ = [\xi, \eta]$, where*

$$\xi = \min_j \tfrac{1}{2} \left(\lambda_n + \rho - x_j + \sqrt{(\lambda_n + \rho + x_j)^2 + 4 z_j} \right)$$

and η is the largest root of the cubic equation

$$\theta^3 + (\delta - (\lambda_1 + \rho))\theta^2 - (\delta(\lambda_1 + \rho) + \sigma_1^2 + z_{\max})\theta - z_{\max}\delta = 0.$$

When Assumption 2.3 holds, ξ reduces to $\lambda_n + \rho$.

Proof. As before, we examine the upper and lower bounds separately.

Upper bound on the positive eigenvalues. We solve for w in the third block row of (4.9) to get $w = -(\theta I + X)^{-1} Z^{\frac{1}{2}} u$, which we substitute into the first block row to obtain

$$(H + \rho I)u + J^T v + Z^{\frac{1}{2}} (\theta I + X)^{-1} Z^{\frac{1}{2}} u = \theta u.$$

Taking the inner product with u and noting that the matrices $Z^{\frac{1}{2}}$ and $(\theta I + X)^{-1}$ are diagonal and therefore commute, gives the following equation for θ:

$$u^T (H + \rho I) u + u^T J^T v + u^T (\theta I + X)^{-1} Z u = \theta \|u\|^2. \tag{4.10}$$

Solving for v in the second block row of (4.9) gives $v = \frac{1}{\theta + \delta} J u$, which we substitute into (4.10) to get

$$u^T (H + \rho I) u + \frac{1}{\theta + \delta} \|J u\|^2 + u^T (\theta I + X)^{-1} Z u = \theta \|u\|^2. \tag{4.11}$$

We use (3.9) and (3.10) to bound the first and second terms in (4.11):

$$(\lambda_1+\rho)\|u\|^2+\frac{\sigma_1^2}{\theta+\delta}\|u\|^2+u^T(\theta I+X)^{-1}Zu\geq\theta\|u\|^2.$$

Since $(\theta I+X)^{-1}Z$ is diagonal, we have

$$u^T(\theta I+X)^{-1}Zu\leq\max_i\frac{z_i}{\theta+x_i}\|u\|^2.$$

We must have $u\neq 0$, since if $u=0$ then the second row of (4.9) implies that $(\theta+\delta)v=0$. Since the matrix (4.3) is nonsingular, $\theta>0$ and thus $\theta+\delta>0$, implying that $v=0$. Then, by the first row of (4.9), $Z^{\frac{1}{2}}w=0$ would imply $w=0$, which must not occur since an eigenvector must be nontrivial. We can thus divide by $\|u\|^2$, and using the last two displayed inequalities we get the relation

$$\lambda_1+\rho+\frac{\sigma_1^2}{\theta+\delta}+\max_i\frac{z_i}{\theta+x_i}\geq\theta.$$

As in the proof for Theorem 3.8, we can bound the maximum term above by $\frac{z_{\max}}{\theta}$:

$$\lambda_1+\rho+\frac{1}{\theta+\delta}\sigma_1^2+\frac{z_{\max}}{\theta}\geq\theta,$$

Multiplying by $(\theta+\delta)\theta$ and rearranging gives

$$\theta^3+(\delta-(\lambda_1+\rho))\,\theta^2-\left(\delta(\lambda_1+\rho)+\sigma_1^2+z_{\max}\right)\theta-z_{\max}\delta\leq 0.$$

As a consequence, θ must be bounded above by the largest real root of the above cubic polynomial. Note that there is exactly one positive real root, since the values of the cubic and its derivative are negative at zero.

Lower bound on the positive eigenvalues. Taking the inner product of v with the second row of (4.9) and rearranging, we have

$$v^TJu=(\theta+\delta)\|v\|^2,$$

which we substitute into (4.10) to give

$$u^T(H+\rho I)u+(\theta+\delta)\|v\|^2+u^T(\theta I+X)^{-1}Zu=\theta\|u\|^2.$$

Then using (3.9), we have

$$(\lambda_n+\rho)\|u\|^2+(\theta+\delta)\|v\|^2+u^T(\theta I+X)^{-1}Zu\leq\theta\|u\|^2.$$

The last term in the left-hand side may be bounded below with a minimum:

$$(\lambda_n+\rho)\|u\|^2+(\theta+\delta)\|v\|^2+\min_i\frac{z_i}{\theta+x_i}\|u\|^2\leq\theta\|u\|^2.$$

We denote the index where the minimum occurs by j, then multiply by $\theta+x_j$ and rearrange into

$$(\theta^2+(x_j-\lambda_n-\rho)\theta-(x_j(\lambda_n+\rho)+z_j))\|u\|^2\geq(\theta+\delta)(\theta+x_j)\|v\|^2\geq 0.$$

Since again $u\neq 0$, we then bound by the positive root of the quadratic. Taking the minimum over all indices j gives the desired bound.

If Assumption 2.3 holds, we use the fact that $\min_i z_i/(\theta + x_i) \geq 0$ and obtain

$$(\theta - \lambda_n - \rho)\|u\|^2 \geq (\theta + \delta)\|v\|^2 \geq 0.$$

The lower bound becomes $\theta \geq \lambda_n + \rho$. □

Note that in the special case where all z_i converge to 0, and thus $z_{\max} = 0$ at the limit, the upper bound at the limit can be written as an explicit quadratic root independent of z_i:

$$\theta \leq \frac{1}{2}\left(\lambda_1 + \rho - \delta + \sqrt{(\delta - \lambda_1 - \rho)^2 + 4\left(\delta(\lambda_1 + \rho) + \sigma_1^2\right)}\right).$$

This parallels the form of Lemma 3.3 and Theorem 3.8.

We begin our investigation of negative eigenvalues with an upper bound, which turns out to depend on the scaling of the problem.

LEMMA 4.8. *Let $\rho \geq 0$, $\delta > 0$ and Assumption 2.1 be satisfied. Suppose also that Assumption 2.2 holds at (x, y, z) where $x_i > \delta$ for $i = 1, \dots, n$. Then the negative eigenvalues of $\hat{K}_{3,\mathrm{reg}}$ are bounded above by $-\delta$, and $\theta = -\delta$ is an eigenvalue if and only if J is rank deficient.*

Proof. We first show that $-\delta$ is an upper bound. Assume by contradiction that there is a negative eigenvalue that satisfies $\theta > -\delta$. Upon extracting $v = \frac{1}{\theta+\delta} Ju$ from the second block row of (4.9) and using the identity $w^T Z^{\frac{1}{2}} u = -w^T(\theta I + X)w$ from the third block row, taking the inner product of the first block row with u gives

$$u^T(H + \rho I)u + \tfrac{1}{\theta+\delta}\|Ju\|^2 + w^T(\theta I + X)w = \theta\|u\|^2.$$

Since $\theta+\delta > 0$ by assumption and all $x_i > \delta > -\theta$, the left-hand side of the last identity is positive. If $u = 0$, then both v and w are also zero, giving a trivial eigenvector and therefore a contradiction. If $u \neq 0$, θ must be positive, which contradicts our initial assumption on the sign of θ. Thus the negative eigenvalues are bounded above by $-\delta$.

We now move on to show when $-\delta$ is an eigenvalue. If J is rank deficient, then $u = 0$, $0 \neq v \in \mathrm{Null}(J^T)$, and $w = 0$ satisfies (4.9) with $\theta = -\delta$. Suppose now that J has full rank. We will show that $\theta \neq -\delta$. By contradiction, assume that $\theta = -\delta$. From the third block row and the assumption that all $x_i > \delta$, we have

$$w^T Z^{\frac{1}{2}} u = w^T(\delta I - X)w \leq 0.$$

Taking the inner product of the first block row of (4.9) with u and using the above inequality and $Ju = 0$ from the second block row, we obtain

$$-\delta\|u\|^2 = u^T(H + \rho I)u - u^T Z^{\frac{1}{2}} w \geq 0.$$

Since $\delta > 0$, this must mean that $u = 0$. The third block row then gives $w = 0$ and we are left with $J^T v = 0$ in the first block row. Since J has full row rank, we conclude that $v = 0$ and that $\theta = -\delta$ cannot be an eigenvalue. □

Interestingly, a similar result holds in the limit. Note however that there seems to be a transition zone between the moment where some components of x drop below δ and the limit when strict complementarity applies. This "gray zone" is necessarily attained if $\mathcal{A} \neq \emptyset$, and it is more difficult to characterize the relationship between θ and $-\delta$ in that zone. See, however, Remark 3. We note that the "gray zone" is strongly tied to the quality of our scaling assumptions; better scaling may shrink that zone.

LEMMA 4.9. *Let $\rho > 0$, $\delta > 0$ and Assumption 2.1 be satisfied. Suppose also Assumption 2.3 holds at (x, y, z), where strictly complementary is satisfied, $x_i > \delta$ for all $i \in \mathcal{I}$, and $\max_i \sqrt{z_i}$ is sufficiently small. Then the negative eigenvalues of $\hat{K}_{3,\mathrm{reg}}$ are bounded above by $-\delta$, and $\theta = -\delta$ is an eigenvalue if and only if J is rank deficient.*

Proof. We first show that $-\delta$ is an upper bound on the negative eigenvalues. Assume by contradiction that there exists a negative eigenvalue that satisfies $\theta > -\delta$. Since $\hat{K}_{3,\mathrm{reg}}$ is nonsingular, there must exist $\epsilon > 0$ such that $\theta \le -\epsilon$ for all negative eigenvalues. If $\delta \le \epsilon$ this implies $-\epsilon \le -\delta$ and the eigenvalues are bounded above by $-\delta$, which would be in line with the statement of the theorem. So let us assume that $\epsilon < \delta$. In the limit, we have $x_{\mathcal{A}} = 0$, $x_{\mathcal{I}} > 0$ and $z_{\mathcal{I}} = 0$. Because strict complementarity is satisfied, we must also have $z_{\mathcal{A}} > 0$. Partitioning the third block row of $\hat{K}_{3,\mathrm{reg}}$ in (4.9) into active and inactive components gives

$$-Z_{\mathcal{A}\mathcal{A}}^{\frac{1}{2}} u_{\mathcal{A}} = \theta w_{\mathcal{A}}, \tag{4.12}$$

$$-X_{\mathcal{I}\mathcal{I}} w_{\mathcal{I}} = \theta w_{\mathcal{I}}. \tag{4.13}$$

We may rewrite (4.13) as $(X_{\mathcal{I}\mathcal{I}} + \theta I) w_{\mathcal{I}} = 0$, which implies $w_{\mathcal{I}} = 0$ because $x_i + \theta > x_i - \delta > 0$ for all $i \in \mathcal{I}$ by assumption. Taking the inner product of both sides of (4.12) with $w_{\mathcal{A}}$ gives

$$-\, w_{\mathcal{A}}^T Z_{\mathcal{A}\mathcal{A}}^{\frac{1}{2}} u_{\mathcal{A}} = \theta \|w_{\mathcal{A}}\|^2. \tag{4.14}$$

Taking now the inner product of the first block row of (4.9) with u, the inner product of the second block row with v and combining them, we write

$$\theta \|u_{\mathcal{I}}\|^2 = u^T (H + \rho I) u + (\theta + \delta) \|v\|^2 + \theta (\|w_{\mathcal{A}}\|^2 - \|u_{\mathcal{A}}\|^2), \tag{4.15}$$

where we used the decomposition $\|u\|^2 = \|u_{\mathcal{A}}\|^2 + \|u_{\mathcal{I}}\|^2$ and (4.14). Note that from (4.12), $u_{\mathcal{A}} = 0$ if and only if $w_{\mathcal{A}} = 0$. If both vanish, necessarily $u_{\mathcal{I}} \ne 0$, and then the right-hand side of (4.15) is strictly positive. This would imply that $\theta > 0$, a contradiction. By our assumption that all $\sqrt{z_i}$ are sufficiently small, we suppose from this point on that $\max_i \sqrt{z_i} < \epsilon$. Suppose now that $u_{\mathcal{A}} \ne 0$ and $w_{\mathcal{A}} \ne 0$. Rearranging (4.12) we find that $w_{\mathcal{A}} = -\frac{1}{\theta} Z_{\mathcal{A}\mathcal{A}}^{\frac{1}{2}} u_{\mathcal{A}}$, and using the upper bounds $\max_i \sqrt{z_i} < \epsilon$ and $\theta \le -\epsilon$ we have $\|w_{\mathcal{A}}\|^2 \le \frac{\epsilon^2}{\theta^2} \|u_{\mathcal{A}}\|^2 \le \|u_{\mathcal{A}}\|^2$. Substituting into (4.15) gives

$$\theta \|u_{\mathcal{I}}\|^2 \ge u^T (H + \rho I) u + (\theta + \delta) \|v\|^2.$$

The right-hand side above is strictly positive, and if $u_{\mathcal{I}} \ne 0$ we have $\theta > 0$, a contradiction. If $u_{\mathcal{I}} = 0$, then $u_{\mathcal{A}} = 0$, again a contradiction. Therefore, we cannot have $\theta > -\delta$, and we conclude that $\theta \le -\delta$.

We now move on to find when $-\delta$ is an eigenvalue. If J is rank deficient, then as in Lemma 4.8, $(0, v, 0)$ with $0 \ne v \in \mathrm{Null}(J^T)$ is an eigenvector for $\theta = -\delta$. Now suppose that J has full rank, and assume by contradiction that $\theta = -\delta$. Partitioning as above gives

$$Z_{\mathcal{A}\mathcal{A}}^{\frac{1}{2}} u_{\mathcal{A}} = \delta w_{\mathcal{A}}, \tag{4.16}$$

$$X_{\mathcal{I}\mathcal{I}} w_{\mathcal{I}} = \delta w_{\mathcal{I}}. \tag{4.17}$$

Upon rearranging (4.17), the matrix $(X_{\mathcal{I}\mathcal{I}} - \delta I)$ has full rank by assumption, so we have $w_{\mathcal{I}} = 0$. If $u = 0$ we also have $w_{\mathcal{A}} = 0$ and there only remains $J^T v = 0$ in the first block row of (4.9), giving a contradiction because we obtain a trivial solution for the eigenvector. Thus $u \neq 0$. Taking the inner product of (4.16) with $w_{\mathcal{A}}$ reveals that

$$w^T Z^{\frac{1}{2}} u = w_{\mathcal{A}}^T Z_{\mathcal{A}\mathcal{A}}^{\frac{1}{2}} u_{\mathcal{A}} = \delta \|w_{\mathcal{A}}\|^2. \tag{4.18}$$

Using the Cauchy–Schwarz inequality and the bound on $\max_i \sqrt{z_i}$, we have

$$w_{\mathcal{A}}^T Z_{\mathcal{A}\mathcal{A}}^{\frac{1}{2}} u_{\mathcal{A}} \leq \epsilon \|w_{\mathcal{A}}\| \|u_{\mathcal{A}}\|,$$

and using (4.18) and rearranging gives

$$\|w_{\mathcal{A}}\| \leq \frac{\epsilon}{\delta} \|u_{\mathcal{A}}\| \leq \|u_{\mathcal{A}}\| \leq \|u\|, \tag{4.19}$$

since $\epsilon < \delta$. Taking the inner product of the first block row with u, using (4.18) and $Ju = 0$ from the second block row, and rearranging, we have

$$u^T (H + \rho I) u = \delta (\|w_{\mathcal{A}}\|^2 - \|u\|^2),$$

which reduces to $u^T (H + \rho I) u \leq 0$ by (4.19). Therefore $u = 0$, a contradiction, and hence $\theta \neq -\delta$ when J is full rank. □

REMARK 3. *By scaling the optimization problem prior to solving, it is possible (in theory) to arrange that the assumptions on scaling of x_i and z_i are satisfied. These results hold trivially when $\delta = 0$.*

Next, we derive a lower bound on the negative eigenvalues.

THEOREM 4.10. *Suppose Assumption 2.1 is satisfied. Assume $\theta I + X$ is nonsingular for all θ throughout the computation. Then the negative eigenvalues θ of the matrix $\hat{K}_{3,\mathrm{reg}}$ satisfy $\theta \geq \zeta$, where*

$$\zeta := \min \left\{ \tfrac{1}{2} \left(\lambda_n + \rho - \delta - \sqrt{(\lambda_n + \rho + \delta)^2 + 4\sigma_1^2} \right), \min_{\theta + x_j < 0} \theta_j^* \right\},$$

and θ_j^* is the smallest negative root of the cubic equation

$$\theta^3 + \left(x_j + \delta - \lambda_n - \rho \right) \theta^2 + \left(\delta x_j - \delta(\lambda_n + \rho) - x_j(\lambda_n + \rho) - \sigma_1^2 - z_j \right) \theta \\ - \left(\delta x_j (\lambda_n + \rho) + \sigma_1^2 x_j + z_j \delta \right) = 0.$$

Proof. We assume that $\theta + \delta < 0$. The case where $\theta \geq -\delta$ poses no difficulty because it is easy to verify that $\zeta \leq -\delta$.

We start from (4.11) with the bounds in (3.9) and (3.10) to get

$$(\lambda_n + \rho)\|u\|^2 + \frac{1}{\theta + \delta} \sigma_1^2 \|u\|^2 + u^T (\theta I + X)^{-1} Z u \leq \theta \|u\|^2.$$

We note that again $u \neq 0$, since if $u = 0$ the second row of (4.9) implies $(\theta + \delta) v = 0$, implying that $v = 0$. The first line yields $Z^{\frac{1}{2}} w = 0$ and thus $w = 0$, a contradiction. Bounding the last term of the left side of the previous inequality by the minimum,

$$\left(\lambda_n + \rho + \frac{1}{\theta + \delta} \sigma_1^2 + \min_i \frac{z_i}{\theta + x_i} \right) \leq \theta.$$

We consider two cases. In case one, $\theta + x_i > 0$ for all indices i, and in this case we can bound the minimum term from below by zero. In case two, some $\theta + x_i < 0$, and there exists an index j such that $\min_i \frac{z_i}{\theta + x_i} = \frac{z_j}{\theta + x_j}$.

Case one: We bound the minimum term below by zero in the inequality above,

$$\lambda_n + \rho + \frac{1}{\theta + \delta}\sigma_1^2 \le \theta, \tag{4.20}$$

which we can multiply by $(\theta + \delta)$ and rearrange to give

$$\theta^2 + (\delta - \lambda_n - \rho)\theta - (\delta(\lambda_n + \rho) + \sigma_1^2) \le 0.$$

Therefore,

$$\theta \ge \tfrac{1}{2}\left(\lambda_n + \rho - \delta - \sqrt{(\delta - \lambda_n - \rho)^2 + 4(\delta(\lambda_n + \rho) + \sigma_1^2)}\right).$$

Case two: We use the index j for the minimum to get

$$\lambda_n + \rho + \frac{1}{\theta + \delta}\sigma_1^2 + \frac{z_j}{\theta + x_j} \le \theta.$$

Multiplying by $(\theta + \delta)(\theta + x_j)$ and rearranging, we get

$$\begin{aligned}\theta^3 + (x_j + \delta - \lambda_n - \rho)\theta^2 + (\delta x_j - \delta(\lambda_n + \rho) - x_j(\lambda_n + \rho) - \sigma_1^2 - z_j)\theta \\ - (\delta x_j(\lambda_n + \rho) + \sigma_1^2 x_j + z_j\delta) \ge 0.\end{aligned}$$

We then define θ_j^* to be the smallest root of the cubic above. The bound is then given by the minimum of these possible bounds from cases one and two. □

Consider now the scenario of Example 2.1. Assuming that the problem has been scaled appropriately, the bounds of the previous results simplify to

$$\zeta \le \theta \le -\delta < 0 \quad \text{or} \quad 0 < \lambda_n + \rho \le \theta \le \eta$$

where ζ and η are both finite. Thus we obtain the asymptotic condition number estimate

$$\kappa(K_{3,\text{reg}}) \le \max(\eta, -\zeta)/\min(\rho + \lambda_n, \delta) = O\left(1/\min(\rho + \lambda_n, \delta)\right).$$

Here we have the validation of our claim that the block 3×3 system sees the best conditioning. Under the usual choices of ρ and δ and unless the conditioning of the problem is such that η or ζ is very large, this condition number will remain within computing limits through convergence of the iteration. Our numerical experiments, presented in §5, verify that the 3×3 matrices remain numerically nonsingular—and reasonably conditioned—throughout.

5. Numerical Experiments. We offer a few examples to validate the analysis of previous sections using a basic MATLAB implementation of Mehrotra's predictor-corrector procedure with an initial point computed as proposed by Friedlander and Orban (2012). Linear systems are solved using MATLAB's backslash operator, which performs a factorization. Eigenvalues, singular values, and condition numbers are computed with MATLAB's built-in functions. Our code is not intended to rival state-of-the-art solvers, but simply to illustrate the analysis in previous sections.

We illustrate the eigenvalue bounds for the traditional systems considered in §3 and the regularized systems of §4. We use two small convex QPs from the TOMLAB[1] optimization software package, problems number 6 and 21, both converted to standard form. Problem 6 has 10 variables, 7 constraints, is strictly convex and its solution satisfies strict complementarity and the LICQ. Problem 21 has 51 variables and 27 constraints, and the Hessian is positive semidefinite whose rank is 3. The solution identified by our interior-point method satisfies strict complementarity but not the LICQ. Those two problems are very small but are nonetheless representative of the situation for several other problems we have tested. The bounds for the regularized systems are illustrated using $\rho = \delta = 10^{-4}$ throughout the interior-point iterations.

We begin with Problem 21. Figure 5.1 shows the eigenvalues of K_1 and $K_{1,\text{reg}}$ on a semilog scale. For K_1, we note that because H is singular, the large eigenvalues grow without bound and the small eigenvalues converge to zero, in accordance with the bounds of Lemma 3.1. The bounds for $K_{1,\text{reg}}$ are fixed after a few initial iterations.

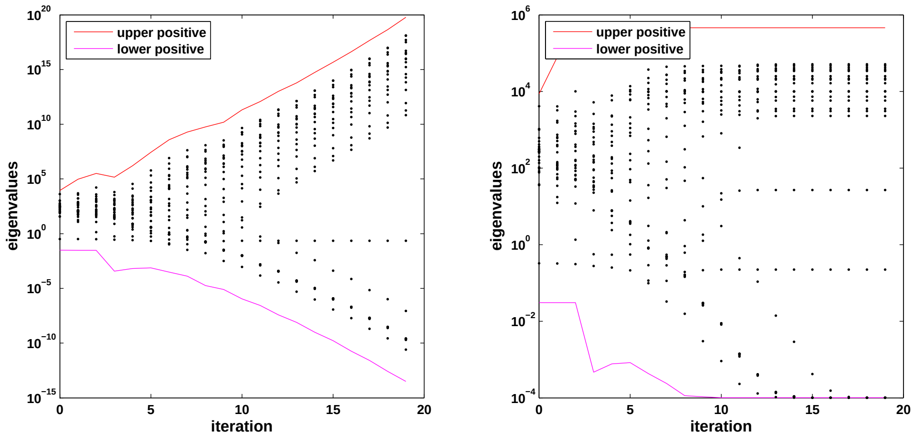
Figure 5.2 shows the positive and negative eigenvalues of K_2 and $K_{2,\text{reg}}$ on separate semilog plots. Note that the absolute value of the negative eigenvalues is represented. For K_2, the upper bound on the negative eigenvalues also decays, but the lower bound on the negative eigenvalues is fixed. This is in accordance with Lemma 3.3 and we see that the bounds are relatively tight in this example. For $K_{2,\text{reg}}$, while the inner bounds are now well away from zero, the upper bound on the positive eigenvalues still increases without bound, as predicted by Lemma 4.3, and all bounds are tight.

Figure 5.3 shows the eigenvalues of K_3 and $K_{3,\text{reg}}$ on separate semilog plots. For K_3, the two outer bounds are fixed and tight. On the other hand, the upper bound on the negative eigenvalues is zero and the lower bound on the positive eigenvalues decays toward zero. The eigenvalues also decay toward zero from both sides. This result is in accordance with Theorems 3.8 and 3.9. For $K_{3,\text{reg}}$, the lower bound on the positive eigenvalues is now well away from zero, and there is an upper bound on the negative eigenvalues at $-\delta$. Interestingly, at iteration 9 the magnitude of the smallest negative eigenvalues drops below δ, illustrating the "gray zone" when the conditions of Lemma 4.8 no longer hold but Lemma 4.9 does not yet apply. In fact, the assumptions of Lemma 4.9 are not satisfied for this problem. Despite these eigenvalues dropping below the bound, they remain away from zero.

We consider the condition numbers for all formulations for the unregularized problem and compare with the same problem regularized with $\rho = \delta = 10^{-4}$. Figure 5.4 shows the condition numbers of the different formulations. For the unregularized problem, all formulations are numerically singular at the end of the iterations, though K_3 and $\hat{K}_3$ have the best condition number. With regularization, $K_{1,\text{reg}}$ and both the unsymmetric and the symmetrized 3×3 formulations have bounded condition numbers, while $K_{2,\text{reg}}$ is still numerically singular.

We now consider the condition numbers for all formulations applied to Problem 6, and compare with the same problem regularized with $\rho = \delta = 10^{-4}$. Figure 5.5 shows the condition numbers of the different formulations. For the unregularized problem, K_3 and $\hat{K}_3$ are well-conditioned throughout, while the reduced forms have exponentially increasing condition numbers. With regularization, the condition number of $K_{1,\text{reg}}$ remains essentially fixed after the initial iterations because of the use of fixed regularization parameters, in accordance with Lemma 4.1. Note that in theory, this condition number could be as large as 10^8 but stabilizes around 10^5 in this instance. The condition numbers of the other formulations exhibit a similar behavior as in

[1] tomopt.com/tomlab



(a) $\rho = \delta = 0$ (b) $\rho = \delta = 10^{-4}$

FIG. 5.1. *TOMLAB problem 21: eigenvalues of* K_1 *and* $K_{1,reg}$.

the unregularized case. Since strict complementarity and LICQ are satisifed for this problem, regularization has little effect on the 3×3 formulations.

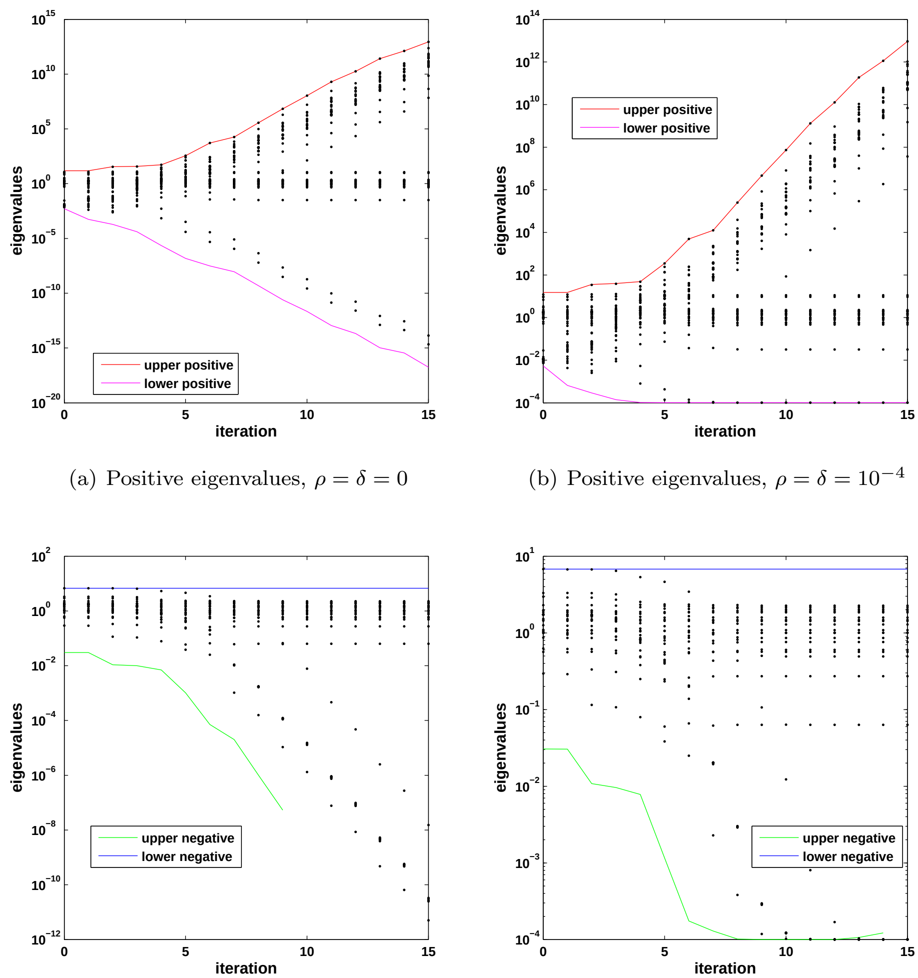
6. Other Formulations. In this section, we briefly mention alternative linear systems that were not covered in the previous sections but to which our techniques appear to generalize. In particular, we examine a formulation adapted to problems that do not satisfy strict complementarity at a solution.

When $H = 0$, the Goldman and Tucker (1956) theorem guarantees that a strictly complementary solution always exists provided there exists at least one solution. Moreover, widely applicable interior-point frameworks guarantee that any limit point of the sequence of iterates determines a strictly complementary solution under mild assumptions—see, e.g., (Wright, 1997, Theorem 6.8). When $H \neq 0$, this desirable result no longer holds in general, i.e., not all problems of the form (1.1a) possess a solution satisfying strict complementarity. A typical counter-example is

$$\underset{x \in \mathbb{R}}{\text{minimize}} \ \tfrac{1}{2} x^2 \quad \text{subject to} \ x \geq 0,$$

for which it is easy to verify that the only primal-dual solution is $(x, z) = (0, 0)$. A difficulty with such problems is that, in the limit, K_3 is singular while K_2 and K_1 do not even appear to be well defined. Our test scenario of Example 2.1 no longer describes the general situation because there is a subset of $\mathcal{A}$ for which $x_i = \Theta(\sqrt{\mu})$ and $z_i = \Theta(\sqrt{\mu})$—see, e.g., (Coulibaly et al., 2012; Monteiro and Wright, 1994; Wright and Orban, 2002).

It is common to study the iterates generated by an interior-point method in the vicinity of a strictly-complementary solution. Consider the typical situation where (1.1a) possesses a solution with at least one zero variable to which is associated a positive multiplier. An immediate difficulty is that K_2 does not converge to a well-defined limit and appears to become arbitrarily ill conditioned—an observation that is confirmed by the results of §3.2. The same holds for K_1. We now outline a strategy to salvage the situation in the case of K_2, and which may be applied to K_3 as well. In the



(a) Positive eigenvalues, $\rho = \delta = 0$

(b) Positive eigenvalues, $\rho = \delta = 10^{-4}$

(c) Negative eigenvalues, $\rho = \delta = 0$

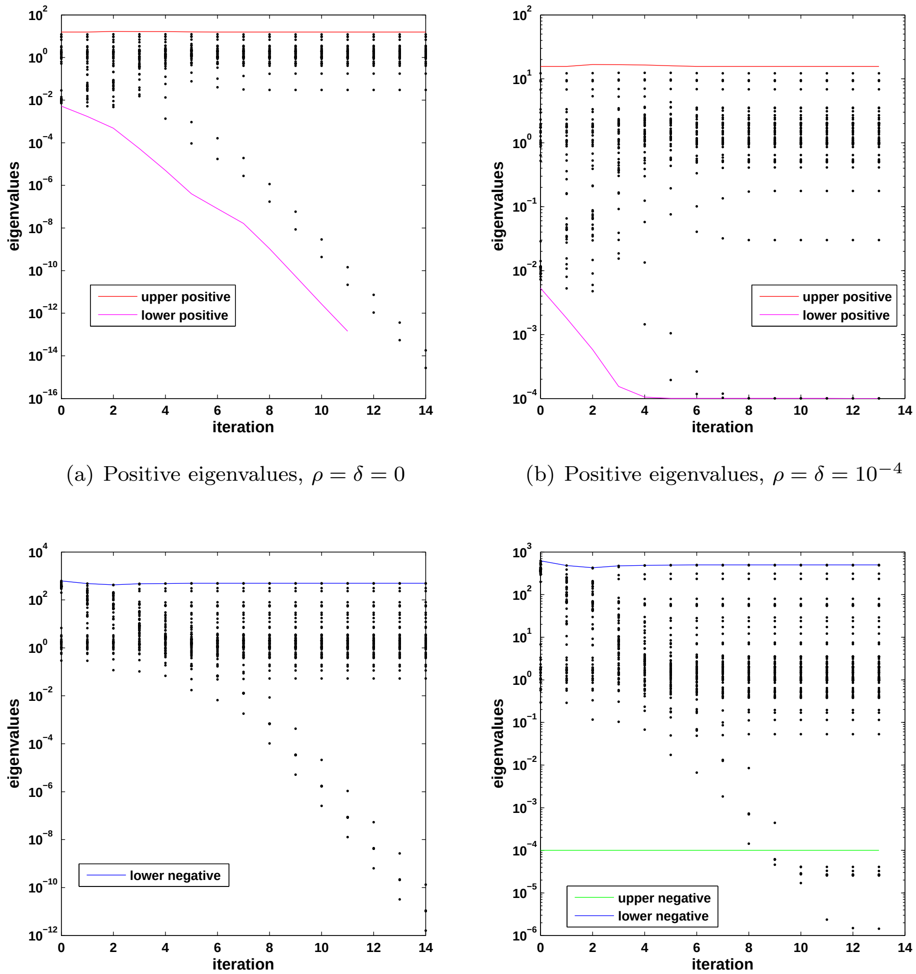
(d) Negative eigenvalues, $\rho = \delta = 10^{-4}$

FIG. 5.2. *TOMLAB problem 21: eigenvalues of* K_2 *and* $K_{2,reg}$.

vicinity of a strictly-complementary solution of (1.1)—assuming one exists—partition the variables according to $\mathcal{A}$ and $\mathcal{I}$ and consider the induced partitioning of the matrices H, J, X and Z. The system (2.2) may be written as

$$\begin{bmatrix} H_{\mathcal{A}\mathcal{A}} & H_{\mathcal{A}\mathcal{I}}^T & J_{\mathcal{A}}^T & -I & \\ H_{\mathcal{A}\mathcal{I}} & H_{\mathcal{I}\mathcal{I}} & J_{\mathcal{I}}^T & & -I \\ J_{\mathcal{A}} & J_{\mathcal{I}} & & & \\ -Z_{\mathcal{A}\mathcal{A}} & & & -X_{\mathcal{A}\mathcal{A}} & \\ & -Z_{\mathcal{I}\mathcal{I}} & & & -X_{\mathcal{I}\mathcal{I}} \end{bmatrix} \begin{bmatrix} \Delta x_{\mathcal{A}} \\ \Delta x_{\mathcal{I}} \\ \Delta y \\ \Delta z_{\mathcal{A}} \\ \Delta z_{\mathcal{I}} \end{bmatrix} = \begin{bmatrix} r_{c,\mathcal{A}} \\ r_{c,\mathcal{I}} \\ r_b \\ r_{\tau,\mathcal{A}} \\ r_{\tau,\mathcal{I}} \end{bmatrix}.$$

Note that $X_{\mathcal{I}\mathcal{I}}^{-1} Z_{\mathcal{I}\mathcal{I}}$ approaches zero. Gould (1986) then proposes to eliminate $\Delta z_{\mathcal{I}}$



(a) Positive eigenvalues, $\rho = \delta = 0$

(b) Positive eigenvalues, $\rho = \delta = 10^{-4}$

(c) Negative eigenvalues, $\rho = \delta = 0$

(d) Negative eigenvalues, $\rho = \delta = 10^{-4}$

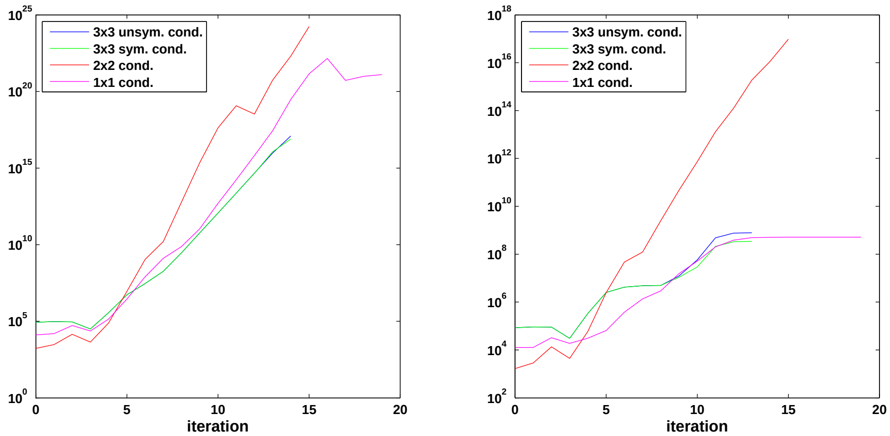
FIG. 5.3. *TOMLAB problem 21: eigenvalues of K_3 and $K_{3,reg}$. Note that the upper bound on the negative eigenvalues is zero for the unregularized system and is thus not visible in graph (c).*

and reformulate the above system as

$$\begin{bmatrix} H_{\mathcal{A}\mathcal{A}} & H_{\mathcal{A}\mathcal{I}}^T & J_{\mathcal{A}}^T & -I \\ H_{\mathcal{A}\mathcal{I}} & H_{\mathcal{I}\mathcal{I}} + X_{\mathcal{I}\mathcal{I}}^{-1} Z_{\mathcal{I}\mathcal{I}} & J_{\mathcal{I}}^T & \\ J_{\mathcal{A}} & J_{\mathcal{I}} & & \\ -Z_{\mathcal{A}\mathcal{A}} & & & -X_{\mathcal{A}\mathcal{A}} \end{bmatrix} \begin{bmatrix} \Delta x_{\mathcal{A}} \\ \Delta x_{\mathcal{I}} \\ \Delta y \\ \Delta z_{\mathcal{A}} \end{bmatrix} = \begin{bmatrix} r_{c,\mathcal{A}} \\ r_{c,\mathcal{I}} - X_{\mathcal{I}\mathcal{I}}^{-1} r_{\tau,\mathcal{I}} \\ r_b \\ r_{\tau,\mathcal{A}} \end{bmatrix}.$$

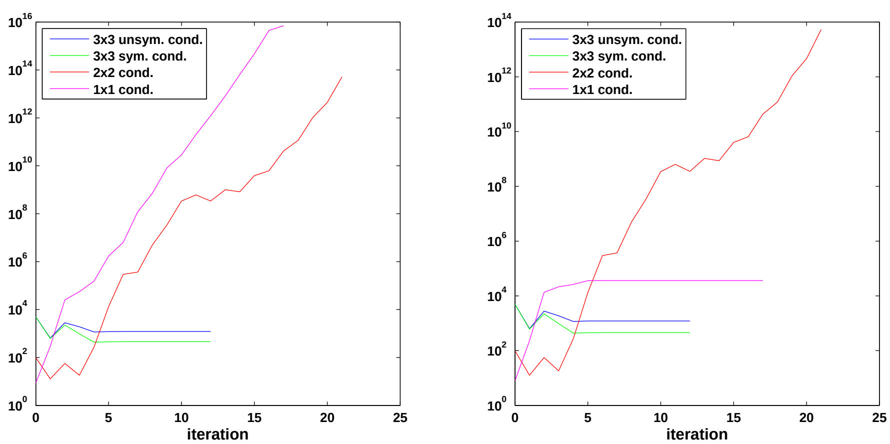
In addition, Gould (1986) symmetrizes the matrix by multiplying the last block row by $Z_{\mathcal{A}\mathcal{A}}^{-1}$. Since $X_{\mathcal{A}\mathcal{A}} Z_{\mathcal{A}\mathcal{A}}^{-1}$ also approaches zero, the resulting matrix then possesses a well-defined limit as long as J has full row rank, and therefore we should expect its condition number to be uniformly bounded, at least in a neighborhood of a strictly-complementary solution.

This partial elimination requires accurate identification of the index sets $\mathcal{A}$ and $\mathcal{I}$.



(a) $\rho = \delta = 0$ (b) $\rho = \delta = 10^{-4}$

FIG. 5.4. *TOMLAB problem 21: comparing the condition numbers for the unregularized and regularized systems.*



(a) $\rho = \delta = 0$ (b) $\rho = \delta = 10^{-4}$

FIG. 5.5. *TOMLAB problem 6: comparing the condition numbers for the unregularized and regularized systems.*

In practice, this can be done by using the predictive index set $\{i \mid x_i/z_i < z_i/x_i\}$ as an approximation to $\mathcal{A}$.

Some of our results below rely on strict complementarity being satisfied in the limit. We now outline a similar partitioning by which our results will apply to problems failing to satisfy strict complementarity.

A possible approach with such problems is to use indicator sets to distinguish between indices $i = 1, \ldots, n$ that are weakly active, strongly active and inactive. Let (x, y, z) be a local approximation to a solution. The set of *strongly active* constraints at x is $\mathcal{A}_S := \{i = 1, \ldots, n \mid x_i = 0 < z_i\}$. The set of *weakly active* constraints at

x is $\mathcal{A}_W := \{i = 1, \dots, n \mid x_i = z_i = 0\}$ and the set of *inactive* constraints at x is $\mathcal{I} := \{i = 1, \dots, n \mid z_i = 0 < x_i\}$. Suppose at each iteration k of an interior-point method, we have a mechanism to identify approximations $\mathcal{A}_S^k$, $\mathcal{A}_W^k$ and $\mathcal{I}^k$ to $\mathcal{A}_S$, $\mathcal{A}_W$ and $\mathcal{I}$, respectively. Such indicator sets can resolve the singular limit difficulty provided they ensure that $z_i^k/x_i^k \to 0$ as $k \to \infty$ for $i \in \mathcal{A}_W^k \cup \mathcal{I}^k$ while $x_i^k/z_i^k \to 0$ as $k \to \infty$ for $i \in \mathcal{A}_S^k$. Indeed if this were the case, upon partitioning x, z, H and J according to $\mathcal{B}^k := \mathcal{A}_W^k \cup \mathcal{I}^k$ and $\mathcal{S}^k := \mathcal{A}_S^k$, (2.2) could be partially eliminated to

$$\begin{bmatrix} H_{\mathcal{SS}} & H_{\mathcal{SB}}^T & J_{\mathcal{S}}^T & -I \\ H_{\mathcal{SB}} & H_{\mathcal{BB}} + X_{\mathcal{BB}}^{-1} Z_{\mathcal{BB}} & J_{\mathcal{B}}^T & \\ J_{\mathcal{S}} & J_{\mathcal{B}} & & \\ -Z_{\mathcal{SS}} & & & -X_{\mathcal{SS}} \end{bmatrix} \begin{bmatrix} \Delta x_{\mathcal{S}} \\ \Delta x_{\mathcal{B}} \\ \Delta y \\ \Delta z_{\mathcal{S}} \end{bmatrix} = \begin{bmatrix} r_{c,\mathcal{S}} \\ r_{c,\mathcal{B}} + X_{\mathcal{BB}}^{-1} r_{\tau,\mathcal{B}} \\ r_b \\ r_{\tau,\mathcal{S}} \end{bmatrix}. \tag{6.1}$$

The matrix of the latter system has a well-defined limit whenever J has full row rank. As above, the typical way to symmetrize the system is to multiply the last block row by $Z_{\mathcal{SS}}^{-1}$. Examples of indicator sets with the requisite properties along with pointers to the literature are given by Coulibaly et al. (2012) and Monteiro and Wright (1994).

In both cases of partial elimination above, should J not have full row rank, dual regularization, as described in §4, will ensure that the partially eliminated matrix converges to a well-defined limit, provided the regularization parameter remains bounded away from zero.

We also note that our results of §3.3 and §4.3 can be extended to the appropriate symmetrization of the two above partially eliminated systems with the benefit of solving smaller systems with similar properties.

There are other formulations of the system that we have not covered in this paper. Forsgren and Gill (1998) propose a *doubly augmented* formulation that has the benefit of being positive definite on and near the central path, allowing the use of more specialized linear solvers. Korzak (1999) and Fourer and Mehrotra (1993) both use scalings of the systems to alleviate ill-conditioning. Gill et al. (1992) proceed similarly to Gould (1986) and partition inequality constraints into active and inactive constraints and note that ill-conditioning is due to the varying sizes of constraint values across the partition. Benzi et al. (2009) discuss the conditioning of the formulations that we cover, focusing on a specific application. They use iterative methods with large-scale problems, where the ill-conditioning of the reduced forms is problematic. In their experiments, they use the original unreduced form and a partially-eliminated formulation.

7. Concluding Remarks. Our analysis indicates that 3×3 block linear systems that arise throughout the iterations of primal-dual interior-point solvers have a favorable spectral structure. The matrix is nonsingular, whereas it tends to singularity when it is reduced using commonly used partial elimination procedures. Regularization is shown to be very effective at shifting the eigenvalues and the eigenvalue bounds away from zero. It is important to note that convergence may be driven by several considerations other than the eigenvalue distribution, especially when direct linear solvers are used. The subspace where the right-hand side lies and the scaling of the problem are two of several issues that affect convergence.

Our "outer" bounds, namely the upper bound on the positive eigenvalues and the lower bound on the negative eigenvalues, seem tight. On the other hand, the "inner" bounds, namely the lower bound on the positive eigenvalues and the upper bound on the negative eigenvalues, are pessimistic in some cases. Specifically, we have not been able to find an effective upper bound on the negative eigenvalues for the unregularized case, but the matrix is analytically known to be nonsingular, and further numerical experiments confirm that the largest eigenvalues in absolute value of the unreduced 3×3 block matrix are bounded away from zero and only modestly grow throughout the iterations. In the regularized case we have found all bounds, and consequently were able to establish an asymptotic condition number. Here our analysis for the upper negative bound involves a region that we term the "gray zone", which represents the transition from the course of iterations to the limit and inevitably deals with the introduction of singularity. Our analysis does not provide an effective bound for the gray zone, but we believe that by tightening our assumptions we will be able to further shrink this zone. This remains an item for further exploration.

We have not addressed the issue of the performance of iterative solvers, but our work in this paper has been done with our eyes on this paradigm. As problems become larger in scale, the importance of studying these solvers is rising. For such solvers the distribution of eigenvalues plays a central role, and the need to find effective preconditioners is critical and challenging. The question of how to utilize inexactness throughout the iterations is also important. See Benzi et al. (2009); Bergamaschi et al. (2007); Forsgren et al. (2007) for relevant and important observations and results. The bounds given in this paper as well as preliminary experimentation lead us to believe that the 3×3 formulation has a strong potential to be very effective in the context of iterative solvers.

Matrices similar to K_1, K_2 and K_3 occur in other branches of optimization, such as cone optimization. The difference in this case is that X and Z may no longer be diagonal but are instead restricted to being symmetric and positive definite. In addition, they may be dense. Whether our results generalize to such a scenario may be an interesting subject for future research.

Acknowledgments. The authors wish to thank Jacek Gondzio and Michael Saunders for carefully reading this paper and for numerous valuable comments. The third author gratefully acknowledges constant moral support from Ludwig Van Beethoven.

References.

A. Altman and J. Gondzio. Regularized symmetric indefinite systems in interior point methods for linear and quadratic optimization. *Optimization Methods and Software*, 11–12:275–302, 1999. DOI: 10.1080/10556789908805754.

P. Armand and J. Benoist. Uniform boundedness of the inverse of a Jacobian matrix arising in regularized interior-point methods. *Mathematical Programming*, pages 1–6, 2011. DOI: 10.1007/s10107-011-0498-3. Online First.

M. Benzi, G. H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numerica*, 14:1–137, 2005. DOI: 10.1017/S0962492904000212.

M. Benzi, E. Haber, and L. Taralli. Multilevel algorithms for large-scale interior point methods. *SIAM Journal on Scientific Computing*, 31(6):4152–4175, 2009. DOI: 10.1137/060650799.

L. Bergamaschi, J. Gondzio, M. Venturin, and G. Zilli. Inexact constraint preconditioners for linear systems arising in interior point methods. *Computational Optimization and Applications*, 36(2-3):137–147, 2007. DOI: 10.1007/s10589-006-9001-0.

R. H. Byrd, M. E. Hribar, and J. Nocedal. An interior point method for large scale nonlinear programming. *SIAM Journal on Optimization*, 9:877–900, 1999. DOI: 10.1137/S1052623497325107.

R. H. Byrd, J.-Ch. Gilbert, and J. Nocedal. A trust region method based on interior point techniques for nonlinear programming. *Mathematical Programming, Series A*, 89(1):149–185, 2000. DOI: 10.1007/PL00011391.

Z. Coulibaly, N. I. M. Gould, and D. Orban. Fast local convergence of interior-point methods in the absence of strict complementarity. Cahier du GERAD G-2012-xx, GERAD, Montréal, QC, Canada, 2012. In preparation.

J. Czyzyk, S. Mehrotra, M. Wagner, and S. J. Wright. PCx: an interior-point code for linear programming. *Optimization Methods and Software*, 11/12(1-4):397–430, 1999. DOI: 10.1080/10556789908805757.

A. V. Fiacco and G. P. McCormick. *Nonlinear programming*, volume 4 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 1990.

A. Forsgren. Inertia-controlling factorizations for optimization algorithms. *Applied Numerical Mathematics*, 43(1):91–107, 2002. DOI: 10.1016/S0168-9274(02)00119-8.

A. Forsgren and Ph. E. Gill. Primal-dual interior methods for nonconvex nonlinear programming. *SIAM Journal on Optimization*, 8(4):1132–1152, 1998. DOI: 10.1137/S1052623496305560.

A. Forsgren, P. E. Gill, and J. R. Shinnerl. Stability of symmetric ill-conditioned systems arising in interior methods for constrained optimization. *SIAM Journal on Matrix Analysis and Applications*, 17(1):187–211, 1996. DOI: 10.1137/S0895479894270658.

A. Forsgren, Ph. E. Gill, and M. H. Wright. Interior methods for nonlinear optimization. *SIAM Review*, 44(4):525–597, 2002. DOI: 10.1137/S0036144502414942.

A. Forsgren, P. E. Gill, and J. D. Griffin. Iterative solution of augmented systems arising in interior methods. *SIAM Journal on Optimization*, 18(2):666–690, 2007. DOI: 10.1137/060650210.

R. Fourer and S. Mehrotra. Solving symmetric indefinite systems in an interior-point method for linear programming. *Mathematical Programming*, 62(1):15–39, 1993. DOI: 10.1007/BF01585158.

M. P. Friedlander and D. Orban. A primal-dual regularized interior-point method for convex quadratic programs. *Mathematical Programming Computation*, 4(1):71–107, 2012. DOI: s12532-012-0035-2.

E. M. Gertz and S. J. Wright. Object-oriented software for quadratic programming. *ACM Transactions on Mathematical Software*, 29(1):58–81, 2003. DOI: 10.1145/641876.641880.

Ph. E. Gill, W. Murray, D. B. Ponceleón, and M. A. Saunders. Preconditioners for indefinite systems arising in optimization. *SIAM Journal on Matrix Analysis and Applications*, 13(1):292–311, 1992. DOI: 10.1137/0613022.

A. J. Goldman and A. W. Tucker. Theory of linear programming. In H. W. Kuhn and A. W. Tucker, editors, *Linear Equalities and Related Systems*, pages 53–97. Princeton University Press, Princeton, N.J., 1956.

J. Gondzio. Matrix-free interior point methods. *Computational Optimization and Applications*, 51:457–480, 2012. DOI: 10.1007/s10589-010-9361-3.

N. I. M. Gould. On practical conditions for the existence and uniqueness of solutions to the general equality quadratic-programming problem. *Mathematical Programming*, 32(1):90–99, 1985. DOI: 10.1007/BF01585660.

N. I. M. Gould. On the accurate determination of search directions for simple differentiable penalty functions. *IMA Journal of Numerical Analysis*, 6(3):357–372, 1986. DOI: 10.1093/imanum/6.3.357.

N. I. M. Gould and V. Simoncini. Spectral analysis of saddle point matrices with indefinite leading blocks. *SIAM Journal on Matrix Analysis and Applications*, 31 (3):1152–1171, 2009. DOI: 10.1137/080733413.

J. Korzak. Eigenvalue relations and conditions of matrices arising in linear programming. *Computing*, 62(1):45–54, 1999. DOI: 10.1007/s006070050012.

S. Mehrotra. On the implementation of a primal-dual interior point method. *SIAM Journal on Optimization*, 2(4):575–601, 1992. DOI: 10.1137/0802028.

R. D. C. Monteiro and S. J. Wright. Local convergence of interior-point algorithms for degenerate monotone LCP. *Computational Optimization and Applications*, 3(2): 131–155, 1994. DOI: 10.1007/BF01300971.

T. Rusten and R. Winther. A preconditioned iterative method for saddlepoint problems. *SIAM Journal on Matrix Analysis and Applications*, 13(3):887–904, 1992. DOI: 10.1137/0613054.

M. A. Saunders. Cholesky-based methods for sparse least squares: The benefits of regularization. In L. Adams and J. L. Nazareth, editors, *Linear and Nonlinear Conjugate Gradient-Related Methods*, pages 92–100. SIAM, Philadelphia, 1996.

D. Silvester and A. Wathen. Fast iterative solution of stabilised Stokes systems. II. Using general block preconditioners. *SIAM Journal on Numerical Analysis*, 31(5): 1352–1367, 1994. DOI: 10.1137/0731070.

A. Wächter and L. T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming, Series A*, 106(1):25–57, 2006. DOI: 10.1007/s10107-004-0559-y.

M. H. Wright. Interior methods for constrained optimization. *Acta Numerica*, 1: 341–407, 1992. DOI: 10.1017/S0962492900002300.

M. H. Wright. Some properties of the Hessian of the logarithmic barrier function. *Mathematical Programming*, 67(1-3):265–295, 1994. DOI: 10.1007/BF01582224.

M. H. Wright. The interior-point revolution in optimization: history, recent developments, and lasting consequences. *Bulletin of the American Mathematical Society*, 42 (1):39–56, 2005. DOI: 10.1090/S0273-0979-04-01040-7.

S. J. Wright. *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia, USA, 1997.

S. J. Wright and D. Orban. Local convergence of the Newton/log-barrier method for degenerate problems. *Mathematics of Operations Research*, 27(3):585–613, 2002. DOI: 10.1287/moor.27.3.585.312.