SPECTRAL ESTIMATES FOR UNREDUCED SYMMETRIC KKT SYSTEMS ARISING FROM INTERIOR POINT METHODS*

BENEDETTA MORINI[§], VALERIA SIMONCINI[†], AND MATTIA TANI[†]

Abstract. We consider symmetrized KKT systems arising in the solution of convex quadratic programming problems in standard form by Interior Point methods. Their coefficient matrices usually have 3×3 block structure and, under suitable conditions on both the quadratic programming problem and the solution, they are nonsingular in the limit.

We present new spectral estimates for these matrices: the new bounds are established for the unpreconditioned matrices and for the matrices preconditioned by symmetric positive definite augmented preconditioners. Some of the obtained results complete the analysis recently given by Greif, Moulding and Orban in [*SIAM J. Optim., 24 (2014), pp. 49-83*]. The sharpness of the new estimates is illustrated by numerical experiments.

Key words. convex quadratic programming, interior point methods, indefinite linear systems, eigenvalue bounds, preconditioners

1. Introduction. Interior Point (IP) methods are effective iterative procedures for solving linear, quadratic and nonlinear programming problems, possibly of very large dimension, see [1, 4, 17, 10, 23, 35] and references therein. Since they are second-order methods, a linear algebra phase constitutes their computational core and its practical implementation is crucial for the efficiency of the overall optimization procedure. Therefore, linear algebra of IP methods has been extensively studied in all algorithmic issues, including formulation of the systems arising at each iteration, employment of direct and iterative solvers, preconditioning and inertia control.

This work is devoted to the study of KKT systems[1] arising in the solution of convex quadratic programming (QP) problems with primal-dual pair of the form

$$\min_{x}\ c^T x + \frac{1}{2}x^T Hx \quad \text{subject to} \quad Jx = b,\ x \geq 0, \tag{1.1}$$

$$\max_{x,y,z}\ b^T y - \frac{1}{2}x^T Hx \quad \text{subject to} \quad J^T y + z - Hx = c,\ z \geq 0, \tag{1.2}$$

where $J \in \mathbb{R}^{m\times n}$ has full row rank $m<n$, $H \in \mathbb{R}^{n\times n}$ is symmetric and positive semidefinite (SPSD in the following), $x,\ z,\ c \in \mathbb{R}^n$, $y,\ b \in \mathbb{R}^m$, and the inequalities are meant componentwise. The application of a primal-dual IP method gives rise, at each iteration, to an unsymmetric 3×3 block matrix of dimension $2n + m$ which allows for alternative formulations of differing dimension, conditioning and definiteness [8, 10, 21, 33]. In fact, the unsymmetric 3×3 matrix can be easily symmetrized without increasing the conditioning of the system [7], and here we will refer to the resulting symmetric matrix as the *unreduced* KKT matrix. On the other hand, by exploiting the structure of the unsymmetric 3×3 matrix and block elimination, it is common

* Work partially supported by INdAM-GNCS under the 2015 Project *Metodi di regolarizzazione per problemi di ottimizzazione e applicazioni*, and the 2014 Project *Metodi di regolarizzazione per problemi di ottimizzazione vincolata*.

[†]Dipartimento di Matematica, Università di Bologna, Piazza di Porta S. Donato 5, 40127 Bologna, Italia. valeria.simoncini@unibo.it, mattia.tani2@unibo.it

[§]Dipartimento di Ingegneria Industriale, Università degli Studi di Firenze, viale G.B. Morgagni 40, 50134 Firenze, Italia, benedetta.morini@unifi.it.

[1]In the context of interior point methods, the coefficient matrix of the system to be solved at each iteration is often referred to as the *barrier* KKT matrix. Following [7, p. 92] we omit the term "barrier" for simplicity and denote the system as KKT system, see also [8, 6].

to use a linear system of dimension $n + m$ with a *reduced* (or *augmented*) symmetric 2×2 KKT matrix. Finally, a further block elimination may yield a *condensed* system (or *normal equations*) where the matrix is a Schur complement of dimension n. The focus of this work is the theoretical study of the unreduced 3×3 formulation and the numerical illustration of the obtained results.

Unlike the reduced 2×2 matrix, under suitable conditions on both the problem (1.1) and the solution, the unreduced KKT matrix has condition number asymptotically uniformly bounded, and typically remains well-conditioned as the solution is approached [7, 10]. Motivated by this feature, in a very recent paper Greif et al. [19] presented a spectral analysis for the 3×3 matrix and claimed that this formulation can be preferable to the reduced one in terms of eigenvalues and conditioning, although a "benign" asymptotically ill-conditioned scaling has to be applied to the right-hand side and to the system variables [10]. Such a study covers also variants of KKT matrices arising from regularization of the optimization problem.

The study conducted in [19] has renewed the interest in the unreduced formulation but leaves some issues open, that are worth investigating and are a prerequisite for a thorough comparison of the 2×2 and 3×3 formulations. Specifically, some eigenvalue bounds presented in [19] may be overly pessimistic and not tight for the unregularized 3×3 matrix; in fact they may not reflect the nonsingularity of the matrix.

Our contribution consists of a spectral analysis of the unpreconditioned and of the diagonally (augmented) preconditioned unreduced matrix. Firstly, by reflecting the potential nonsingularity of the matrices at the limit, we complement the results of [19] providing two missing bounds: an upper bound on the negative eigenvalues and a lower bound on positive eigenvalues when H is SPSD and IP iterations progress. These results give rise to new estimates of the condition number for the unpreconditioned matrix and sharply characterize different stages of the Interior Point method. This piece of information can be helpful for both the direct and iterative unpreconditioned solution of the system. Secondly, we present a new spectral analysis for a general class of KKT matrices preconditioned by an augmented block diagonal preconditioner, and specialize such results to the unreduced matrix. Preconditioners that offer spectral intervals of the preconditioned matrix independent of the problem data are discussed, while a more detailed analysis of economical variants can be found in [26].

The outline of this paper is as follows. In Section 2 we introduce the KKT systems studied and summarize existing results on their properties. In Section 3 we give new estimates on the bounds of the unreduced KKT matrix and perform the analysis for the early and middle stage of the IP method, and for the late stage of the IP method, separately. A numerical validation of the bounds obtained is also given. In Section 4 we provide the spectral analysis of the unreduced matrix preconditioned by a class of augmented block diagonal preconditioners, and a numerical validation of the sharpness of the obtained results. Final conclusions are drawn in Section 5.

Notation. In the following, $\|\cdot\|$ denotes the vector 2-norm or its induced matrix norm. For any vector x, the ith component is denoted as either x_i or $(x)_i$; furthermore, $x_{\min}$ and $x_{\max}$ (or $(x)_{\min}$, $(x)_{\max}$) denote the minimum and maximum components of x in absolute value. Given $x \in \mathbb{R}^n$, $X = \text{diag}(x)$ is the diagonal matrix with diagonal entries $x_1, \ldots, x_n$. Given column vectors x and y, we write (x, y) for the column vector given by their concatenation instead of using $[x^T, y^T]^T$; analogously for (x, y, z). For any $x \in \mathbb{R}^n$ and set of indices $\mathcal{A} \subset \{1, 2, \ldots, n\}$, we write $x_{\mathcal{A}}$ for the subvector of x having components x_i with $i \in \mathcal{A}$. Further, if B is a matrix we write $B_{\mathcal{A}}$ for the submatrix of the columns of B with indices in $\mathcal{A}$. Given a matrix

A, the set of its eigenvalues is indicated by $\Lambda(A)$. For an arbitrary symmetric matrix A, unless explicitly stated, $\lambda_i(A)$ denotes the ith eigenvalue of A sorted in increasing order, while $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ denote its minimum and maximum eigenvalues. The symbols $\sigma_{\min}(B)$ and $\sigma_{\max}(B)$ represent the minimum and maximum singular values of a given matrix B. In particular, due to their recurrence throughout the manuscript, we let $\lambda_{\min}$ and $\lambda_{\max}$ be the minimum and maximum eigenvalues of H (without parentheses), and analogously $\sigma_{\min}$ and $\sigma_{\max}$ be the minimum and maximum singular values of J.

Finally, given two square matrices A and B of the same dimension, $A \succeq B$ means that $A - B$ is SPSD. For any positive integer p, the identity matrix of dimension p is indicated by I_p.

2. Preliminaries. First-order optimality conditions for the problems (1.1) and (1.2) are described by a mildly nonlinear system where x and z are bounded to be nonnegative and the nonlinearity arises entirely from the complementarity condition between such variables. Primal-dual IP methods find solutions $(\hat{x}, \hat{y}, \hat{z})$ for problems (1.1)-(1.2) by generating iterates $(x_k, y_k, z_k) \in \mathbb{R}^{2n+m}$ where x_k and z_k are strictly feasible with respect to the simple bounds, i.e. $x_k,\ z_k > 0$. Search directions biased toward the interior of the nonnegative orthant $(x, z) \geq 0$ are computed by perturbing the complementarity condition and applying Newton method to the resulting system. We refer the reader to [35] for details on primal-dual IP methods.

The primal-dual Newton direction solves, possibly approximately, the linear system of dimension $2n + m$

$$\begin{bmatrix} H & J^T & -I_n \\ J & 0 & 0 \\ -Z_k & 0 & -X_k \end{bmatrix} \begin{bmatrix} \Delta x_k \\ -\Delta y_k \\ \Delta z_k \end{bmatrix} = \begin{bmatrix} -c - Hx_k + J^T y_k + z_k \\ b - Jx_k \\ X_k Z_k 1_n - \tau_k 1_n \end{bmatrix}, \tag{2.1}$$

where $X_k = \operatorname{diag}(x_k)$, $Z_k = \operatorname{diag}(z_k)$ are diagonal positive definite, $1_n \in \mathbb{R}^n$ is the vector of ones and the positive scalar τ_k controls the distance to optimality and it is gradually reduced during the IP iterations. When predictor-corrector schemes are applied, the system (2.1) is solved for different right-hand sides. For notational convenience the iteration subscript will be dropped in the following.

The computational core of these methods consists of the sequence of linear systems arising during the iterative procedure and different formulations for such systems are allowed, as discussed below. The matrix in (2.1), say $K_{3,\text{uns}}$, is symmetrizable by setting (see [7])

$$K_3 = R^{-1} K_{3,\text{uns}} R = \begin{bmatrix} H & J^T & -Z^{\frac{1}{2}} \\ J & 0 & 0 \\ -Z^{\frac{1}{2}} & 0 & -X \end{bmatrix}, \quad \text{where} \quad R = \begin{bmatrix} I_n & 0 & 0 \\ 0 & I_m & 0 \\ 0 & 0 & Z^{\frac{1}{2}} \end{bmatrix}. \tag{2.2}$$

Thus, we can consider the system equivalent to (2.1) with matrix[2] K_3. Due to the presence of zero and diagonal blocks in (2.1), it is very common to eliminate Δz from the third equation and to obtain a KKT system of dimension $n + m$ with matrix

$$K_2 = \begin{bmatrix} H + X^{-1}Z & J^T \\ J & 0 \end{bmatrix}. \tag{2.3}$$

[2] There are other ways to symmetrize $K_{3,\text{uns}}$; matrix K_3 considered here does not suffer inevitable ill-conditioning as the solution is approached [7, 10].

One further block elimination step yields the normal equation with matrix $K_1 = J(H + X^{-1}Z)^{-1}J^T$. With a proper block partitioning, K_3 can be cast into a KKT formulation as well. Under our assumptions on H and J, and as long as X and Z are diagonal with positive entries, K_2 and K_3 are nonsingular; K_2 has n positive and m negative eigenvalues, while K_3 has n positive and $n+m$ negative eigenvalues, see e.g. [4, Lemma 4.1], [19, Lemma 3.5, 3.8]. The distinctive feature of these two matrices is their behavior in the limit of the IP procedure. Specifically, if $(\hat{x}, \hat{y}, \hat{z}) \in \mathbb{R}^{2n+m}$ solves the QP problem (1.1)-(1.2), then $\hat{x}$, $\hat{z}$ are nonnegative and satisfy the complementarity condition $\hat{x}_i\hat{z}_i = 0, \quad i = 1, \ldots, n$. Let

$$\mathcal{A} := \{i = 1, \ldots, n \mid \hat{x}_i = 0\}, \qquad \mathcal{I} := \{1, \ldots, n\} \setminus \mathcal{A}. \tag{2.4}$$

The vectors $\hat{x}$, $\hat{z}$ are strictly complementary when $\hat{z}_i > 0$, for all $i \in \mathcal{A}$. As a consequence, when the IP iterates approach a solution, some entries of $X^{-1}Z$ tend to zero while others tend to infinity and the eigenvalues of the $(1,1)$ block in K_2 may spread from zero to infinity. The effect of this feature on the conditioning of K_2 can be formally described in the situation where

$$\min_{1\le i\le n} \frac{z_i}{x_i} = \mathcal{O}(\mu), \qquad \text{and} \qquad \max_{1\le i\le n} \frac{z_i}{x_i} = \mathcal{O}(\mu^{-1}),$$

and $\mu = x^T z/n$ is the duality measure. These asymptotic estimates hold when strict complementary is in place, $\mathcal{A} \neq \emptyset$, $\mathcal{I} \neq \emptyset$, and the iterates are restricted to a suitable neighborhood of the central path, see, e.g., [19, 17]. As a consequence of these assumptions, the asymptotic condition number of K_2 may get as large as $\mathcal{O}(\mu^{-2})$, [17, Lemma 2.2], [19, Corollary 5.2]. Remedies to this occurrence may consist either in scalings of K_2 [11] or in regularization strategies [17, 12, 30]. For the sake of completeness, we recall here that ill-conditioning of the matrix is usually harmless when direct methods are applied [9, 34].

Under suitable conditions stated below, the unreduced matrix K_3 can be well conditioned and nonsingular in the limit although the diagonal scaling (2.2) used for forming the right-hand side of the system and unscaling the variables remains *benignly* ill-conditioned [7, 10]. Therefore, a spectral analysis of the original K_3 may give insight into both its conditioning and the possible need for regularization.

Let q be the cardinality of the set $\mathcal{A}$, $q = \text{card}(\mathcal{A})$. Without loss of generality, suppose that the zero components of $\hat{x}$ are its first q elements. Hence, if $\hat{x}$ and $\hat{z}$ are strictly complementary by (2.4) we have

$$\hat{x} = (0, \hat{x}_{\mathcal{I}}), \quad \hat{z} = (\hat{z}_{\mathcal{A}}, 0), \quad \hat{x}_{\mathcal{I}} > 0, \quad \hat{z}_{\mathcal{A}} > 0, \tag{2.5}$$

where $\hat{x}_{\mathcal{I}} \in \mathbb{R}^{n-q}$, $\hat{z}_{\mathcal{A}} \in \mathbb{R}^q$. The next definition will be used in the following.

Definition 1. *The Linear Independence Constraint Qualification (LICQ) is satisfied at $\hat{x}$ if the matrix $\begin{bmatrix} J^T & -I_{\mathcal{A}} \end{bmatrix}$ has full column rank.*

Note that a necessary condition for the LICQ condition to be satisfied at any point is that J has full row rank.

It is useful to make some comments on the matrices $K_{3,\text{uns}}$ and K_3 evaluated at $x = \hat{x}$, $z = \hat{z}$. To this end, we let

$$\widehat{K}_{3,\text{uns}} = \begin{bmatrix} H & J^T & -I_n \\ J & 0 & 0 \\ -\widehat{Z} & 0 & -\widehat{X} \end{bmatrix}, \quad \widehat{K}_3 = \begin{bmatrix} H & J^T & -\widehat{Z}^{\frac{1}{2}} \\ J & 0 & 0 \\ -\widehat{Z}^{\frac{1}{2}} & 0 & -\widehat{X} \end{bmatrix}, \tag{2.6}$$

where $\widehat{X} = \text{diag}(\hat{x})$, $\widehat{Z} = \text{diag}(\hat{z})$. Throughout the paper, $\widehat{K}_3$ and K_3 will denote the coefficient matrices at the QP solution and during the iterations, respectively.

The systems involving matrices $K_{3,\text{uns}}$ and K_3 are formally equivalent also at the exact solution $(\hat{x}, \hat{z})$, at least after the natural elimination of some equations. Indeed, let us assume for simplicity that $\hat{x}$ and $\hat{z}$ are partitioned as in (2.5) and strictly complementary. Then, if in equation (2.1) we substitute $x = \hat{x}$ and $z = \hat{z}$, upon reduction of the components of Δz with indices in $\mathcal{I}$, we get a system with matrix

$$\begin{bmatrix} \multicolumn{2}{c}{\multirow{2}{*}{H}} & J_{\mathcal{A}}^T & -I_q \\ & & J_{\mathcal{I}}^T & 0 \\ J_{\mathcal{A}} & J_{\mathcal{I}} & 0 & 0 \\ -Z_{\mathcal{A}} & 0 & 0 & 0 \end{bmatrix},$$

and $J_{\mathcal{A}} \in \mathbb{R}^{m\times q}$, $J_{\mathcal{I}} \in \mathbb{R}^{m\times(n-q)}$ and $Z_{\mathcal{A}} = \text{diag}(\hat{z}_{\mathcal{A}}) \in \mathbb{R}^{q\times q}$. By using the similarity transformation with a matrix of the form (2.2), namely

$$\begin{bmatrix} I_n & 0 & 0 \\ 0 & I_m & 0 \\ 0 & 0 & Z_{\mathcal{A}}^{\frac{1}{2}} \end{bmatrix},$$

the resulting system is symmetric with matrix obtained by removing in $\widehat{K}_3$ the last block row and column associated to the set $\mathcal{I}$.

We also observe that $\widehat{K}_3$ and $\widehat{K}_{3,\text{uns}}$ have the same eigenvalues. Indeed, by (2.2) K_3 and $K_{3,\text{uns}}$ have the same eigenvalues for every strictly positive x and z. A continuity argument shows that they also coincide when taking the limit as $x \to \hat{x}$ and $z \to \hat{z}$. The following theorem states conditions under which $\widehat{K}_3$ is nonsingular, see, e.g., [19, Theorem 3.10].

Theorem 2.1. *Suppose H is SPSD, $\widehat{X}$ and $\widehat{Z}$ are diagonal with nonnegative entries. Then $\widehat{K}_3$ in (2.6) is nonsingular if and only if*

1. *$\hat{x}$ and $\hat{z}$ are strictly complementary,*
2. *the LICQ is satisfied at $\hat{x}$,*
3. *the null spaces of matrices H, J, $\widehat{Z}$ satisfy*

$$\ker(H) \cap \ker(J) \cap \ker(\widehat{Z}) = \{0\}. \tag{2.7}$$

In the next theorem we summarize the bounds for the eigenvalues of K_3 given in [19, Corollary 5.3 and Corollary 5.4]. We recall that we denote by $\lambda_{\min}$ and $\lambda_{\max}$ the minimum and maximum eigenvalues of H and by $\sigma_{\min}$ and $\sigma_{\max}$ the minimum and maximum singular values of J.

Theorem 2.2. *Suppose H is SPSD and let K_3 be as in (2.2).*

i) If $\theta^- I_n + X$ is nonsingular for all the negative eigenvalues θ^- of K_3, then $\theta^- \in [\zeta, 0)$, where

$$\zeta = \min\left\{\frac{1}{2}\left(\lambda_{\min} - \sqrt{\lambda_{\min}^2 + 4\sigma_{\max}^2}\right), \min_{\{j | x_j+\theta^-<0\}} \theta_j^*\right\},$$

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

$$p_j(\theta) = \theta^3 + (x_j - \lambda_{\min})\theta^2 - (\sigma_{\max}^2 + z_j + x_j\lambda_{\min})\theta - x_j\sigma_{\max}^2. \tag{2.8}$$

ii) If J has full rank and X and Z are diagonal with positive entries, then the positive eigenvalues θ^+ of K_3 satisfy $\theta^+ \in [\theta_3,\, \theta_4]$, where

$$\theta_3 = \min_{1\le j\le n} \frac{1}{2}\left(\lambda_{\min} - x_j + \sqrt{(\lambda_{\min} + x_j)^2 + 4z_j}\right), \tag{2.9}$$

$$\theta_4 = \frac{1}{2}\left(\lambda_{\max} + \sqrt{\lambda_{\max}^2 + 4(\sigma_{\max}^2 + z_{\max})}\right). \tag{2.10}$$

A computable lower bound on the negative eigenvalues of K_3 is ([19, p. 68])

$$\theta_1 := \min\left\{\frac{1}{2}\left(\lambda_{\min} - \sqrt{\lambda_{\min}^2 + 4\sigma_{\max}^2}\right), \min_j \theta_j^*\right\}. \tag{2.11}$$

In [19, §5.2] Greif et al. observe that the lower bound in Theorem 2.2(i) is established excluding that some eigenvalue θ^- of K_3 belong to the spectrum of $-X$ but this assumption may fail both in the course of the iterations and in the limit if there are inactive bounds. They also note that the zero upper bound in Theorem 2.2(i) is not particularly meaningful either in the case where $(x, z) > 0$ or in the limit. Finally, they point out that the lower bound θ_3 in Theorem 2.2(ii) is strictly positive as long as $(x, z) > 0$ but in the limit it reduces to $\lambda_{\min}$ and may be overly pessimistic if $\lambda_{\min} = 0$. In particular, the nonsingularity of $\widehat{K}_3$ stated in Theorem 2.1 is not reflected by this spectral analysis. In the next section we find new bounds for the spectrum of K_3 which improve upon the existing results.

3. Spectral estimates. In this section we give new bounds for the eigenvalues of K_3 and distinguish between the matrix arising at a generic IP iteration and the matrix arising asymptotically or in the limit of the IP method. Therefore, first we only assume strict positivity of x and z. Then, we suppose that the assumptions in Theorem 2.1 hold and that (x, z) is either a positive vector approaching $(\hat{x}, \hat{z})$ or that it coincides with $(\hat{x}, \hat{z})$.

General IP iterations. For positive x and z we fill in the incomplete analysis on the negative eigenvalues of K_3 given in [19]. If the leading block of K_3 is positive definite, an upper bound for the negative eigenvalues can be found in [31, Lemma 2.2], but this analysis does not apply to our case where H is only positive semidefinite. In [18, Proposition 3.2, Proposition 3.3], the authors derive eigenvalue bounds for general saddle point systems with possibly indefinite and also singular (1,1) block. However, their results require that the (1,2) block of the saddle point matrix have full rank; this can be satisfied in our setting by a simple reordering of the blocks. However, they also require strong assumptions on the norm of the (2,2) block, which in our numerical experiments (see section 3.1) do not hold except during the very first few iterations. In the following theorem we determine an upper bound for the negative eigenvalues of K_3 under weaker hypotheses, by exploiting the structure of the blocks.

THEOREM 3.1. *Suppose that H is SPSD, J has full rank and X and Z have positive diagonal entries. Then the negative eigenvalues θ^- of K_3 given in (2.2) satisfy*

$$\theta^- \le \theta_2 = \gamma, \tag{3.1}$$

where γ is the largest negative root of the cubic polynomial

$$p(\theta) = \theta^3 + (x_{\min} - \lambda_{\max})\theta^2 - (x_{\min}\lambda_{\max} + \sigma_{\min}^2 + z_{\max})\theta - \sigma_{\min}^2 x_{\min}, \tag{3.2}$$

with $\gamma > -x_{\min}$.

We explicitly notice that the polynomial $p(\theta)$ above does admit at least one real negative root, since $p(0) = -\sigma_{\min}^2 x_{\min} < 0$ and $p(-x_{\min}) = z_{\max} x_{\min} > 0$.

Proof. Let $\theta < 0$, and let us split the eigenvalue problem

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

into its block equations,

$$Hu + J^T v - Z^{\frac{1}{2}} w = \theta u, \tag{3.3}$$

$$Ju = \theta v, \tag{3.4}$$

$$-Z^{\frac{1}{2}} u - Xw = \theta w. \tag{3.5}$$

Necessarily $u \neq 0$; otherwise (3.4) gives $v = 0$, and by the positive definiteness of Z, (3.3) yields $w = 0$, which is a contradiction. Similarly, w must be nonzero since otherwise, (3.5) implies $u = 0$. We first assume that $u \in \ker(J)$. From (3.4) we infer that $v = 0$. The first and second equations now read

$$Hu - Z^{\frac{1}{2}} w = \theta u, \qquad -Z^{\frac{1}{2}} u - Xw = \theta w.$$

If we determine u from the first equation above, substitute it in the second one, and multiply the resulting equation from the left by w^T, we obtain

$$w^T Z^{\frac{1}{2}} (H - \theta I_n)^{-1} Z^{\frac{1}{2}} w + w^T X w + \theta \|w\|^2 = 0.$$

Thus, using Rayleigh quotient arguments, we obtain $z_{\min}/(\lambda_{\max} - \theta) + x_{\min} + \theta \leq 0$, and

$$\theta \leq \gamma_1 := \frac{1}{2} \left(\lambda_{\max} - x_{\min} - \sqrt{(\lambda_{\max} + x_{\min})^2 + 4 z_{\min}} \right). \tag{3.6}$$

We now suppose $u \notin \ker(J)$, and write $u = u_1 + u_2$, with $u_1 \in \ker(J)$ and $0 \neq u_2 \in \ker(J)^{\perp}$. Moreover, we suppose $\theta > -x_{\min}$ (otherwise, $-x_{\min}$ is the sought after upper bound), so that the matrix $X + \theta I_n$ is also positive definite. From (3.4) and (3.5) we respectively obtain

$$v = \frac{1}{\theta} J u, \qquad w = -(X + \theta I_n)^{-1} Z^{\frac{1}{2}} u.$$

If we substitute in (3.3) and premultiply it by u_1^T and u_2^T, we respectively obtain:

$$u_1^T H(u_1 + u_2) + u_1^T Z^{\frac{1}{2}} (X + \theta I_n)^{-1} Z^{\frac{1}{2}} (u_1 + u_2) - \theta \|u_1\|^2 = 0,$$

$$u_2^T H(u_1 + u_2) + \frac{1}{\theta} \|J u_2\|^2 + u_2^T Z^{\frac{1}{2}} (X + \theta I_n)^{-1} Z^{\frac{1}{2}} (u_1 + u_2) - \theta \|u_2\|^2 = 0.$$

Subtracting the two equations,

$$u_2^T H u_2 - u_1 H u_1 + \frac{1}{\theta} \|J u_2\|^2 + u_2^T Z^{\frac{1}{2}} (X + \theta I_n)^{-1} Z^{\frac{1}{2}} u_2 +$$
$$- u_1^T Z^{\frac{1}{2}} (X + \theta I_n)^{-1} Z^{\frac{1}{2}} u_1 - \theta \|u_2\|^2 + \theta \|u_1\|^2 = 0.$$

Since $-u_1^T H u_1$, $-u_1^T Z^{\frac{1}{2}} (X + \theta I_n)^{-1} Z^{\frac{1}{2}} u_1$ and $\theta \|u_1\|^2$ are nonpositive, it holds

$$u_2^T \left(H + \frac{1}{\theta} J^T J + Z^{\frac{1}{2}} (X + \theta I_n)^{-1} Z^{\frac{1}{2}} - \theta I_n \right) u_2 \geq 0,$$

from which we obtain

$$\left(\lambda_{\max} + \frac{\sigma_{\min}^2}{\theta} + \frac{z_{\max}}{x_{\min} + \theta} - \theta \right) \|u_2\|^2 \geq 0.$$

Dividing by $\|u_2\|^2$ and multiplying by $-\theta(\theta + x_{\min})$, we find that θ satisfies $p(\theta) \geq 0$ where $p(\theta)$ is the cubic polynomial in (3.2). Noting that $p(0) = -\sigma_{\min}^2 x_{\min} < 0$ and $p(-x_{\min}) = z_{\max} x_{\min} > 0$, it follows that $\theta \leq \gamma$, where γ is the largest negative root of $p(\theta)$, and $\gamma > -x_{\min}$. By (3.6) and $\gamma_1 < -x_{\min} < \gamma$, we can conclude that $\theta \leq \max\{\gamma_1, \gamma\} = \gamma$. □

Combining the above result with the bounds given in Theorem 2.2, we obtain

$$\Lambda(K_3) \subseteq I^- \cup I^+ = [\theta_1,\ \theta_2] \cup [\theta_3,\ \theta_4], \tag{3.7}$$

where θ_1 is as in (2.11), and θ_2 is the new estimate in (3.1). Clearly, the estimate obtained depends on the scaling of the optimization problem, as noted also in [19]. Note that θ_2 provides an improved upper bound for the negative eigenvalues, as compared with Theorem 2.2 taken from [19], whose upper bound was simply zero; see section 3.1 for some illustration. On the other hand, we observe that the estimate θ_2 in Theorem 3.1 may be inadequate in the final stage of the IP method, as it goes to zero with $x_{\min}$. This fact can also be appreciated by writing the polynomial $p(\theta)$ in the theorem statement as

$$p(\theta) = (\theta + x_{\min})(\theta^2 - \lambda_{\max}\theta - \sigma_{\min}^2 - z_{\max}) + z_{\max} x_{\min},$$

so that $p(\theta)$ differs by $z_{\max} x_{\min}$ from a polynomial having $-x_{\min}$ as one of its roots. Such a property shows the inadequacy of this technique to derive spectral estimates at later stages of the IP iterations, when the coefficient matrix remains fairly well conditioned.

Asymptotic IP iterations and limit point. The bounds in (3.7) are meaningful as long as (x, z) are either early or middle stage iterates of the IP method, or (x, z) are late-stage iterates and K_3 tends to singularity. However, if (x, z) approaches a solution $(\hat{x}, \hat{z})$ satisfying the conditions in Theorem 2.1, then the bounds are unsatisfactory. Indeed, $\hat{K}_3$ is nonsingular whereas the upper negative eigenvalue θ_2 tends to 0 as $x_{\min}$ tends to 0, and so does the lower bound θ_3 on the positive eigenvalues if $\lambda_{\min} = 0$. We thus make a further step and focus on the case when $\hat{K}_3$ is nonsingular. It is therefore useful to analyze the assumptions made in Theorem 2.1. Considering the partitioning in (2.5) we can write

$$J = \begin{bmatrix} J_{\mathcal{A}} & J_{\mathcal{I}} \end{bmatrix}, \qquad \begin{bmatrix} J^T & -I_{\mathcal{A}} \end{bmatrix} = \begin{bmatrix} J_{\mathcal{A}}^T & -I_q \\ J_{\mathcal{I}}^T & 0 \end{bmatrix},$$

with $J_{\mathcal{A}} \in \mathbb{R}^{m \times q}$ and $J_{\mathcal{I}} \in \mathbb{R}^{m \times (n-q)}$. The LICQ condition is satisfied at $\hat{x}$ if and only if $J_{\mathcal{I}}^T$ has full column rank. This fact implies that $J_{\mathcal{I}}$ is a "fat" or square matrix, i.e. $q \leq n - m$, and that $\sigma_{\min}(J_{\mathcal{I}}) > 0$.

Concerning condition (2.7), we have that $\ker(\widehat{Z}) = \{(0,y) \in \mathbb{R}^n \mid y \in \mathbb{R}^{(n-q)}\}$, and the vectors of $\ker(J) \cap \ker(\widehat{Z})$ are of the form $(0,y)$ with $y \in \ker(J_{\mathcal{I}})$. If $q = n-m$ then $J_{\mathcal{I}}$ is square and $\ker(J_{\mathcal{I}}) = \{0\}$; thus $\ker(J) \cap \ker(\widehat{Z}) = \{0\}$ and (2.7) is met. Otherwise, if $q < n-m$, then $\ker(J) \cap \ker(\widehat{Z})$ is a nontrivial subspace and condition (2.7) is equivalent to

$$\min_{0 \neq x \in \ker(J) \cap \ker(\widehat{Z})} \frac{x^T H x}{x^T x} = \lambda^* > 0. \tag{3.8}$$

Using the above properties, we prove nontrivial and sharp bounds for K_3 in the late stage of the IP method and for $\hat{K}_3$. To this end, the following technical lemma is needed. It provides bounds for the singular values of a matrix B, which will be used for later estimates; its proof is postponed to the Appendix.

LEMMA 3.2. *Suppose that $\hat{x}$ and $\hat{z}$ are strictly complementary, and $\mathcal{A}$ and $\mathcal{I}$ are the index sets of active and inactive bounds at $\hat{x}$ defined in (2.4). Further, suppose that $\hat{x}$ and $\hat{z}$ are partitioned as in (2.5), the LICQ condition is satisfied at $\hat{x}$, and (2.7) holds. Let $Z_{\mathcal{A}} \in \mathbb{R}^{q \times q}$ be a diagonal positive definite matrix and*

$$B = \begin{bmatrix} J_{\mathcal{A}} & J_{\mathcal{I}} \\ -Z_{\mathcal{A}}^{\frac{1}{2}} & 0 \end{bmatrix}. \tag{3.9}$$

Then

$$\begin{aligned}
\sigma^2_{\min}(B) &\geq \frac{1}{2}\left(\chi - \sqrt{\chi^2 - 4\sigma^2_{\min}(J_{\mathcal{I}})(z_{\mathcal{A}})_{\min}}\right), \\
\sigma^2_{\max}(B) &\leq \frac{1}{2}\left((z_{\mathcal{A}})_{\max} + \sigma^2_{\max} + \sqrt{\left((z_{\mathcal{A}})_{\max} - \sigma^2_{\max}\right)^2 + 4(z_{\mathcal{A}})_{\max}\sigma^2_{\max}(J_{\mathcal{A}})}\right) \\
&\leq \sigma^2_{\max} + (z_{\mathcal{A}})_{\max},
\end{aligned}$$

with $\chi = \sigma^2_{\max}(J_{\mathcal{A}}) + \sigma^2_{\min}(J_{\mathcal{I}}) + (z_{\mathcal{A}})_{\min}$.

The following theorem provides bounds for all eigenvalues of K_3 under the stated assumptions; these bounds are based on perturbation theory results for symmetric matrices and on estimates in [18, 22].

THEOREM 3.3. *Let H be SPSD with nontrivial null space, $\hat{x}$ and $\hat{z}$ strictly complementary, $\mathcal{A}$ and $\mathcal{I}$ be the index sets of active and inactive bounds at $\hat{x}$ defined in (2.4). Further, suppose that the cardinality of $\mathcal{A}$ is equal to q, $\hat{x}$ and $\hat{z}$ are partitioned as in (2.5), the LICQ condition is satisfied at $\hat{x}$, and condition (2.7) holds. Let x and z be sufficiently close to $\hat{x}$ and $\hat{z}$ and be such that $x = (x_{\mathcal{A}},\ x_{\mathcal{I}})$, $z = (z_{\mathcal{A}},\ z_{\mathcal{I}})$ with $x_{\mathcal{A}} \geq 0$, $x_{\mathcal{I}} > 0$, $z_{\mathcal{A}} > 0$, $z_{\mathcal{I}} \geq 0$. Then*

$$\Lambda(K_3) \subseteq [\mu_1,\ \mu_2] \cup [\mu_3,\ \mu_4],$$

where $\mu_1, \mu_2 < 0$ and $\mu_3, \mu_4 > 0$ are given by

$$\begin{aligned}
\mu_1 &= \min\left\{-(x_{\mathcal{I}})_{\max},\ \frac{1}{2}\left(\lambda_{\min} - \sqrt{\lambda^2_{\min} + 4\sigma^2_{\max}(B)}\right)\right\} - \max\left\{(x_{\mathcal{A}})_{\max},\ \sqrt{(z_{\mathcal{I}})_{\max}}\right\}, \\
\mu_2 &= \max\left\{-(x_{\mathcal{I}})_{\min},\ \frac{1}{2}\left(\lambda_{\max} - \sqrt{\lambda^2_{\max} + 4\sigma^2_{\min}(B)}\right)\right\} + \sqrt{(z_{\mathcal{I}})_{\max}}, \\
\mu_3 &= \mu_3^* - (x_{\mathcal{A}})_{\max}, \\
\mu_4 &= \frac{1}{2}\left(\lambda_{\max} + \sqrt{\lambda^2_{\max} + 4\sigma^2_{\max}(B)}\right) + \sqrt{(z_{\mathcal{I}})_{\max}}.
\end{aligned}$$

If $q < n-m$, the scalar μ_3^* is the smallest positive root of the cubic equation

$$\mu^3 - \lambda_{\max}\mu^2 - \sigma_{\min}^2(B)\mu + \lambda^*\sigma_{\min}^2(B) = 0,$$

where λ^* is defined as in (3.8)*. If $q = n-m$ we have instead*

$$\mu_3^* = \frac{1}{2}\left(\lambda_{\min} + \sqrt{\lambda_{\min}^2 + 4\sigma_{\min}^2(B)}\right).$$

Once again, we explicitly notice that the polynomial $p(\mu) = \mu^3 - \lambda_{\max}\mu^2 - \sigma_{\min}^2(B)\mu + \lambda^*\sigma_{\min}^2(B)$ does admit a positive real root, since $p(0) > 0$ and $p(\sigma_{\min}) < 0$.

Proof. We write K_3 in extended form

$$K_3 = \begin{bmatrix} \multicolumn{2}{c}{\multirow{2}{*}{H}} & J_{\mathcal{A}}^T & -Z_{\mathcal{A}}^{\frac{1}{2}} & 0 \\ & & J_{\mathcal{I}}^T & 0 & -Z_{\mathcal{I}}^{\frac{1}{2}} \\ J_{\mathcal{A}} & J_{\mathcal{I}} & 0 & 0 & 0 \\ -Z_{\mathcal{A}}^{\frac{1}{2}} & 0 & 0 & -X_{\mathcal{A}} & 0 \\ 0 & -Z_{\mathcal{I}}^{\frac{1}{2}} & 0 & 0 & -X_{\mathcal{I}} \end{bmatrix},$$

with $X_{\mathcal{A}} = \mathrm{diag}(x_{\mathcal{A}}) \in \mathbb{R}^{q\times q}$, $X_{\mathcal{I}} = \mathrm{diag}(x_{\mathcal{I}}) \in \mathbb{R}^{n-q\times n-q}$, $Z_{\mathcal{A}} = \mathrm{diag}(z_{\mathcal{A}}) \in \mathbb{R}^{q\times q}$, $Z_{\mathcal{I}} = \mathrm{diag}(z_{\mathcal{I}}) \in \mathbb{R}^{n-q\times n-q}$, and observe that

$$\begin{aligned} K_3 &= \widetilde{K}_3 + \Delta_K \qquad (3.10) \\ &= \begin{bmatrix} \multicolumn{2}{c}{\multirow{2}{*}{H}} & J_{\mathcal{A}}^T & -Z_{\mathcal{A}}^{\frac{1}{2}} & 0 \\ & & J_{\mathcal{I}}^T & 0 & 0 \\ J_{\mathcal{A}} & J_{\mathcal{I}} & 0 & 0 & 0 \\ -Z_{\mathcal{A}}^{\frac{1}{2}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -X_{\mathcal{I}} \end{bmatrix} + \begin{bmatrix} \multicolumn{2}{c}{\multirow{2}{*}{0}} & 0 & 0 & 0 \\ & & 0 & 0 & -Z_{\mathcal{I}}^{\frac{1}{2}} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -X_{\mathcal{A}} & 0 \\ 0 & -Z_{\mathcal{I}}^{\frac{1}{2}} & 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Standard perturbation arguments for symmetric matrices ensure that an eigenvalue θ of K_3 satisfies (see, e.g., [13, Theorem 8.1.5])

$$\lambda_i(\widetilde{K}_3) + \lambda_{\min}(\Delta_K) \le \theta \le \lambda_i(\widetilde{K}_3) + \lambda_{\max}(\Delta_K), \quad i = 1, \dots 2n+m. \qquad (3.11)$$

Thus, estimates for θ can be derived from spectral information on $\widetilde{K}_3$ and Δ_K, where

$$\lambda_{\min}(\Delta_K) = -\max\left\{(x_{\mathcal{A}})_{\max}, \sqrt{(z_{\mathcal{I}})_{\max}}\right\}, \quad \lambda_{\max}(\Delta_K) = \sqrt{(z_{\mathcal{I}})_{\max}}.$$

As for $\widetilde{K}_3$, we have that $\Lambda\left(\widetilde{K}_3\right) = \Lambda\left(-X_{\mathcal{I}}\right) \cup \Lambda\left(\check{K}\right)$, where $\check{K}$ is the saddle point matrix

$$\check{K} = \begin{bmatrix} H & B^T \\ B & 0 \end{bmatrix},$$

with B given in (3.9). By Lemma 3.2 we know that B^T has full column rank. Moreover, $\ker(B) = \ker(J) \cap \ker(\widehat{Z})$, where we recall that $\widehat{Z} = \mathrm{diag}(\hat{z})$ and $\hat{z}$ is defined in (2.5). In addition, by (2.7) either $\ker(B) = \{0\}$ (if B is square, i.e. $q = n-m$) or H is positive definite on $\ker(B)$. Thus, $\check{K}$ satisfies the hypothesis of [18, Proposition 2.2], and the expressions for μ_1, μ_2 and μ_4 are a direct consequence of that result.

A slightly different approach is needed to obtain μ_3. We consider the principal submatrix $\bar{K}$ of K_3 obtained by taking its first $n+m+q$ rows and columns, i.e.

$$\bar{K} = \begin{bmatrix} \multicolumn{2}{c}{\multirow{2}{*}{H}} & J_{\mathcal{A}}^T & -Z_{\mathcal{A}}^{\frac{1}{2}} \\ & & J_{\mathcal{I}}^T & 0 \\ J_{\mathcal{A}} & J_{\mathcal{I}} & 0 & 0 \\ -Z_{\mathcal{A}}^{\frac{1}{2}} & 0 & 0 & -X_{\mathcal{A}} \end{bmatrix}.$$

It holds that K_3 has n positive and $n+m$ negative eigenvalues, and $\bar{K}$ has n positive eigenvalues and $m+q$ negative ones [19, Lemma 3.8]. Using interlacing properties of the eigenvalues and again the standard perturbation bounds for symmetric matrices, we infer

$$\lambda_{\min}^{+}(K_3) \geq \lambda_{\min}^{+}(\bar{K}) \geq \lambda_{\min}^{+}(\check{K}) - (x_{\mathcal{A}})_{\max},$$

where the symbol $\lambda_{\min}^{+}(\cdot)$ indicates the smallest positive eigenvalue of a matrix. If $q = n - m$ we can use again [18, Proposition 2.2] to obtain the expression of μ_3^*. If $q < n - m$, since we supposed H singular we can instead use the lower bound of the positive eigenvalues of $\check{K}$ given in [22, Theorem 2] to obtain the final result. □

It is interesting to observe that, whenever x and z are sufficiently close to $\hat{x}$ and $\hat{z}$, then $(x_{\mathcal{A}})_{\max}$ and $(z_{\mathcal{I}})_{\max}$ are small enough to guarantee that the intervals $[\mu_1,\ \mu_2]$ and $[\mu_3,\ \mu_4]$ are nontrivial, i.e., μ_2 is strictly negative and μ_3 is strictly positive.

Theorem 3.3 covers both the case where (x, z) is strictly positive and close enough to $(\hat{x}, \hat{z})$, and the case where $(x, z) = (\hat{x}, \hat{z})$. Thus, these bounds are valid for the matrices K_3 occurring at the late stage of the IP method, and also for $\hat{K}_3$. The proof of this theorem relies on the perturbation theory for symmetric eigenvalue problems, and involves $\widetilde{K}_3$ and the scalars $(x_{\mathcal{A}})_{\max}$ and $(z_{\mathcal{I}})_{\max}$ which approach zero when (x, z) tends to $(\hat{x}, \hat{z})$. Hence, the smaller $(x_{\mathcal{A}})_{\max}$ and $(z_{\mathcal{I}})_{\max}$, the closer $\mu_1,\ \mu_2,\ \mu_3,\ \mu_4$ are to the spectral bounds for $\widetilde{K}_3$, for which bounds are available [18, 22].

Remark 3.4. In Theorem 3.3, for the case $q < n - m$, the value of μ_3^* relies on results from [22] and it holds for H singular. If H is nonsingular and $q < n - m$ then it holds that $\mu_3^* = \max\{\lambda_{\min}, \gamma\}$, where γ is the smallest positive root of the cubic polynomial $p_3(\mu) = \mu^3 - (\lambda_{\min} + \lambda^*)\,\mu^2 + \left(\lambda_{\min}\lambda^* - \lambda_{\max}^2 - \sigma_{\min}(B)^2\right)\mu + \lambda^*\sigma_{\min}^2(B)$. This follows from applying [18, Proposition 2.2] and [29, Lemma 2.1] to the matrix $\check{K}$ in the proof of Theorem 3.3. Furthermore, note that μ_3^* remains bounded away from zero as the IP iteration converges, as the polynomial coefficients do not depend on $(x_{\mathcal{A}})_{\max}$ and $(z_{\mathcal{I}})_{\max}$. All other bounds given by Theorem 3.3 still hold when H is positive definite.

3.1. Validation of the spectral bounds for K_3. We analyze the quality of our bounds by first using two examples with small matrices. In the first case, H is positive definite (see Remark 3.4), while in the second H is only positive semidefinite.

We want to stress that the new bounds are chiefly of theoretical interest, because they emphasize how the eigenvalues of the matrix depend on the various matrices appearing in the blocks, and thus, in turn, on the problem data. The bounds can be computed for small scale problems; we do not advise their explicit computation in the large scale case, however knowing that these bounds exist may be helpful in guiding the selection of acceleration procedures.

Example 1. Given positive scalars $\lambda,\ \sigma,\ \rho$, let

$$H = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}, \quad J^T = \begin{bmatrix} 0 \\ \sigma \end{bmatrix}, \quad x = \begin{bmatrix} 0 \\ \rho \end{bmatrix}, \quad z = \begin{bmatrix} \sigma \\ 0 \end{bmatrix}, \text{ so that } B = \begin{bmatrix} 0 & \sigma \\ -\sigma & 0 \end{bmatrix}.$$

The characteristic polynomial of K_3 is given by $\pi(\theta) = (\theta + \rho)\left(\theta^2 - \lambda\theta - \sigma^2\right)^2$. The eigenvalues of K_3 are $-\rho$, $\frac{1}{2}\left(\lambda - \sqrt{\lambda^2 + 4\sigma^2}\right)$, $\frac{1}{2}\left(\lambda + \sqrt{\lambda^2 + 4\sigma^2}\right)$, and the bounds in Theorem 3.3 are sharp (note that $q = n - m$). □

In this second example, we have $q < n - m$ and show that the estimates μ_1 and μ_3 can be sharp.

EXAMPLE 2. Given positive scalars $\lambda_{\max}$, λ^*, σ, $x_{\min}$ $x_{\max}$, $\lambda_{\max} > \lambda^*$, let

$$H = \begin{bmatrix} \lambda_{\max} - \lambda^* & \sqrt{\lambda^*(\lambda_{\max} - \lambda^*)} & 0 \\ \sqrt{\lambda^*(\lambda_{\max} - \lambda^*)} & \lambda^* & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad J^T = \begin{bmatrix} 0 \\ 0 \\ \sigma \end{bmatrix},$$

$$z = \begin{bmatrix} \sigma \\ 0 \\ 0 \end{bmatrix}, \quad x = \begin{bmatrix} 0 \\ x_{\min} \\ x_{\max} \end{bmatrix}, \text{ so that } B = \begin{bmatrix} 0 & 0 & \sigma \\ -\sigma & 0 & 0 \end{bmatrix}.$$

The characteristic polynomial of K_3 is

$$\pi(\theta) = (\theta + x_{\min})(\theta + x_{\max})(\sigma^2 - \theta^2)(\theta^3 - \lambda_{\max}\theta^2 - \sigma^2\theta + \lambda^*\sigma^2).$$

Since $\lambda_{\min} = 0$ and $q < n - m$, the bounds μ_1 and μ_3 are sharp. □

We then proceed by analyzing the quality of our spectral estimates on benchmark Linear Programming problems:

$$\min_{x \in \mathbb{R}^n} c^T x, \qquad \text{subject to} \quad Jx = b, \quad x \geq 0,$$

where $n = 185$, $m = 129$, $J \in \mathbb{R}^{m \times n}$ is the matrix in LPnetlib/lp_scagr7 [32] with full row rank, b and c are fixed so that the $\hat{x} = (0, 1_{n-q})$ and $\hat{z} = (1_q, 0)$ are exact primal and dual solutions. The value of q is varied and it may affect the fulfillment of the LICQ condition at $\hat{x}$.

The problems were solved with the PDCO solver [27] and sequence of iterates approaching $\hat{x}$ and $\hat{z}$ were computed and stored. Then, for each iterate we formed matrix K_3 letting $H = \rho I_n$ with $\rho = 10^{-6}$; this amounts to applying a primal positive definite regularization. The eigenvalues of the resulting matrices were computed and compared with the bounds given in (3.7) and in Theorem 3.3 with μ_3^* as in Remark 3.4. Regarding the actual computation of the bounds, singular values $\sigma_{\min}(J)$, $\sigma_{\max}(J)$, $\sigma_{\min}(B)$ and $\sigma_{\max}(B)$ were computed while $\lambda_{\min} = \lambda_{\max} = \lambda^* = \rho$.

In our numerical experiments, the known bounds θ_1 and θ_4 from Theorem 2.1 are very similar to our new bounds μ_1 and μ_4 from Theorem 3.3, and both pairs seem to be good approximations of the extreme eigenvalues of K_3. Since a numerical validation of θ_1 and θ_4 was already given in [19], we do not show any of these bounds in the plots. On the other hand, μ_3, μ_2 were plotted only when meaningful, that is only when appearing with positive and negative sign, respectively.

We start by reporting on the accuracy of θ_2 in Theorem 3.1. For this purpose we set $q = n - m$, which makes the matrix $J_\jmath$, and thus B, rank deficient. The absolute value of the largest negative eigenvalue $\lambda_{n+m}(K_3)$ (solid line), and its bound θ_2 are displayed in Figure 3.1, showing that the estimate matches quite well the true eigenvalue of K_3.

Let $\text{fix}(\cdot)$ be the function that rounds its argument to the nearest integer towards zero. We then set $q = \text{fix}((n - m)/2)$ so that the assumptions of Theorem 3.3

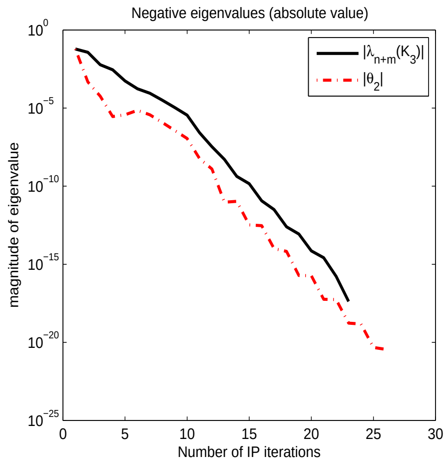

FIG. 3.1. *Negative eigenvalue of K_3 closest to zero (solid line) and its bound at every iteration, $q = n - m$.*

hold; indeed, with this choice the matrix $J_{\mathcal{I}}$, and thus B, have full rank. In the left plot of Figure 3.2, for each iterate on the x-axis, the minimum positive eigenvalue $\lambda_{n+m+1}(K_3)$, and its bounds θ_3 and μ_3 are displayed; θ_3 is a good lower bound and μ_3 is sharp as well during the later ones. Similarly, in the right plot of Figure 3.2 we report the absolute value of the negative eigenvalue $\lambda_{n+m}(K_3)$ closest to zero, along with the bounds θ_2 and μ_2. As expected, μ_2 is sharp during the final iterations, unlike θ_2.

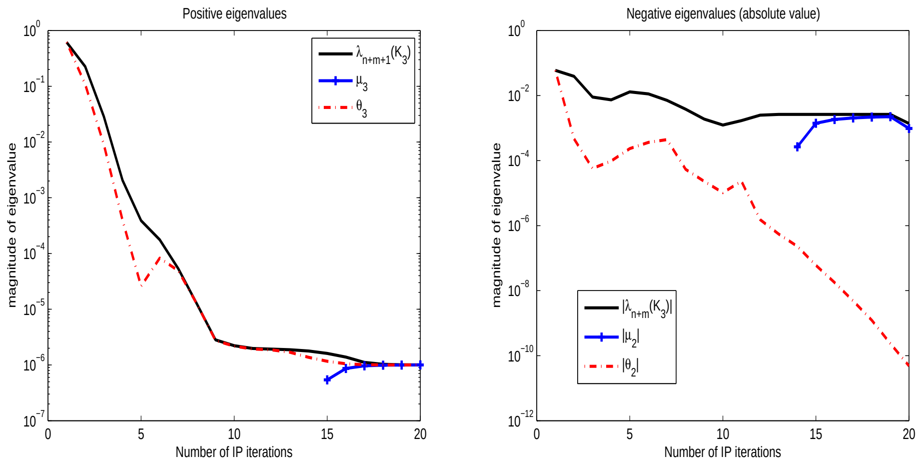

FIG. 3.2. *Eigenvalues of K_3 closest to zero (solid line) and their bounds at every iteration, $q = fix((n - m)/2)$ and H nonsingular. Left: positive eigenvalues. Right: negative eigenvalues.*

It is of interest testing the validity of the lower bound μ_3 to the positive eigenvalues when H is singular. For this reason, let us consider a QP problem where J is the same matrix as before and let the orthonormal columns of V span $\ker(J)$. Then by taking H in the form

$$H = \begin{bmatrix} V & Q \end{bmatrix} \begin{bmatrix} \rho I_{n-m} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V^T \\ Q^T \end{bmatrix}, \tag{3.12}$$

where ρ is positive and $[V\,Q]$ is an orthogonal matrix, we ensure that $\lambda^* = \lambda_{\max}$. It holds that $\lambda_{\min} = 0$ and $\lambda_{\max} = \rho$. Finally, the QP problem is built setting $\rho = 1$, so that $\hat{x} = (0, 1_{n-q})$ and $\hat{z} = (1_q, 0)$ are exact primal and dual solutions with $q = \text{fix}((n-m)/2)$. The QP problem was solved with PDCO and a sequence of iterates approaching $\hat{x}, \hat{z}$ was formed. Figure 3.3 displays the positive eigenvalue $\lambda_{n+m+1}(K_3)$ and the bounds θ_3, μ_3, as the iterations proceed. Since $\lambda^* = \lambda_{\max}$ and $\mu_3^* = \min\{\sigma_{\min}(B), \rho\}$, during the last iterations of the interior point method, μ_3 gets close to μ_3^* and is sharp whereas θ_3 is not representative of the minimum eigenvalue.

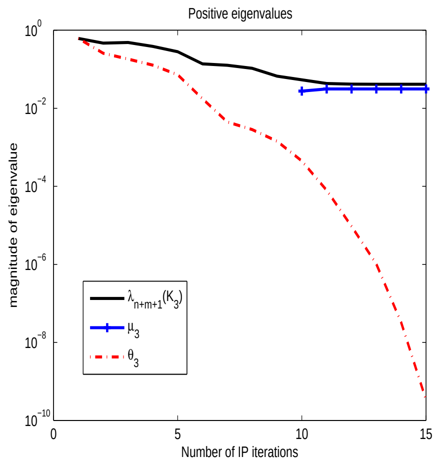

FIG. 3.3. *Minimum positive eigenvalue of* K_3 *and its bounds for* H *singular as in* (3.12), $\rho = 1$, $q = \mathit{fix}((n-m)/2)$.

We then consider larger QP problems from the Maros and Meszaros collection [24], but small enough so that we can still explicitly compute spectral quantities: STCQP2 of dimensions $n = 4097$, $m = 2052$, and CONT-050 of dimensions $n = 2597$, $m = 2401$. In STCQP2, variables x and z were scaled by a factor 10^{-1} and 10^3 respectively. As before, we solved these problems using PDCO and compared the eigenvalues of the matrices generated through the IP procedure with the bounds θ_2, θ_3 from [19], and μ_2 and μ_3 given in section 3; we remark that K_3 is nonsingular at the limit.

Figures 3.4 and 3.5 show the eigenvalues of K_3 closest to zero, namely λ_{m+n} and λ_{m+n+1}, together with their bounds θ_2, θ_3 from [19], and the new bounds μ_2 and μ_3. Our conclusions are analogous to those obtained above: θ_2 and θ_3 are good approximations at early and middle IP iterations; moreover in problem STCQP2 matrix H is positive definite and θ_3 is meaningful also in the late iterations, see left plot of Figure 3.4. In late iterations μ_2 and μ_3 are representative for both problems.

A special case is depicted in the left plot of Figure 3.5 for problem CONT-050 where H is positive definite. In the final stage of the IP method, the estimate θ_3 from [19] satisfies $\theta_3 \approx \lambda_{\min}(H) > 0$, see Theorem 2.2; however, this bound is not tight, as the minimum positive eigenvalue $\lambda_{m+n+1}(K_3)$ is several orders of magnitude larger

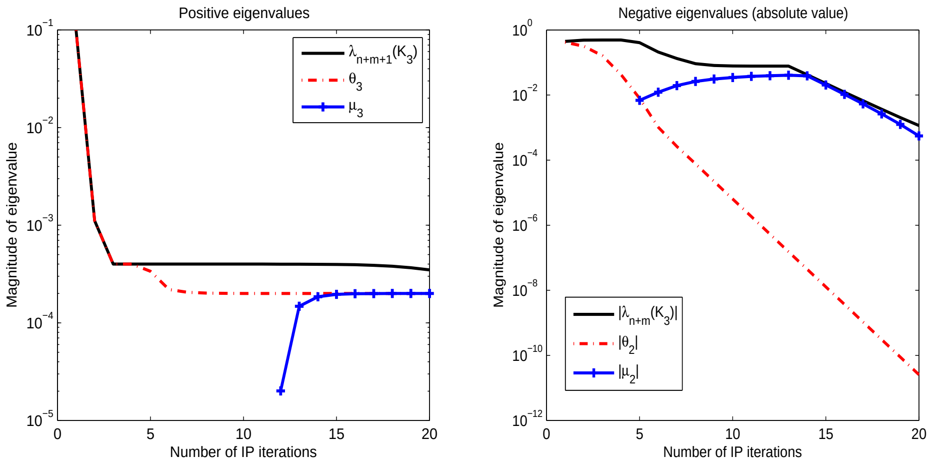

FIG. 3.4. *Problem* STCQP2*: eigenvalues of* K_3 *closest to zero (solid line) and their bounds at every iteration. Left: positive eigenvalues. Right: negative eigenvalues.*

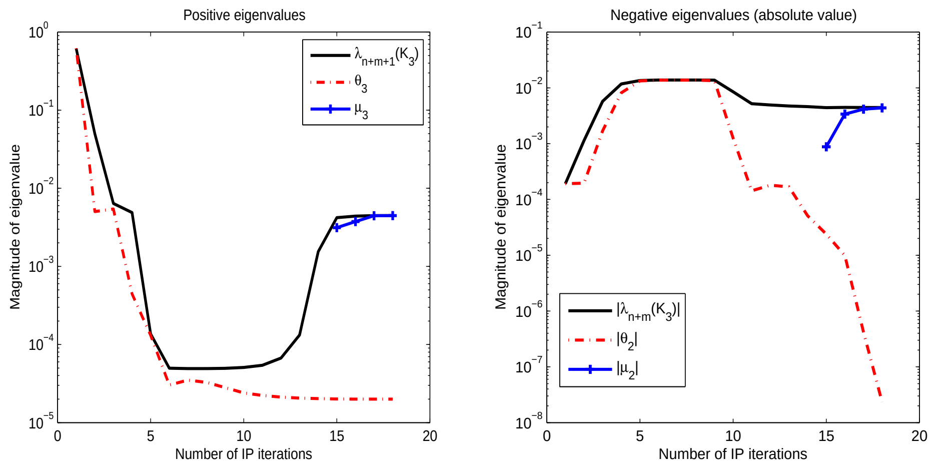

FIG. 3.5. *Problem* CONT-050*: eigenvalues of* K_3 *closest to zero (solid line) and their bounds at every iteration. Left: positive eigenvalues. Right: negative eigenvalues.*

than $\lambda_{\min}(H)$. On the other hand, the new lower bound μ_3 approximates $\lambda_{m+n+1}(K_3)$ accurately (since $q = n - m$, the second expression of μ_3^* from Theorem 3.3 was used.)

4. A block diagonal augmented preconditioner for K_3.

An accurate spectral analysis of the coefficient matrix K_3 must be accompanied by a theoretical discussion of the effectiveness of ideal preconditioners for the given problem. These two parts together represent a necessary starting point for a more computationally-oriented study of effective acceleration strategies. When we first addressed the problem of deepening our knowledge of the unreduced formulation, we realized that specific standard preconditioning strategies were lacking of basic reference spectral information for our setting. We aim to fill this gap in this section.

A variety of preconditioning techniques have been proposed for KKT systems arising in optimization; see the surveys [3, 6, 16]. We provide the spectral analysis of matrix K_3 preconditioned by an augmented block diagonal preconditioner; since the symmetry is preserved preconditioned MINRES can be used as a solver. This choice of preconditioner is motivated by the work in [28], where the effectiveness of augmented block diagonal preconditioners on linear systems arising from IP methods is demonstrated for highly singular (1,1) block and null (2,2) block. To the best of our knowledge there are no results on spectral properties for nonzero (2,2) block in the literature, as is the case for our matrix K_3. This section is devoted to such a study. We shall mainly discuss *ideal* preconditioners, that is preconditioners that lead to spectral intervals of the preconditioned matrix that are independent of the problem data, and small enough so as to ensure fast convergence. To make these "ideal" choices computationally attractive, approximation strategies may have to be performed; see, e.g., [26].

The sparsity pattern and spectral properties of the blocks of K_3 should be carefully exploited when devising the block preconditioner. In particular, in some cases the (1,1) block may be more suitable for preconditioning purposes than the other diagonal blocks, whereas in other settings it may be the opposite. Therefore, in the following we shall consider two block reorderings, namely

$$\left[\begin{array}{c|cc} H & J^T & -Z^{\frac{1}{2}} \\ \hline J & 0 & 0 \\ -Z^{\frac{1}{2}} & 0 & -X \end{array}\right], \quad \left[\begin{array}{cc|c} 0 & 0 & J \\ 0 & -X & -Z^{\frac{1}{2}} \\ \hline J^T & -Z^{\frac{1}{2}} & H \end{array}\right]. \tag{4.1}$$

For the sake of generality, we use the following general notation for the saddle point matrix

$$M = \begin{bmatrix} F & G^T \\ G & -C \end{bmatrix}, \tag{4.2}$$

where $0 \neq F \in \mathbb{R}^{n_1 \times n_1}$, $C \in \mathbb{R}^{n_2 \times n_2}$ are SPSD matrices and $0 \neq G \in \mathbb{R}^{n_2 \times n_1}$. Each of the given results will be specialized to our setting in a subsequent corollary.

Depending on the block reordering matrix G in (4.2) may be either tall or fat. In either case, it must hold that $\ker(F) \cap \ker(G) = \{0\} = \ker(G^T) \cap \ker(C)$ to ensure nonsingularity of M (see, e.g., [3]). Since F is SPSD, we consider the following block diagonal preconditioner

$$P_D = \begin{bmatrix} F + G^T W^{-1} G & 0 \\ 0 & W \end{bmatrix}, \tag{4.3}$$

which is based on the augmentation of the (1,1) block of M by a symmetric and positive definite matrix $W \in \mathbb{R}^{n_2 \times n_2}$; due to the hypotheses on F and G, the matrix $F + G^T W^{-1} G$ is symmetric and positive definite. In the case where $C = 0$ and $n_2 \leq n_1$ the preconditioner P_D was originally proposed in [20] and later extended in [28, 5]. Here we assume that C is SPSD and we allow any relation between n_1 and n_2. We stress that P_D is an *ideal* preconditioner, and that further efforts need be put into selecting computationally effective variants; nonetheless, spectral intervals obtained with P_D will drive the choices of more economical approximations. Our spectral estimates no longer apply when approximations to P_D are performed, although we expect small perturbations of these estimates as the approximate preconditioner slightly deviates from the ideal one.

We first give spectral bounds for the preconditioned matrix and then discuss the choice of the matrix W.

4.1. Spectral estimates for G^T of full column rank. We first consider the case when $\ker(G^T) = \{0\}$.

THEOREM 4.1. *Let $F \in \mathbb{R}^{n_1 \times n_1}$ and $C \in \mathbb{R}^{n_2 \times n_2}$ be SPSD, $W \in \mathbb{R}^{n_2 \times n_2}$ be symmetric and positive definite, $G \in \mathbb{R}^{n_2 \times n_1}$. Suppose that the matrix M in (4.2) is nonsingular and that* $\ker(G^T) = \{0\}$.

Let $c_0 = \lambda_{\min}\left(W^{-1}C\right) \geq 0$, $c_1 = \lambda_{\max}\left(W^{-1}C\right)$, $g_0 = \lambda_{\min}\left(W^{-1}GG^T\right) > 0$ and $g_1 = \lambda_{\max}\left(W^{-1}GG^T\right)$. Then, given P_D in (4.3), it holds

$$\Lambda(P_D^{-1}M) \subseteq I^- \cup I^+ = [\xi_1,\ \xi_2] \cup [\xi_3,\ 1]\,,$$

where

$$\xi_1 = \frac{(1-c_1)s - c_1 - \sqrt{((1-c_1)s-c_1)^2 + 4(1+c_1 s)(1+s)}}{2(1+s)},$$
$$\xi_2 = \frac{1 - tc_0 - \sqrt{(1+tc_0)^2 + 4t(t-1)}}{2t},$$
$$\xi_3 = \frac{(1-c_1)s - c_1 + \sqrt{((1-c_1)s-c_1)^2 + 4(1+c_1 s)(1+s)}}{2(1+s)},$$

with $s = \dfrac{\lambda_{\min}(F)}{g_1} \geq 0$ and $t = 1 + \dfrac{g_0}{\|F\|}$.

Moreover, if ℓ denotes the nullity of G, then $P_D^{-1}M$ has the eigenvalue 1 with multiplicity ℓ.

Proof. Consider the generalized eigenvalue problem $M \begin{bmatrix} u \\ v \end{bmatrix} = \theta P_D \begin{bmatrix} u \\ v \end{bmatrix}$, i.e.

$$Fu + G^T v = \theta\left(F + G^T W^{-1} G\right) u, \tag{4.4}$$
$$Gu - Cv = \theta W v \tag{4.5}$$

and first suppose that $\theta > 0$. Since any vector $(u, 0)$ with $u \in \ker(G)$ satisfies (4.4) and (4.5) with $\theta = 1$, $P_D^{-1}M$ has the eigenvalue one with multiplicity ℓ.

Suppose now $u \notin \ker(G)$. Since $\theta W + C$ is positive definite, we eliminate v from (4.5) and substitute it into (4.4). Further, we premultiply the resulting equation by u^T and obtain

$$u^T(F + G^T\left(\theta W + C\right)^{-1} G)u = \theta u^T\left(F + G^T W^{-1} G\right) u. \tag{4.6}$$

By using the inequality $(\theta W + C)^{-1} \preceq \dfrac{1}{\theta} W^{-1}$, and after some rearrangement we get

$$(1-\theta) u^T\left(\theta F + (1+\theta) G^T W^{-1} G\right) u \geq 0.$$

Noting that $\left(\theta F + (1+\theta) G^T W^{-1} G\right)$ is positive definite, it follows $\theta \leq 1$.

To show the lower bound for the positive eigenvalues, we reformulate (4.6) as

$$u^T G^T W^{-\frac{1}{2}}\left(\theta I_{n_2} - \left(\theta I_{n_2} + \widetilde{C}\right)^{-1}\right) W^{-\frac{1}{2}} G u = (1-\theta) u^T F u,$$

where $\widetilde{C} = W^{-\frac{1}{2}} C W^{-\frac{1}{2}}$ and note that

$$F \succeq \lambda_{\min}(F) I_{n_1}, \quad \theta I_{n_2} - \left(\theta I_{n_2} + \widetilde{C}\right)^{-1} \preceq \left(\theta - \frac{1}{\theta + c_1}\right) I_{n_2}, \quad G^T W^{-1} G \preceq g_1 I_{n_1},$$

where the last inequality follows form the fact that $G^T W^{-1} G$ and $W^{-1} G G^T$ admit the same maximum eigenvalue. Using these inequalities and dividing by $\|u\|^2$, we find that θ satisfies

$$\left(\theta - \frac{1}{\theta + c_1}\right) g_1 \geq (1 - \theta) \lambda_{\min}(F),$$

which is equivalent to $\theta^2 (1+s) + \theta ((c_1 - 1)s + c_1) - (1 + c_1 s) \geq 0$, with $s = \lambda_{\min}(F)/g_1$. The expression of the left extreme of I^+ readily follows.

Let now $\theta < 0$. Equation (4.4) can be rewritten as $\left((1-\theta)F - \theta G^T W^{-1} G\right) u = -G^T v$, and the matrix on the left-hand side is now positive definite. This implies that $v \neq 0$ otherwise we would also have $u = 0$. If we eliminate u from (4.4), substitute into (4.5) and premultiply the resulting equation by $W^{-\frac{1}{2}}$ we get

$$\widetilde{G} \left((1-\theta)F - \theta \widetilde{G}^T \widetilde{G}\right)^{-1} \widetilde{G}^T w + \widetilde{C} w = -\theta w, \tag{4.7}$$

where $v = W^{-\frac{1}{2}} w$, $\widetilde{C} = W^{-\frac{1}{2}} C W^{-\frac{1}{2}}$, $\widetilde{G} = W^{-\frac{1}{2}} G$. It holds

$$\begin{aligned} \widetilde{G} \left((1-\theta)F - \theta \widetilde{G}^T \widetilde{G}\right)^{-1} \widetilde{G}^T &\preceq \widetilde{G} \left((1-\theta)\lambda_{\min}(F) I_{n_1} - \theta \widetilde{G}^T \widetilde{G}\right)^{-1} \widetilde{G}^T \\ &\preceq \frac{g_1}{(1-\theta)\lambda_{\min}(F) - \theta g_1} I_{n_2}, \end{aligned}$$

since the eigenvalues of $\widetilde{G} \left(\|F\| (1-\theta) I_{n_1} - \theta \widetilde{G}^T \widetilde{G}\right)^{-1} \widetilde{G}^T$ are of the form $\sigma^2/((1-\theta)\|F\| - \theta\sigma^2)$, with σ being a singular value of $\widetilde{G}^T$. We now multiply (4.7) by w^T and use the above inequality as well as $\widetilde{C} \preceq c_1 I_{n_2}$. Then, after some rearrangements we obtain

$$\theta^2 (1+s) + \theta ((c_1 - 1)s + c_1) - (1 + c_1 s) \leq 0,$$

with s as above, from which the lower bound for the negative eigenvalues follows.

Finally, we prove the upper bound for the negative eigenvalues. We have $\widetilde{C} \succeq c_0 I_{n_2}$ and

$$\begin{aligned} \widetilde{G} \left((1-\theta)F - \theta \widetilde{G}^T \widetilde{G}\right)^{-1} \widetilde{G}^T &\succeq \widetilde{G} \left(\|F\| (1-\theta) I_{n_1} - \theta \widetilde{G}^T \widetilde{G}\right)^{-1} \widetilde{G}^T \\ &\succeq \frac{g_0}{(1-\theta)\|F\| - \theta g_0} I_{n_2}, \end{aligned}$$

using again the form of the eigenvalues of $\widetilde{G} \left(\|F\| (1-\theta) I_{n_1} - \theta \widetilde{G}^T \widetilde{G}\right)^{-1} \widetilde{G}^T$. We now multiply (4.7) from the left by w^T. Using the above inequalities and dividing by $\|w\|^2$, we obtain

$$\frac{g_0}{(1-\theta)\|F\| - \theta g_0} + c_0 \leq -\theta,$$

which implies $t\theta^2 - (1 - tc_0)\,\theta - t + 1 - c_0 \geq 0$, with $t = 1 + \dfrac{g_0}{\|F\|}$, and the stated expression for the right extreme of I^-. $\square$

When $C = 0$, the bounds of Theorem 4.1 reduce to $I^+ = \{1\}$ and $I^- = \left[-\frac{1}{1+s},\, -\frac{t-1}{t}\right]$; moreover, the interval I^+ and the left extreme of I^- coincide with the corresponding bounds in [20, Theorem 2.2], when F is singular, i.e. $s = 0$, while the right extreme is an improvement of that in [20, Theorem 2.2]; we refer the reader to [25, section 4] for more details.

We next specialize Theorem 4.1 to our setting. The full rank assumption on G^T is satisfied with the second of the two reorderings in (4.1). The result follows by noticing that here $s = 0$.

COROLLARY 4.2. *Assume that* $\begin{bmatrix} J \\ -Z^{\frac{1}{2}} \end{bmatrix}$ *is full column rank. Let*

$$M = \begin{bmatrix} F & G^T \\ G & -C \end{bmatrix} := -\left[\begin{array}{cc|c} 0 & 0 & J \\ 0 & -X & -Z^{\frac{1}{2}} \\ \hline J^T & -Z^{\frac{1}{2}} & H \end{array}\right],$$

and P_D defined accordingly. Let $c_0 = \lambda_{\min}(W^{-1}H)$, $c_1 = \lambda_{\max}(W^{-1}H)$, $g_0 = \lambda_{\min}(W^{-1}(J^TJ + Z))$, $g_1 = \lambda_{\max}(W^{-1}(J^TJ + Z))$. Then

$$\Lambda(P_D^{-1}M) \subset [\xi_1, \xi_2] \cup [\xi_3, 1],$$

where $\xi_1 = \frac{1}{2}(-c_1 - \sqrt{c_1^2 + 4})$, $\xi_2 = \frac{1}{2t}(1 - tc_0 - \sqrt{(1 + tc_0)^2 + 4t(t-1)})$, $\xi_3 = \frac{1}{2}(-c_1 + \sqrt{c_1^2 + 4})$, and $t = 1 + g_0/\max_{1\leq i\leq n}|x_i|$. Moreover, if ℓ denotes the nullity of $[J^T, -Z^{\frac{1}{2}}]$, then $P_D^{-1}M$ has the eigenvalue 1 with multiplicity ℓ.

4.2. Spectral estimates for G^T column-rank deficient. The proofs of Theorem 4.1 and of its corollary rely on the full rank assumption of $G^T = [J^T, -Z^{\frac{1}{2}}]^T$. In the case when the first ordering in (4.1) is used, this assumption no longer holds, since the matrix $[J^T, -Z^{\frac{1}{2}}]$ is not full-column rank in general. Nonetheless, we can still provide insightful spectral bounds for the preconditioned problem. Once again, we first state the result in general, and then specialize it to our setting.

THEOREM 4.3. *Let $F \in \mathbb{R}^{n_1\times n_1}$ and $C \in \mathbb{R}^{n_2\times n_2}$ be SPSD, $G \in \mathbb{R}^{n_2\times n_1}$, M in (4.2) nonsingular. Let $W \in \mathbb{R}^{n_2\times n_2}$ be symmetric and positive definite and $c_1 = \lambda_{\max}\left(W^{-1}C\right)$. Suppose that G^T has a nontrivial null space, define*

$$\min_{0\neq x\in\ker(G^T)} \frac{x^TCx}{x^TWx} = c^* > 0,$$

and let g_+ be the minimum positive eigenvalue of $W^{-1}GG^T$. For P_D as in (4.3), it holds

$$\Lambda(P_D^{-1}M) \subseteq I^- \cup I^+ = [\xi_1,\ \min\{\eta, \xi_2\}] \cup [\xi_3,\ 1],$$

where ξ_1, ξ_2 and ξ_3 are given in Theorem 4.1, and $\eta \geq -c^*$ is the largest negative root of the cubic polynomial

$$q(\theta) = t_+\theta^3 + \theta^2\left((c_1 + c^*)t_+ - 1\right) - \theta\left(c_1 + c^* - 1 + t_+\right) - (t_+ - 1)\,c^*, \tag{4.8}$$

with $t_+ = 1 + \dfrac{g_+}{\|F\|}$.

Proof. We only need to prove the upper bound for the negative eigenvalues. From the proof of Theorem 4.1 we infer that if θ is a negative eigenvalue of $P_D^{-1}M$, then $\theta \leq \xi_2 = -c_0$.

Consider equation (4.7) and suppose that $w \in \ker(\widetilde{G}^T)$. Hence, we have $w^T\widetilde{C}w = -\theta\,\|w\|^2$, which implies $\theta \leq -c^*$.

We now suppose $\theta > -c^*$, (hence, $w \notin \ker(\widetilde{G}^T)$), and write $w = w_0 + w_1$, with $w_0 \in \ker(\widetilde{G}^T)$ and $w_1 \in \ker(\widetilde{G}^T)^\perp$. We premultiply equation (4.7) by w_0^T to get

$$w_0^T\widetilde{C}w_1 = -w_0^T\widetilde{C}w_0 - \theta\,\|w_0\|^2 \leq -\left(1+\frac{\theta}{c^*}\right)w_0^T\widetilde{C}w_0,$$

from which, using the inequality $w_0^T\widetilde{C}w_1 \geq -\left(w_0^T\widetilde{C}w_0\right)^{\frac{1}{2}}\left(w_1^T\widetilde{C}w_1\right)^{\frac{1}{2}}$ we infer

$$-\left(w_0^T\widetilde{C}w_0\right)^{\frac{1}{2}} \geq -\frac{c^*}{c^*+\theta}\left(w_1^T\widetilde{C}w_1\right)^{\frac{1}{2}}.$$

Note that this inequality holds also when $\widetilde{C}w_0 = 0$. Thus,

$$w_0^T\widetilde{C}w_1 \geq -\left(w_0^T\widetilde{C}w_0\right)^{\frac{1}{2}}\left(w_1^T\widetilde{C}w_1\right)^{\frac{1}{2}} \geq -\frac{c^*}{c^*+\theta}w_1^T\widetilde{C}w_1. \tag{4.9}$$

We then premultiply equation (4.7) by w_1^T, and we bound the leftmost term as follows:

$$\begin{aligned} w_1^T\widetilde{G}\left((1-\theta)F - \theta\widetilde{G}^T\widetilde{G}\right)^{-1}\widetilde{G}^Tw_1 &\geq w_1^T\widetilde{G}\left(\|F\|\,(1-\theta)I_{n_1} - \theta\widetilde{G}^T\widetilde{G}\right)^{-1}\widetilde{G}^Tw_1 \\ &\geq \frac{g_+}{(1-\theta)\,\|F\| - \theta g_+}\,\|w_1\|^2, \end{aligned}$$

where the last inequality is justified by the fact that w_1 is orthogonal to the null space of $\widetilde{G}\left(\|F\|\,(1-\theta)I_{n_1} - \theta\widetilde{G}^T\widetilde{G}\right)^{-1}\widetilde{G}^T$, and that the eigenvalues of such matrix are of the form $\dfrac{\sigma^2}{(1-\theta)\,\|F\| - \theta\sigma^2}$, where σ is a singular value of $\widetilde{G}^T$. Therefore, we obtain

$$\frac{g_+}{(1-\theta)\,\|F\| - \theta g_+}\,\|w_1\|^2 + w_1^T\widetilde{C}(w_0+w_1) \leq -\theta\,\|w_1\|^2. \tag{4.10}$$

According to the inequality (4.9), it holds

$$w_1^T\widetilde{C}(w_0+w_1) \geq \left(1-\frac{c^*}{c^*+\theta}\right)w_1^T\widetilde{C}w_1 \geq \frac{c_1\theta}{c^*+\theta}\,\|w_1\|^2.$$

Thus, after dividing (4.10) by $\|w_1\|^2$, we obtain

$$\frac{g_+}{(1-\theta)\,\|F\| - \theta g_+} + \frac{c_1\theta}{c^*+\theta} \leq -\theta.$$

After some algebra, $\theta^3 t_+ + \theta^2\left((c_1+c^*)t_+ - 1\right) - \theta\left(c_1+c^*+t_+-1\right) - (t_+-1)\,c^* \geq 0$, where $t_+ = 1 + \dfrac{g_+}{\|F\|}$. If we call $q(\theta)$ the above cubic polynomial and η its largest negative root, then it holds that $\theta \leq \eta$. Since $q(-c^*) \geq 0$, then $-c^* \leq \eta$. $\square$

Finally, we specialize the result to our setting, by using the first ordering in (4.1) with which $\xi_2 = 0$, so that the rightmost negative eigenvalue is bounded by η in Theorem 4.3.

Corollary 4.4. *Assume that* $\begin{bmatrix} J \\ -Z^{\frac{1}{2}} \end{bmatrix}$ *has a non-trivial null space, and let*

$$M = \begin{bmatrix} F & G^T \\ G & -C \end{bmatrix} := \left[\begin{array}{c|cc} H & J^T & -Z^{\frac{1}{2}} \\ \hline J & 0 & 0 \\ -Z^{\frac{1}{2}} & 0 & -X \end{array}\right].$$

Let g_+ be the minimum positive eigenvalue of $W^{-1}GG^T$, and c^* be as defined in Theorem 4.3. Then

$$\Lambda(P_D^{-1}M) \subseteq [\xi_1,\ \eta] \cup [\xi_3,\ 1]\,,$$

where ξ_1 and ξ_3 are given in Theorem 4.1, and $\eta \geq -c^*$ is the largest negative root of the cubic polynomial in (4.8) with $t_+ = 1 + \dfrac{g_+}{\|H\|}$.

4.3. On the choice of W. We wish to identify choices of W by which the eigenvalues of the preconditioned matrix are bounded by constants independent of the problem parameters (e.g., the spectral properties of the blocks). A thorough algebraic analysis of the choice of W can be found in [14, 15], where however the focus was not the use of W within the context in (4.3).

If C were positive definite, as in the regularized case, the choice $W = C$ would do and would lead to the standard block diagonal preconditioner; see, e.g., [3]. Then, we would have $c_0 = c_1 = c^* = 1$, and the expression for I^- and I^+ would be ideal and consistent with the known results in [18, Proposition 4.2]. In our setting, and for both orderings, C is singular or very ill-conditioned, and we show in the following theorem that the alternative choice $W = C + \frac{1}{\|F\|}GG^T$ also yields an ideal preconditioner. Clearly, this choice has computational drawbacks, as its effectiveness depends on the cost of solving with W.

Theorem 4.5. *Let $F \in \mathbb{R}^{n_1 \times n_1}$ and $C \in \mathbb{R}^{n_2 \times n_2}$ be SPSD, $G \in \mathbb{R}^{n_2 \times n_1}$, M in (4.2) nonsingular. Given $W = C + \dfrac{1}{\|F\|}GG^T$ and P_D in (4.3), it holds that $\Lambda(P_D^{-1}M) \subseteq I^- \cup I^+$ with*

$$I^- = \left[-\frac{1+\sqrt{1+4(1+f)^2}}{2(1+f)},\ -\frac{1}{2}\right] \subseteq \left[-\frac{1}{2}\left(1+\sqrt{5}\right),\ -\frac{1}{2}\right],$$

$$I^+ = \left[\frac{-1+\sqrt{1+4(1+f)^2}}{2(1+f)},\ 1\right] \subseteq \left[\frac{1}{2}\left(-1+\sqrt{5}\right),\ 1\right]$$

and $f = \dfrac{\lambda_{\min}(F)}{\|F\|} \geq 0$. Moreover, if ℓ denotes the nullity of G^T, then $P_D^{-1}M$ has the eigenvalue -1 with multiplicity ℓ.

Proof. Direct calculation shows that any vector of the form $(0, v)$, with $v \in \ker(G^T)$ is an eigenvector of $P_D^{-1}M$ with associated eigenvalue equal to -1.

Let $c_1 = \lambda_{\max}\left(W^{-1}C\right)$. Since $W \succeq C$, it follows $c_1 \leq 1$. Moreover,

$$c_1 \geq c^* = \min_{0 \neq x \in \ker(G^T)} \frac{x^T C x}{x^T\left(C + \frac{1}{\|F\|}GG^T\right)x} = 1.$$

and as a consequence $c_1 = c^* = 1$.

From $W \succeq \frac{1}{\|F\|} GG^T$, we obtain $G^T W^{-1} G \preceq \|F\| \, I_{n_1}$ and thus $s = \lambda_{\min}(F)/g_1 \geq \lambda_{\min}(F)/\|F\|$. Then all bounds except the right extreme of I^- follow from Theorem 4.1.

Regarding the upper bound for the negative eigenvalues, we start again from equations (4.4) and (4.5) and suppose $v \notin \ker(G^T)$ (if $v \in \ker(G^T)$ we immediately have that $\theta = -1$). If we eliminate u from (4.4) and substitute into (4.5), we obtain

$$G \left(\left((1-\theta)F - \theta G^T W^{-1} G\right)^{-1} + \frac{\theta}{\|F\|} I_{n_1} \right) G^T v + (1+\theta) C v = 0.$$

Let us now multiply the above equation by v^T from the left. Assuming that $\theta \geq -1$, it holds that $(1+\theta) v^T C v \geq 0$ and we get

$$v^T G \left(\left((1-\theta)F - \theta G^T W^{-1} G\right)^{-1} + \frac{\theta}{\|F\|} I_{n_1} \right) G^T v \leq 0.$$

Since $v \notin \ker(G^T)$, the minimum eigenvalue of $\left((1-\theta)F - \theta G^T W^{-1} G\right)^{-1} + \frac{\theta}{\|F\|} I_{n_1}$ must be nonpositive. Thus, again by $G^T W^{-1} G \preceq \|F\| \, I_n$ and $F \preceq \|F\| \, I_n$, it follows

$$0 \geq \lambda_{\min} \left(\left((1-\theta)F - \theta G^T W^{-1} G\right)^{-1} + \frac{\theta}{\|F\|} I_{n_1} \right) \geq \frac{1}{(1-2\theta)\,\|F\|} + \frac{\theta}{\|F\|}.$$

Rearranging the above equation we obtain $-2\theta^2 + \theta + 1 \leq 0$ from which $\theta \leq -\frac{1}{2}$. □

4.4. Numerical illustration. In this section we illustrate the quality of the spectral estimates derived for the proposed class of preconditioners. As a model problem, we used the first test problem in Section 3.1 with $q = \mathrm{fix}((n-m)/2)$, taken from the fifth iteration of the IP method.

In the first set of experiments we explore the setting of Corollary 4.2, so that

$$F = \begin{bmatrix} 0 & 0 \\ 0 & X \end{bmatrix}, \quad C = H, \quad G = \begin{bmatrix} -J^T & Z^{\frac{1}{2}} \end{bmatrix}, \tag{4.11}$$

with $n_1 > n_2$ and $\ker(G^T) = \{0\}$. Table 4.1 reports the true spectral intervals and the bounds from Corollary 4.2 and Theorem 4.5, for the following choices for W:

$$W = C = H, \qquad W = W_S = C + \frac{1}{\|F\|} GG^T = H + \frac{1}{\max\limits_{1 \leq i \leq n} |x_i|} (J^T J + Z). \tag{4.12}$$

For both selections, the bounds are rather sharp.

W	I^-	$[\lambda_1,\ \lambda_{n_2}]$	I^+	$[\lambda_{n_2+1},\ \lambda_{n_1+n_2}]$
H	$[-1.618,\ -1.615]$	$[-1.618,\ -1.618]$	$[0.618,\ 1]$	$[0.618,\ 1.00]$
W_S	$[-1.001,\ -0.500]$	$[-1.000,\ -0.527]$	$[0.998,\ 1]$	$[0.999,\ 1.00]$

TABLE 4.1

True spectral intervals and bounds from Corollary 4.2 and Theorem 4.5; W_S is as in (4.12).

In the second set we employ Corollary 4.4, which uses

$$F = H, \quad C = \begin{bmatrix} 0 & 0 \\ 0 & X \end{bmatrix}, \qquad G = \begin{bmatrix} J \\ -Z^{\frac{1}{2}} \end{bmatrix}, \tag{4.13}$$

so that $n_1 < n_2$ and $\ker(G^T) \neq \{0\}$. The following typical choices for the matrix W were considered as nonsingular completion of C:

i) Schur-complement type augmentation:

$$W = W_S = C + \frac{1}{\|F\|} GG^T = \begin{bmatrix} \frac{1}{\|H\|} JJ^T & -\frac{1}{\|H\|} JZ^{\frac{1}{2}} \\ -\frac{1}{\|H\|} Z^{\frac{1}{2}} J^T & X + \frac{1}{\|H\|} Z \end{bmatrix}, \tag{4.14}$$

ii) $W = W_\delta = C$, for the regularized problem with $C = \begin{bmatrix} \delta I_m & 0 \\ 0 & X \end{bmatrix}$, with $\delta = 10^{-6}$.

Table 4.2 shows the comparison between the true spectral intervals, and the sets I^- and I^+ obtained in Corollary 4.4 and Theorem 4.5. Rather sharp bounds are obtained for the considered Ws.

W	I^-	$[\lambda_1,\ \lambda_{n_2}]$	I^+	$[\lambda_{n_2+1},\ \lambda_{n_1+n_2}]$
W_S	$[-1.280,\ -0.500]$	$[-1.000,\ -0.500]$	$[0.780,\ 1]$	$[0.999,\ 1.000]$
W_δ	$[-1.618,\ -1.000]$	$[-1.618,\ -1.000]$	$[0.618,\ 1]$	$[0.618,\ 0.618]$

Table 4.2

True spectral intervals and bounds from Corollary 4.4 and Theorem 4.5; W_S is as defined in (4.14), W_δ is applied to the regularized problem.

As already mentioned, a major consideration in the choice of W is that both W and $F + G^T W^{-1} G$ should be cheap to invert. While the constraint on W can be easily dealt with by using, e.g., a diagonal matrix, the constraint on $F + G^T W^{-1} G$ cannot be resolved without taking into account the specific application data. Indeed, $F + G^T W^{-1} G$ could be much denser than each of its terms [28], and some approximation of its inverse should be considered. Depending on the chosen setting, matrices F and G can change with the IP iteration, so that any (incomplete) factorization needs to be recomputed. Strategies that avoid this new computation from scratch could be fruitfully employed, see, e.g., [2, 36].

All these issues are of great importance in devising practical preconditioning strategies that can be applied to a large variety of constrained optimization problems; a more detailed analysis of these issues can be found in, e.g., [26].

5. Conclusions. We have studied symmetric unreduced KKT systems, as they arise in the solution of convex quadratic programming problems solved by IP methods, and we have characterized the spectrum of the corresponding matrices.

In the unpreconditioned case, we distinguished between two stages of the IP method: generic iterations, and late or final stage. A spectral analysis should be able to reflect the peculiarities of each of these two phases, and in particular to capture the potential nonsingularity of the matrices at the limit. For the generic iteration, we were able to measure the distance from singularity for the negative eigenvalues. By properly partitioning the coefficient matrix, we were also able to characterize the spectral properties in the late stage of the IP iterations and at the solution, giving novel reliable estimates in this delicate case.

We also addressed the use of positive definite augmented preconditioners, that preserve the symmetry of the coefficient matrix, while coping with the possible high singularity of the diagonal blocks. General spectral bounds for the preconditioned matrix were derived; they significantly expand results in the literature, covering the case of all nonzero diagonal blocks.

Appendix. In this appendix we prove Lemma 3.2.

Proof. Matrix B has dimension $(m+q)\times n$ where q is the cardinality of the active set $\mathcal{A}$ at $\hat{x}$ (see the discussion on q before (3.8)). By the LICQ condition, $q \le n-m$ and $J_{\mathcal{J}}^T$ has full column rank. Consequently, B^T has full column rank.

We provide estimates for $\sigma_{\max}(B)$ and $\sigma_{\min}(B)$ by using the relations $\sigma^2_{\max}(B) = \lambda_{\max}(BB^T)$ and $\sigma^2_{\min}(B) = \lambda_{\min}(BB^T)$ and considering the eigenvalue problem for BB^T, that is

$$\begin{bmatrix} J_{\mathcal{A}}J_{\mathcal{A}}^T + J_{\mathcal{J}}J_{\mathcal{J}}^T & -J_{\mathcal{A}}Z_{\mathcal{A}}^{\frac{1}{2}} \\ -Z_{\mathcal{A}}^{\frac{1}{2}}J_{\mathcal{A}}^T & Z_{\mathcal{A}} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} u \\ v \end{bmatrix}. \tag{5.1}$$

If $v=0$, then from the second equation $J_{\mathcal{A}}^T u = 0$ and from the first equation we find $\sigma^2_{\min}(J_{\mathcal{J}}) \le \lambda \le \sigma^2_{\max}(J_{\mathcal{J}})$. Then, we first focus on $\sigma_{\min}(B)$ and consider the case where $v \neq 0$ and $\lambda < \sigma^2_{\min}(J_{\mathcal{J}})$, otherwise $\sigma^2_{\min}(J_{\mathcal{J}})$ is the requested bound. By the first block equation in (5.1), $u = \left(J_{\mathcal{A}}J_{\mathcal{A}}^T + J_{\mathcal{J}}J_{\mathcal{J}}^T - \lambda I_m\right)^{-1} J_{\mathcal{A}}Z_{\mathcal{A}}^{\frac{1}{2}}v$. Then, the second block equation of (5.1) becomes

$$-Z_{\mathcal{A}}^{\frac{1}{2}}J_{\mathcal{A}}^T\left(J_{\mathcal{A}}J_{\mathcal{A}}^T + J_{\mathcal{J}}J_{\mathcal{J}}^T - \lambda I_m\right)^{-1} J_{\mathcal{A}}Z_{\mathcal{A}}^{\frac{1}{2}}v + Z_{\mathcal{A}}v - \lambda v = 0,$$

and premultiplying it by v^T we get

$$v^T Z_{\mathcal{A}}^{\frac{1}{2}}\left[I_q - J_{\mathcal{A}}^T\left(J_{\mathcal{A}}J_{\mathcal{A}}^T + J_{\mathcal{J}}J_{\mathcal{J}}^T - \lambda I_m\right)^{-1} J_{\mathcal{A}}\right] Z_{\mathcal{A}}^{\frac{1}{2}}v - \lambda\|v\|^2 = 0, \tag{5.2}$$

Now we observe that $\left(J_{\mathcal{A}}J_{\mathcal{A}}^T + J_{\mathcal{J}}J_{\mathcal{J}}^T - \lambda I_m\right) \succeq \left(J_{\mathcal{A}}J_{\mathcal{A}}^T + (\sigma^2_{\min}(J_{\mathcal{J}}) - \lambda)I_m\right)$, and $J_{\mathcal{A}}^T\left(J_{\mathcal{A}}J_{\mathcal{A}}^T + J_{\mathcal{J}}J_{\mathcal{J}}^T - \lambda I_m\right)^{-1} J_{\mathcal{A}} \preceq J_{\mathcal{A}}^T\left(J_{\mathcal{A}}J_{\mathcal{A}}^T + (\sigma^2_{\min}(J_{\mathcal{J}}) - \lambda)I_m\right)^{-1} J_{\mathcal{A}}$. Then

$$\begin{aligned} w^T J_{\mathcal{A}}^T\left(J_{\mathcal{A}}J_{\mathcal{A}}^T + (\sigma^2_{\min}(J_{\mathcal{J}}) - \lambda)I_m\right)^{-1} J_{\mathcal{A}}w &\le \max_i \frac{\sigma_i^2(J_{\mathcal{A}})}{\sigma_i^2(J_{\mathcal{A}}) + \sigma^2_{\min}(J_{\mathcal{J}}) - \lambda}\|w\|^2 \\ &= \frac{\sigma^2_{\max}(J_{\mathcal{A}})}{\sigma^2_{\max}(J_{\mathcal{A}}) + \sigma^2_{\min}(J_{\mathcal{J}}) - \lambda}\|w\|^2 \end{aligned}$$

where $w \in \mathbb{R}^q$ and (5.2) gives

$$(z_{\mathcal{A}})_{\min} - \frac{(z_{\mathcal{A}})_{\min}\sigma^2_{\max}(J_{\mathcal{A}})}{\sigma^2_{\max}(J_{\mathcal{A}}) + \sigma^2_{\min}(J_{\mathcal{J}}) - \lambda} - \lambda \le 0,$$

which is equivalent to $r(\lambda) = \lambda^2 - (\sigma^2_{\max}(J_{\mathcal{A}}) + \sigma^2_{\min}(J_{\mathcal{J}}) + (z_{\mathcal{A}})_{\min})\lambda + (z_{\mathcal{A}})_{\min}\sigma^2_{\min}(J_{\mathcal{J}}) \le 0$. Since $r(0) > 0$ and $r\left(\sigma^2_{\min}(J_{\mathcal{J}})\right) < 0$, then $\sigma^2_{\min}(J_{\mathcal{J}})$ is greater than the smallest root of $r(\lambda)$ and the stated bound on $\sigma_{\min}(B)$ follows.

Finally, if $w \in \mathbb{R}^q$, $\hat{w} \in \mathbb{R}^{n-q}$ we have

$$\left\|B\begin{bmatrix} w \\ \hat{w}\end{bmatrix}\right\|^2 = \left\|J\begin{bmatrix} w \\ \hat{w}\end{bmatrix}\right\|^2 + \left\|Z_{\mathcal{A}}^{\frac{1}{2}}w\right\|^2 \le \left(\|J\|^2 + \|Z_{\mathcal{A}}\|\right)\left\|\begin{bmatrix} w \\ \hat{w}\end{bmatrix}\right\|^2,$$

from which the looser bound for $\sigma_{\max}(B)$ follows.

The sharper bound for $\sigma_{\max}(B)$, although more complicated, can be derived as follows. We start again from (5.1) and suppose $\lambda > \sigma^2_{\max}$, which in particular implies

$v \neq 0$. As before, we find u from the first equation (note that $J_{\mathcal{A}} J_{\mathcal{A}}^T + J_{\mathcal{I}} J_{\mathcal{I}}^T = JJ^T$) and substitute into the second one. Premultiplying for v^T we obtain again equation (5.2). This time we use

$$J_{\mathcal{A}}^T \left(\lambda I_m - JJ^T\right)^{-1} J_{\mathcal{A}} \preceq \frac{\sigma_{\max}^2(J_{\mathcal{A}})}{\lambda - \sigma_{\max}^2} I_q.$$

Proceeding as above, we derive the inequality

$$(z_{\mathcal{A}})_{\max} + \frac{(z_{\mathcal{A}})_{\max}\sigma_{\max}^2(J_{\mathcal{A}})}{\lambda - \sigma_{\max}^2} - \lambda \geq 0,$$

which is equivalent to $q(\lambda) := \lambda^2 - \left((z_{\mathcal{A}})_{\max} + \sigma_{\max}^2\right)\lambda + (z_{\mathcal{A}})_{\max}\left(\sigma_{\max}^2 - \sigma_{\max}^2(J_{\mathcal{A}})\right) \leq 0$. Since $q\left(\sigma_{\max}^2\right) < 0$, then $\sigma_{\max}^2$ is smaller than the largest root of $q(\lambda)$. We thus obtain

$$\sigma_{\max}^2(B) \leq \frac{1}{2}\left((z_{\mathcal{A}})_{\max} + \sigma_{\max}^2 + \sqrt{\left((z_{\mathcal{A}})_{\max} - \sigma_{\max}^2\right)^2 + 4(z_{\mathcal{A}})_{\max}\sigma_{\max}^2(J_{\mathcal{A}})}\right).$$

Since $\sigma_{\max}^2 - \sigma_{\max}^2(J_{\mathcal{A}}) \geq 0$, this bound is sharper than the simpler one. □

REFERENCES

[1] S. Bellavia. An inexact interior point method. *Journal of Optimization Theory and Applications*, 96:109–121, 1998.

[2] S. Bellavia, V. De Simone, D. di Serafino, and B. Morini. A preconditioning framework for sequences of diagonally modified linear systems arising in optimization. *SIAM Journal on Numerical Analysis*, 50:3280–3302, 2012.

[3] M. Benzi, G.H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numerica*, 14:1–137, 2005.

[4] M. Benzi, E. Haber, and L. Taralli. Multilevel algorithms for large-scale interior point methods. *SIAM Journal on Scientific Computing*, 31:4152–4175, 2009.

[5] Zhi-Hao Cao. Augmentation block preconditioners for saddle point-type matrices with singular (1, 1) blocks. *Numerical Linear Algebra with Applications*, 15:515–533, 2008.

[6] M. D'Apuzzo, V. De Simone, and D. di Serafino. On mutual impact of linear algebra and large-scale optimization with focus on interior point methods. *Computational Optimization and Applications*, 45:283–310, 2010.

[7] A. Forsgren. Inertia-controlling factorizations for optimization algorithms. *Applied Numerical Mathematics*, 43:91–107, 2002.

[8] A. Forsgren, P.E. Gill, and J.D. Griffin. Iterative solution of augmented systems arising in interior methods. *SIAM Journal on Optimization*, 18:666–690, 2007.

[9] 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 Optimization*, 17:187–211, 1996.

[10] A. Forsgren, P.E. Gill, and M.H. Wright. Interior methods for nonlinear optimization. *SIAM Review*, 443:525–597, 2002.

[11] R. Fourer and S. Mehrotra. Solving symmetric indefinite systems in an interior-point method for convex quadratic programs. *Mathematical Programming*, 62:15–39, 1993.

[12] M.P. Friedlander and D. Orban. A primal-dual regularization interior-point method for convex quadratic programs. *Mathematichal Programming Computation*, 4:71–107, 2012.

[13] G. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, 3rd edition, 1996.

[14] G.H. Golub and C. Greif. On solving block-structured indefinite linear systems. *SIAM Journal on Scientific Computing*, 24(6):2076–2092, 2003.

[15] G.H. Golub, C. Greif, and J.M. Varah. An algebraic analysis of a block diagonal preconditioner for saddle point systems. *SIAM Journal on Matrix Analysis and Applications*, 27(3):779–792, 2005.

[16] J. Gondzio. Interior point methods 25 years later. *European Journal on Operational Research*, 218:587–601, 2012.

[17] J. Gondzio. Matrix-free interior point method. *Computational Optimization and Applications*, 51:457–480, 2012.

[18] N.I.M. Gould and V. Simoncini. Spectral analysis of saddle point matrices with indefinite leading blocks. *SIAM Journal on Matrix Analysis and Applications*, 31:1152–1171, 2009.

[19] C. Greif, E. Moulding, and D. Orban. Bounds on eigenvalues of matrices arising from interior-point methods. *SIAM Journal on Optimization*, 24:49–83, 2014.

[20] C. Greif and D. Schötzau. Preconditioners for saddle point linear systems with highly singular (1,1) blocks. *Electronic Transactions on Numerical Analysis*, 22:114–121, 2006.

[21] F. Jarre and C. Hergenroeder. A symmetric reduction of the NT direction. *SIAM Journal on Optimization*, 24:714–732, 2014.

[22] W. Krendl, V. Simoncini, and W. Zulehner. Stability estimates and structural spectral properties of saddle point problems. *Numerische Mathematik*, 124:183–213, 2013.

[23] L. Lukšan and J. Vlček. Indefinitely preconditioned inexact Newton method for large sparse equality constrained nonlinear programming problems. *Numerical Linear Algebra with Applications*, 5:219–247, 1998.

[24] I. Maros and C. Meszaros. A repository of convex quadratic programming problems. *Optimization Methods and Software*, 11&12:671–681, 1999.

[25] B. Morini, V. Simoncini, and M. Tani. Unreduced symmetric KKT systems arising from interior point methods. Part I: spectral estimates. *Optimization Online, http://www.optimization-online.org/DB_HTML/2014/05/4356.html*, 2014.

[26] B. Morini, V. Simoncini, and M. Tani. Unreduced symmetric KKT systems arising from interior point methods. part II: preconditioning. *Optimization Online, http://www.optimization-online.org/DB_HTML/2014/07/4418.html*, 2014.

[27] PDCO: Primal-Dual interior method for Convex Objectives. http://convexoptimization.com/sol-software/pdco.html.

[28] T. Rees and C. Greif. A preconditioner for linear systems arising from interior point optimization methods. *SIAM Journal on Scientific Computing*, 29:1992–2007, 2007.

[29] T. Rusten and R. Winther. A preconditioned iterative method for saddlepoint problems. *SIAM Journal on Matrix Analysis and Applications*, 13:887–904, 1992.

[30] M.A. Saunders. Cholesky-based methods for sparse least-squares: the benefit of regularization. In L.Adams and J.L. Nazareth, editors, *Linear and nonlinear conjugate gradient-related methods*, pages 92–100, Philadelphia, 1994. SIAM.

[31] D. Silvester and A. Wathen. Fast iterative solution of stabilised Stokes systems part II. Using general block preconditoners. *SIAM Journal on Numerical Analysis*, 31:1352–1367, 1994.

[32] The University of Florida Sparse Matrix Collection. http://www.cise.ufl.edu/research/sparse/matrices/.

[33] M.H. Wright. Ill-conditioning and computational error in interior methods for nonlinear programming. *SIAM Journal on Optimization*, 9:84–111, 1998.

[34] S.J. Wright. Stability of linear equations solvers in interior-point methods. *SIAM Journal on Matrix Analysis and Applications*, 16:1287 – 1307, 1995.

[35] S.J. Wright. *Primal-dual interior-point methods*. SIAM, Philadelphia, USA, 1997.

[36] W.Wang and D.P. O'Leary. Adaptive use of iterative methods in predictor-corrector interior point methods for linear programming. *Numerical Algorithms*, 25:387–406, 2000.