A COMPARISON OF REDUCED AND UNREDUCED KKT SYSTEMS ARISING FROM INTERIOR POINT METHODS*

BENEDETTA MORINI§, VALERIA SIMONCINI†, AND MATTIA TANI¶

Abstract. We address the iterative solution of KKT systems arising in the solution of convex quadratic programming problems. Two strictly related and well established formulations for such systems are studied with particular emphasis on the effect of preconditioning strategies on their relation. Constraint and augmented preconditioners are considered, and the choice of the augmentation matrix is discussed. A theoretical and experimental analysis is conducted to assess which of the two formulations should be preferred for solving large-scale problems.

Key words. convex quadratic programming, interior point methods, KKT systems, preconditioners.

1. Introduction.

Interior Point (IP) methods for solving large-scale Quadratic Programming (QP) problems have been an intensive area of research in the past decades. Their efficiency depends heavily on the per-iteration cost and this is mainly due to the solution of a structured algebraic linear system. Therefore, much effort has been devoted to developing properly tailored preconditioned iterative solvers, whose computational cost and memory requirements may be lower than those of direct solvers. The analysis and development of these resulting *inexact* IP methods have covered several different aspects such as the level of accuracy in the solution of linear systems, the design of suitable iterative methods and preconditioners, and the convergence analysis of the inexact IP solver, including worst-case iteration complexity, see, e.g., [4, 7, 8, 11, 14, 16, 19, 22, 34].

We consider the sequence of symmetric KKT systems arising in the solution of convex QP problems in standard form and analyze two well established formulations for such systems. The former is the system originally formed during the application of an IP method and its symmetrized version; the latter is a widely used formulation of smaller dimension obtained by an inexpensive variable reduction. In the following, we refer to these systems as *unreduced* and *reduced* respectively. Relevant alternative formulations, possibly definite, such as condensed systems and doubly augmented systems (see, e.g., [19]) are not considered in this work. It is well-known that the reduced systems become increasingly ill-conditioned in the progress of the IP iterations while the unreduced systems may be well-conditioned throughout the IP iteration and nonsingular even in the limit. This distinguishing feature, observed in [18] and recently supported by spectral analysis in [25, 31], motivates our interest in this possibly less exercised formulation.

On the one hand, an in-depth study of the ill-conditioning and computational error of direct system solvers in the reduced formulation has shown that, in spite of possible increasing ill-conditioning at the late stage of the IP method, the steps remain sufficiently accurate to be useful [20, 39]. This analysis has supported the

†Dipartimento di Matematica, Piazza di Porta S. Donato 5, 40127 Bologna, Italia. valeria.simoncini@unibo.it

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

¶Dipartimento di Matematica, Università di Pavia, via A. Ferrata 1, 27100 Pavia, Italia. mattia.tani@unipv.it

*Version of June 29, 2016. Work partially supported by INdAM-GNCS under the 2015 Project *Metodi di regolarizzazione per problemi di ottimizzazione e applicazioni*.

practical use of the reduced systems, due to their lower dimension. On the other hand, a theoretical and experimental comparison of the iterative solution of reduced and unreduced formulations in the preconditioned regime has not been performed and it is the goal of this work. First we aim to assess whether the use of unreduced systems may still offer some advantage over the reduced ones in terms of eigenvalues and conditioning as occurs in the unpreconditioned case [25, 31]; successively we conduct a computational comparison between the unreduced and reduced formulations.

For the unreduced systems we use both the unsymmetric and symmetric formulations, while the reduced systems are genuinely symmetric. Regularization of the QP problem is also considered. We relate the systems by either matrix transformations or congruence transformations. Then, we consider some frequently employed constraint preconditioners and augmented preconditioners and show that in the unreduced case, their condition number may vary slowly with the IP iterations and remain considerably smaller than in the reduced formulation. We also show that the two formulations remain strictly related in terms of spectra of the preconditioned matrices, block structure of the linear equations and, possibly, convergence behaviour of the iterative solver's residuals.

Numerical experiments with standard QP problems and the IP solver PDCO [33] confirm that the conditioning of the unreduced systems, as well as of some corresponding preconditioners, may be considerably smaller than in the reduced formulation. However, these better spectral properties may not play a role as long as the reduced systems are numerically nonsingular, which is a realistic situation if the QP problem is well-scaled and the solution accuracy is not very high. Specifically, in such a case our experience shows that the IP implementations with the unreduced and reduced formulations are both successful and the preconditioned linear solvers and the IP solvers behave similarly. Thus the smaller dimension of the reduced systems makes such formulation preferable in terms of computation time.

In principle the unreduced formulation may maintain better spectral properties than the reduced one and its use may enhance the robustness of the interior point solver when the reduced systems become severely ill-conditioned, or even numerically singular, although we were not able to see this in practice.

The paper is organized as follows. In §2 we review the linear system formulations under study and well-known classes of structured preconditioners suitable for such matrices, and establish connections between the systems. In §3 we show that the unreduced and reduced formulations are closely related when some types of preconditioners are used, and characterize the spectra of the preconditioned matrices. In §4 we experimentally compare the two formulations and their iterative solution in an IP solver, and in §5 we draw our conclusions.

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 $(x)_i$. Given $x \in \mathbb{R}^n$, $X = \text{diag}(x)$ is the diagonal matrix with diagonal entries $x_1, \dots, x_n$. Given column vectors x and y, we write (x, y) for the column vector given by their concatenation instead of using $[x^T, y^T]^T$. Given a matrix A, the set of its eigenvalues is indicated by $\Lambda(A)$. The notation $\text{diag}(A)$ denotes the diagonal matrix derived from the diagonal of A. Finally, for any positive integer p, the identity matrix of dimension p is indicated by I_p.

2. Preliminaries. We consider the primal-dual pair of QP problems

$$\min_{x} c^T x + \frac{1}{2} x^T H x \quad \text{subject to} \quad Jx = b,\ x \geq 0, \tag{2.1}$$

$$\max_{x,y,z} b^T y - \frac{1}{2} x^T H x \quad \text{subject to} \quad J^T y + z - Hx = c,\ z \geq 0, \tag{2.2}$$

where $J \in \mathbb{R}^{m\times n}$ has dimensions $m \leq n$, $H \in \mathbb{R}^{n\times n}$ is symmetric and positive semidefinite, $x,\ z,\ c \in \mathbb{R}^n$, $y,\ b \in \mathbb{R}^m$, and the inequalities are meant componentwise.

A primal-dual IP method [40] is defined by replacing the inequality constraint $x \geq 0$ in (2.1) by a logarithmic barrier function and then using the Lagrangian duality theory to derive first-order optimality conditions. These conditions take the form of a system of nonlinear equations and are solved by Newton's method. In this section we review different formulations of the linear systems arising in Newton's method, and the features of the associated coefficient matrices taking the KKT form

$$M = \begin{bmatrix} E & F^T \\ G & -C \end{bmatrix}, \tag{2.3}$$

with $E \in \mathbb{R}^{n_1\times n_1}$, $C \in \mathbb{R}^{n_2\times n_2}$ both positive semidefinite, $F \in \mathbb{R}^{n_2\times n_1}$, $G \in \mathbb{R}^{n_2\times n_1}$, and any relation allowed between n_1 and n_2.

Omitting for simplicity the iteration index k in the IP method and letting $(x,\ y,\ z)$ be the current iterate with $x,\ z > 0$, we may write the Newton direction $(\Delta x,\ \Delta y,\ \Delta z)$ as the solution to the linear system

$$\begin{bmatrix} H & J^T & -I_n \\ J & 0 & 0 \\ -Z & 0 & -X \end{bmatrix} \begin{bmatrix} \Delta x \\ -\Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} -c - Hx + J^T y + z, \\ b - Jx, \\ XZ\mathbf{1}_n - \mu \mathbf{1}_n \end{bmatrix}, \tag{2.4}$$

where $X = \mathrm{diag}(x)$, $Z = \mathrm{diag}(z)$ are diagonal positive definite matrices, $\mathbf{1}_n \in \mathbb{R}^n$ is the vector of ones, and the positive scalar μ is the barrier parameter, which controls the distance to optimality and is gradually reduced during the IP iterations. If correctors are used, systems of the form (2.4) need to be solved with different right-hand sides [40].

The above system is unsymmetric, 3×3 block structured, and of dimension $2n+m$. Letting[1] $\widehat{K}_3$ be the matrix in (2.4) and R be the diagonal positive definite matrix

$$R = \begin{bmatrix} I_n & 0 & 0 \\ 0 & I_m & 0 \\ 0 & 0 & Z^{\frac{1}{2}} \end{bmatrix}, \tag{2.5}$$

we can transform the system (2.4) into a symmetric one with matrix

$$K_3 = R^{-1} \widehat{K}_3 R = \begin{bmatrix} H & J^T & -Z^{\frac{1}{2}} \\ J & 0 & 0 \\ -Z^{\frac{1}{2}} & 0 & -X \end{bmatrix}, \tag{2.6}$$

and proper right-hand side [18]. We refer to [18, 25, 31] for a detailed analysis of $\widehat{K}_3$ and K_3 and their relations.

[1] Throughout we use the hat symbol, $\widehat{\ }$, for matrices corresponding to the nonsymmetric 3×3 formulation.

Alternatively to the above 3×3 formulations, it is common to eliminate Δz in (2.4) and use a system of smaller dimension $n+m$. The resulting system has matrix

$$K_2 = \begin{bmatrix} H + X^{-1}Z & J^T \\ J & 0 \end{bmatrix}, \tag{2.7}$$

and right-hand side formed accordingly.

In order to provide a comprehensive analysis of the 3×3 systems, it is useful to consider the case where regularizations are applied to the optimization problems (2.1), (2.2). Several regularization techniques have been proposed in order to improve the numerical properties of the KKT systems, and for details we refer to [1, 10, 21, 33, 37]. Here we focus on primal-dual regularizations such that system (2.4) becomes

$$\widehat{K}_{3,\text{reg}}\, \widehat{\Delta}_3 = \widehat{f}_3 \tag{2.8}$$

with

$$\widehat{K}_{3,\text{reg}} = \begin{bmatrix} H + \rho I_n & J^T & -I_n \\ J & -\delta I_m & 0 \\ -Z & 0 & -X \end{bmatrix}, \quad \widehat{\Delta}_3 = \begin{bmatrix} \Delta x \\ -\Delta y \\ \Delta z \end{bmatrix}, \quad \widehat{f}_3 = \begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix}, \tag{2.9}$$

where $\delta, \rho \geq 0$ and the right-hand side vectors f_x, $f_z \in \mathbb{R}^n$ and $f_y \in \mathbb{R}^m$ are appropriately computed, see, e.g., [21, 25, 33]. Using again the block diagonal matrix R in (2.5) gives the corresponding symmetric formulation

$$K_{3,\text{reg}}\, \Delta_3 = f_3, \tag{2.10}$$

where

$$K_{3,\text{reg}} = \begin{bmatrix} H + \rho I_n & J^T & -Z^{\frac{1}{2}} \\ J & -\delta I_m & 0 \\ -Z^{\frac{1}{2}} & 0 & -X \end{bmatrix}, \; \Delta_3 = \begin{bmatrix} \Delta x \\ -\Delta y \\ Z^{-\frac{1}{2}} \Delta z \end{bmatrix}, \; f_3 = \begin{bmatrix} f_x \\ f_y \\ Z^{-\frac{1}{2}} f_z \end{bmatrix}. \tag{2.11}$$

Upon reduction of Δz, system (2.8) becomes

$$K_{2,\text{reg}}\, \Delta_2 = f_2, \tag{2.12}$$

with

$$K_{2,\text{reg}} = \begin{bmatrix} H + \rho I_n + X^{-1}Z & J^T \\ J & -\delta I_m \end{bmatrix}, \; \Delta_2 = \begin{bmatrix} \Delta x \\ -\Delta y \end{bmatrix}, \; f_2 = \begin{bmatrix} f_x - X^{-1} f_z \\ f_y \end{bmatrix}. \tag{2.13}$$

For $\delta, \rho \geq 0$, throughout the paper we refer to $K_{2,\text{reg}}$ as the *reduced* matrix and to $\widehat{K}_{3,\text{reg}}$ and $K_{3,\text{reg}}$ as the unsymmetric and symmetric *unreduced* matrices and recover K_2, $\widehat{K}_3$ and K_3 by setting $\delta = \rho = 0$. A similar terminology is used for the corresponding systems. For later convenience, we observe that as long as X is nonsingular, the reduced and unreduced matrices are mathematically related by means of suitable transformations. Indeed, setting

$$\widehat{L}_1 = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ X^{-1} & 0 & I \end{bmatrix}, \quad \widehat{L}_2 = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ X^{-1}Z & 0 & I \end{bmatrix}, \tag{2.14}$$

gives

$$\widehat{K}_{3,\text{reg}} = \widehat{L}_1^T \begin{bmatrix} \multicolumn{2}{c}{\multirow{2}{*}{$K_{2,\text{reg}}$}} & 0 \\ & & 0 \\ 0 & 0 & -X \end{bmatrix} \widehat{L}_2, \tag{2.15}$$

while setting

$$L = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ X^{-1}Z^{\frac{1}{2}} & 0 & I \end{bmatrix}, \tag{2.16}$$

gives the congruence transformation (see, e.g., [25])

$$K_{3,reg} = L^T \begin{bmatrix} \multicolumn{2}{c}{\multirow{2}{*}{$K_{2,\text{reg}}$}} & 0 \\ & & 0 \\ 0 & 0 & -X \end{bmatrix} L. \tag{2.17}$$

We use (2.15) and (2.17) only for theoretical purposes, but it is interesting to observe that the possibly ill-conditioned matrices $\widehat{L}_1$, $\widehat{L}_2$ and L do not in fact necessarily transfer their conditioning to $K_{3,\text{reg}}$.

The following theorem from [25, Corollary 3.4, Corollary 3.6, Theorem 3.9] characterizes the nonsingularity of the considered matrices.

Theorem 2.1. *Suppose that H is symmetric and positive semidefinite, $\rho \geq 0$, and X, Z are diagonal with positive entries. Then $K_{2,\text{reg}}$ in (2.13), $\widehat{K}_{3,\text{reg}}$ in (2.9), and $K_{3,\text{reg}}$ in (2.11) are nonsingular if and only if either $\delta > 0$, or $\delta = 0$ and J has full row rank.*

The use of the 2×2 formulation is supported by its reduced dimension and by the variety of direct solvers, iterative solvers and preconditioners available for its numerical solution [5]. However, the presence of $X^{-1}Z$ may cause ill-conditioning of $K_{2,\text{reg}}$ as a solution is approached, and it represents the key difference between $K_{2,\text{reg}}$ and the matrices $\widehat{K}_{3,\text{reg}}$, $K_{3,\text{reg}}$.

The unreduced matrices $\widehat{K}_{3,\text{reg}}$, $K_{3,\text{reg}}$ and $\widehat{K}_3$ are nonsingular throughout the IP iterations, and remarkably remain nonsingular at a solution of (2.1)-(2.2) if proper conditions, including strict complementarity and Linear Independence Constraint Qualification are satisfied. The conditions for nonsingularity stated in [25, Theorem 3.11] are reported below for completeness.

Theorem 2.2. *Suppose that H is symmetric and positive semidefinite, X, Z are diagonal with nonnegative entries, x and z satisfy the complementarity condition $x_iz_i = 0$, for all i. Necessary and sufficient conditions for $\widehat{K}_{3,\text{reg}}$ to be nonsingular are that: $x_i > 0$ if $z_i = 0$, and $z_i > 0$ if $x_i = 0$ (strict complementarity); $Ker(H) \cap Ker(J) \cap Ker(Z) = \{0\}$ if $\rho = 0$; $[J^T \ -(I_n)_{\mathcal{A}}]$ is full rank if $\delta = 0$, with $(I_n)_{\mathcal{A}}$ being the submatrix of I_n of the column corresponding to the active set $\mathcal{A}$ for x (Linear Independence Constraint Qualification).*

These properties may favor the use of unreduced systems and have indeed motivated the study of their spectral estimates [25, 31]. Regarding the use of $K_{3,\text{reg}}$, it is important to note that ill-conditioning occurs in the matrix R defined in (2.5) but the square root in the $(3, 3)$ block has a damping effect on ill-conditioning at the final stage of the IP process. On the other hand, when either of the systems in (2.8) and

(2.10) is solved iteratively and preconditioning is required, assessing the advantages of the unreduced formulation over the reduced one is still an open issue. We address this problem by studying, both theoretically and computationally, the systems (2.8), (2.10) and (2.12) when corresponding preconditioners are applied, and we establish their distinguishing features.

We conclude this section by recalling a few suitable preconditioners for our systems that will be analyzed in the following: constraint preconditioners and preconditioners based on augmentation of the $(1,1)$ block. In order to handle the above systems simultaneously, we refer to the matrix in (2.3).

A block diagonal preconditioner for M based on augmentation is

$$P_{\mathcal{D}} = \begin{bmatrix} E + F^T W^{-1} G & 0 \\ 0 & W \end{bmatrix}, \tag{2.18}$$

where $W \in \mathbb{R}^{n_2 \times n_2}$ is a symmetric positive definite weight matrix [9, 24, 26, 31]. The advantages of using $P_{\mathcal{D}}$ over preconditioners based on the Schur complement of E are visible for E singular or ill-conditioned. Interestingly, this may be the case for both the reduced and unreduced matrices, as E may be ill-conditioned in $K_{2,\text{reg}}$ at the late stage of the IP method, and singular in K_3 at any IP iteration both if E is the leading 1×1 or the leading 2×2 block. Clearly, $P_{\mathcal{D}}$ is positive definite if $F = G$ and $\ker(E) \cap \ker(F) = \{0\}$.

Alternatively, block triangular preconditioners based on augmentation of E can be defined; see, e.g., [2, 6, 9, 34, 38], where, however, most results are for a zero (2,2) block C. In this nonsymmetric framework, one option is to set

$$P_{\mathcal{T}} = \begin{bmatrix} E + F^T W^{-1} G & \kappa F^T \\ 0 & -W \end{bmatrix}, \tag{2.19}$$

where κ is a scalar and $W \in \mathbb{R}^{n_2 \times n_2}$ is a symmetric positive definite weight matrix. If $\kappa = 0$, $P_{\mathcal{T}}$ is block diagonal and indefinite. Another possibility is to consider

$$T_{\mathcal{W}} = \begin{bmatrix} E + F^T W^{-1} G & \kappa F^T \\ 0 & -(W + C) \end{bmatrix}, \tag{2.20}$$

where κ is a scalar and $W \in \mathbb{R}^{n_2 \times n_2}$ is symmetric positive definite. This preconditioner was introduced for symmetric matrices in [38] with $\kappa = 2$.

Finally, indefinite constraint preconditioners for M are commonly used in optimization, see, e.g., [7, 8, 13–16, 19, 28, 29], and take the form

$$P_{\mathcal{C}} = \begin{bmatrix} \tilde{E} & F^T \\ G & -C \end{bmatrix}, \tag{2.21}$$

where $\tilde{E}$ is a symmetric approximation to E.

3. Standard preconditioners for the unreduced systems. In this section we consider some of the preconditioners of type (2.18)–(2.21). We establish two different results on the relationship between preconditioned 2×2 and 3×3 formulations. The first result is given in §3.1 and holds for specific constraint preconditioners and augmented triangular preconditioners; invariance of the spectra of the preconditioned matrices and correspondence of block equations in the preconditioned systems are shown. The second result, given in §3.2, is valid for specific augmented diagonal and triangular preconditioners and indicates that, with the same elimination of Δz as in (2.10), the preconditioned 3×3 systems reduce to the preconditioned 2×2 systems.

Despite the spectral invariance of the preconditioned matrices $K_{2,\text{reg}}$, $\widehat{K}_{3,\text{reg}}$, $K_{3,\text{reg}}$ via specific constraint and triangular preconditioners, the fact that a preconditioner is invertible in the limit and possibly well-conditioned may be preferable in finite precision arithmetic, making the unreduced system more appealing. This topic is further explored in §4.

3.1. Equivalence properties of 2×2 and 3×3 formulations under certain preconditioners. The considered constraint preconditioners have the form

$$P_{2,\mathcal{C}} = \begin{bmatrix} \text{diag}(H + \rho I_n + X^{-1}Z) & J^T \\ J & -\delta I_m \end{bmatrix}, \tag{3.1}$$

$$\widehat{P}_{3,\mathcal{C}} = \begin{bmatrix} \text{diag}(H + \rho I_n) & J^T & -I_n \\ J & -\delta I_m & 0 \\ -Z & 0 & -X \end{bmatrix} = \widehat{L}_1^T \begin{bmatrix} P_{2,\mathcal{C}} & & 0 \\ & & 0 \\ 0 & 0 & -X \end{bmatrix} \widehat{L}_2, \tag{3.2}$$

$$P_{3,\mathcal{C}} = \begin{bmatrix} \text{diag}(H + \rho I_n) & J^T & -Z^{\frac{1}{2}} \\ J & -\delta I_m & 0 \\ -Z^{\frac{1}{2}} & 0 & -X \end{bmatrix} = L^T \begin{bmatrix} P_{2,\mathcal{C}} & & 0 \\ & & 0 \\ 0 & 0 & -X \end{bmatrix} L, \tag{3.3}$$

where the factorizations in (3.2) and (3.3) are analogous to those for $\widehat{K}_{3,\text{reg}}$ and $K_{3,reg}$ in (2.15), (2.17) and follow from the equality $\text{diag}(H+\rho I_n+X^{-1}Z) = \text{diag}(H+\rho I_n)+X^{-1}Z$. Constraint preconditioners where the $(1,1)$ block is approximated by its main diagonal are frequently used in optimization, see, e.g., [7, 8, 16, 28, 29].

Augmentation in $P_{\mathcal{T}}$ (see (2.19)) is performed on the $(1,1)$ block of $K_{2,\text{reg}}$ and $K_{3,\text{reg}}$ as follows. Let $R_d \in \mathbb{R}^{m\times m}$ be positive definite and such that

$$R_d = \delta I_m, \quad \text{with} \quad \delta > 0,$$

and let $W = \begin{bmatrix} R_d & 0 \\ 0 & X \end{bmatrix}$ for system (2.10), and $W = R_d$ for system (2.12). The augmented block triangular preconditioners $P_{2,\mathcal{T}}$, $\widehat{P}_{3,\mathcal{T}}$, $P_{3,\mathcal{T}}$ for $K_{2,\text{reg}}$, $\widehat{K}_{3,\text{reg}}$, $K_{3,\text{reg}}$ are of the form

$$P_{2,\mathcal{T}} = \begin{bmatrix} H + \rho I_n + X^{-1}Z + J^T R_d^{-1} J & \kappa J^T \\ 0 & -R_d \end{bmatrix}, \tag{3.4}$$

$$\widehat{P}_{3,\mathcal{T}} = \begin{bmatrix} H + \rho I_n + X^{-1}Z + J^T R_d^{-1} J & \kappa J^T & -\kappa I_n \\ 0 & -R_d & 0 \\ 0 & 0 & -X \end{bmatrix} = \begin{bmatrix} P_{2,\mathcal{T}} & & -\kappa I_n \\ & & 0 \\ 0 & 0 & -X \end{bmatrix}, \tag{3.5}$$

$$P_{3,\mathcal{T}} = \begin{bmatrix} H + \rho I_n + X^{-1}Z + J^T R_d^{-1} J & \kappa J^T & -\kappa Z^{\frac{1}{2}} \\ 0 & -R_d & 0 \\ 0 & 0 & -X \end{bmatrix} = \begin{bmatrix} P_{2,\mathcal{T}} & & -\kappa Z^{\frac{1}{2}} \\ & & 0 \\ 0 & 0 & -X \end{bmatrix}. \tag{3.6}$$

If regularization is applied, the above choice of W amounts to $W = C$, and GMRES [36] converges in at most two iterations if $P_{2,\mathcal{T}}$ (or $P_{3,\mathcal{T}}$) is applied exactly and round-off errors are ignored, see [38, Remark 2.1].

Now we establish connections between the spectra of the systems (2.12), (2.8) and (2.10) preconditioned by the constraint preconditioners in (3.1)–(3.3) and the triangular preconditioners (3.4)–(3.6). Such results follow from the transformations (2.15), (2.17). We prove that, apart from the multiplicity of the unit eigenvalue, the eigenvalues of preconditioned $K_{3,\mathrm{reg}}$ and $K_{2,\mathrm{reg}}$ coincide with either the constraint or the triangular (with $\kappa = 1$) preconditioners. Second we show that the first two block equations of $P_{3,\mathcal{C}}^{-1}K_{3,\mathrm{reg}}\Delta_3 = P_{3,\mathcal{C}}^{-1}f_3$ ($P_{3,\mathcal{T}}^{-1}K_{3,\mathrm{reg}}\Delta_3 = P_{3,\mathcal{T}}^{-1}f_3$) are equivalent to $P_{2,\mathcal{C}}^{-1}K_{2,\mathrm{reg}}\Delta_2 = P_{2,\mathcal{C}}^{-1}f_2$ ($P_{2,\mathcal{T}}^{-1}K_{2,\mathrm{reg}}\Delta_2 = P_{2,\mathcal{T}}^{-1}f_2$) and the third equation is equivalent to the third equation in (2.10). Similar results hold for the unsymmetric 3×3 system.

THEOREM 3.1. *Suppose that H is symmetric positive semidefinite, X and Z are diagonal with positive entries, $K_{2,\mathrm{reg}}$, $\widehat{K}_{3,\mathrm{reg}}$ and $K_{3,\mathrm{reg}}$ in (2.13), (2.9) and (2.11) are nonsingular. Let $P_{2,\mathcal{C}}$, $\widehat{P}_{3,\mathcal{C}}$ and $P_{3,\mathcal{C}}$ be the preconditioners in (3.1), (3.2) and (3.3) respectively, and consider systems (2.12), (2.8) and (2.10). Then*

i) $\theta \in \Lambda(P_{3,\mathcal{C}}^{-1}K_{3,\mathrm{reg}})$ if and only if either $\theta = 1$ or $\theta \in \Lambda(P_{2,\mathcal{C}}^{-1}K_{2,\mathrm{reg}})$.

ii) Solving $P_{3,\mathcal{C}}^{-1}K_{3,\mathrm{reg}}\Delta_3 = P_{3,\mathcal{C}}^{-1}f_3$ is equivalent to solving $P_{2,\mathcal{C}}^{-1}K_{2,\mathrm{reg}}\Delta_2 = P_{2,\mathcal{C}}^{-1}f_2$ and recovering Δz from the third equation in (2.10).

iii) $\theta \in \Lambda(\widehat{P}_{3,\mathcal{C}}^{-1}\widehat{K}_{3,\mathrm{reg}})$ if and only if either $\theta = 1$ or $\theta \in \Lambda(P_{2,\mathcal{C}}^{-1}K_{2,\mathrm{reg}})$.

iv) The same feature in (ii) holds with $\widehat{K}_{3,\mathrm{reg}}$ and $\widehat{P}_{3,\mathcal{C}}$.

Proof. To show *i)*, consider the generalized eigenvalue problem $K_{3,reg}u = \theta P_{3,\mathcal{C}}u$ with $u \in \mathbb{R}^{2n+m}$. Using (2.17) and (3.3), we have $K_{3,reg}u = \theta P_{3,\mathcal{C}}u$ if and only if

$$L^T \begin{bmatrix} \multicolumn{2}{c}{\multirow{2}{*}{$K_{2,\mathrm{reg}}$}} & 0 \\ & & 0 \\ 0 & 0 & -X \end{bmatrix} Lu = \theta L^T \begin{bmatrix} \multicolumn{2}{c}{\multirow{2}{*}{$P_{2,\mathcal{C}}$}} & 0 \\ & & 0 \\ 0 & 0 & -X \end{bmatrix} Lu,$$

and the result readily follows from the nonsingularity of L.

Concerning *ii)* we use again (2.17) and (3.3) and note that $P_{3,\mathcal{C}}^{-1}K_{3,\mathrm{reg}}\Delta_3 = P_{3,\mathcal{C}}^{-1}f_3$ if and only if

$$L^{-1} \begin{bmatrix} \multicolumn{2}{c}{\multirow{2}{*}{$P_{2,\mathcal{C}}$}} & 0 \\ & & 0 \\ 0 & 0 & -X \end{bmatrix}^{-1} \begin{bmatrix} \multicolumn{2}{c}{\multirow{2}{*}{$K_{2,\mathrm{reg}}$}} & 0 \\ & & 0 \\ 0 & 0 & -X \end{bmatrix} (L\Delta_3) = L^{-1} \begin{bmatrix} \multicolumn{2}{c}{\multirow{2}{*}{$P_{2,\mathcal{C}}$}} & 0 \\ & & 0 \\ 0 & 0 & -X \end{bmatrix}^{-1} L^{-T}f_3.$$

The result follows from explicitly writing

$$L\Delta_3 = (\Delta x, -\Delta y, X^{-1}Z^{\frac{1}{2}}\Delta x + Z^{-\frac{1}{2}}\Delta z) = (\Delta_2, X^{-1}Z^{\frac{1}{2}}\Delta x + Z^{-\frac{1}{2}}\Delta z), \quad (3.7)$$
$$L^{-T}f_3 = (f_x - X^{-1}f_z, f_y, Z^{-\frac{1}{2}}f_z) = (f_2, Z^{-\frac{1}{2}}f_z), \quad (3.8)$$

with Δ_2 and f_2 given in (2.13).

From their definition, we obtain

$$\begin{aligned} \widehat{P}_{3,\mathcal{C}}^{-1}\widehat{K}_{3,\mathrm{reg}} &= \widehat{L}_2^{-1} \begin{bmatrix} \multicolumn{2}{c}{\multirow{2}{*}{$P_{2,\mathcal{C}}^{-1}$}} & 0 \\ & & 0 \\ 0 & 0 & -X^{-1} \end{bmatrix} \widehat{L}_1^{-T}\widehat{L}_1^T \begin{bmatrix} \multicolumn{2}{c}{\multirow{2}{*}{$K_{2,\mathrm{reg}}$}} & 0 \\ & & 0 \\ 0 & 0 & -X \end{bmatrix} \widehat{L}_2 \\ &= \widehat{L}_2^{-1} \begin{bmatrix} \multicolumn{2}{c}{\multirow{2}{*}{$P_{2,\mathcal{C}}^{-1}K_{2,\mathrm{reg}}$}} & 0 \\ & & 0 \\ 0 & 0 & I_n \end{bmatrix} \widehat{L}_2, \end{aligned}$$

showing that the two preconditioned matrices are similar, leading to (*iii*).

The proof of (*iv*) parallels the arguments of (*ii*) and is based on (3.2). ∎

The spectral properties of $P_{2,\mathcal{C}}^{-1}K_{2,\mathrm{reg}}$ have been studied in a variety of papers, e.g., [7, 13–15, 28, 29], and in light of Theorem 3.1, these results apply to $P_{3,\mathcal{C}}^{-1}K_{3,\mathrm{reg}}$. In terms of performance, Theorem 3.1 shows that in exact arithmetic little can be gained by the unreduced formulation when the popular constraint preconditioner is employed; the situation might differ in finite precision arithmetic.

The comparison of the two approaches in the unsymmetric case, given in the above item (iv), requires further comments, as the performance of iterative preconditioned solvers such as Krylov subspace methods does not only depend on the eigenvalues of the two matrices. To this end, let us consider the right-preconditioned matrices $\widehat{K}_{3,\text{reg}}\widehat{P}_{3,\mathcal{C}}^{-1}$ and $K_{2,\text{reg}}P_{2,\mathcal{C}}^{-1}$, which is what we will use in our numerical experiments. In the 3×3 case in (2.8), for instance, the system to be solved is $\widehat{K}_{3,\text{reg}}\widehat{P}_{3,\mathcal{C}}^{-1}\Delta = \widehat{f}_3$, with $\widehat{\Delta}_3 = \widehat{P}_{3,\mathcal{C}}^{-1}\Delta$; analogously for the 2×2 system. We have that

$$\widehat{K}_{3,\text{reg}}\widehat{P}_{3,\mathcal{C}}^{-1} = \widehat{L}_1^T \begin{bmatrix} \begin{matrix} K_{2,\text{reg}}P_{2,\mathcal{C}}^{-1} \end{matrix} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & \quad 0 \end{matrix} & I_n \end{bmatrix} \widehat{L}_1^{-T}.$$

An interesting relation between the residuals obtained by using a Krylov subspace solver and the constraint preconditioner on the 3×3 and 2×2 systems can be established. Consider an approximate solution $\widehat{\Delta}^{(k)}$ in the Krylov subspace of dimension k generated with $\widehat{K}_{3,\text{reg}}\widehat{P}_{3,\mathcal{C}}^{-1}$ and $\widehat{f}_3 = (f_x, f_y, f_z)$ and zero starting approximate solution. For some polynomial φ_k of degree at most k, such that $\varphi_k(0) = 1$, the residual can be written as

$$\widehat{f}_3 - \widehat{K}_{3,\text{reg}}\widehat{P}_{3,\mathcal{C}}^{-1}\widehat{\Delta}^{(k)} = \begin{bmatrix} \varphi_k(K_{2,\text{reg}}P_{2,\mathcal{C}}^{-1})f_2 + \varphi_k(1)\begin{bmatrix} X^{-1}f_z \\ 0 \end{bmatrix} \\ \varphi_k(1)f_z \end{bmatrix}, \tag{3.9}$$

where f_2 is given in (2.13). Indeed,

$$\begin{aligned}
\widehat{f}_3 - \widehat{K}_{3,\text{reg}}\widehat{P}_{3,\mathcal{C}}^{-1}\widehat{\Delta}^{(k)} &= \varphi_k(\widehat{K}_{3,\text{reg}}\widehat{P}_{3,\mathcal{C}}^{-1})\widehat{f}_3 \\
&= \widehat{L}_1^T \varphi_k\left(\begin{bmatrix} \begin{matrix} K_{2,\text{reg}}P_{2,\mathcal{C}}^{-1} \end{matrix} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & \quad 0 \end{matrix} & I_n \end{bmatrix} \right) \widehat{L}_1^{-T}\widehat{f}_3 \\
&= \widehat{L}_1^T \begin{bmatrix} \begin{matrix} \varphi_k\big(K_{2,\text{reg}}P_{2,\mathcal{C}}^{-1}\big) \end{matrix} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & \quad 0 \end{matrix} & \varphi_k(1)I_n \end{bmatrix} \widehat{L}_1^{-T} f_3 \\
&= \widehat{L}_1^T \begin{bmatrix} \begin{matrix} \varphi_k\big(K_{2,\text{reg}}P_{2,\mathcal{C}}^{-1}\big) \end{matrix} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & \quad 0 \end{matrix} & \varphi_k(1)I_n \end{bmatrix} \begin{bmatrix} f_x - X^{-1}f_z \\ f_y \\ f_z \end{bmatrix} \\
&= \widehat{L}_1^T \begin{bmatrix} \varphi_k(K_{2,\text{reg}}P_{2,\mathcal{C}}^{-1})f_2 \\ \varphi_k(1)f_z \end{bmatrix},
\end{aligned}$$

and the last vector is precisely (3.9). The quantity $\varphi_k(K_{2,\text{reg}}P_{2,\mathcal{C}}^{-1})f_2$ is the same as that obtained by using a Krylov subspace on the 2×2 system, except that the polynomial φ_k is different, in general. Indeed, for a minimal residual method such as GMRES [36], φ_k is obtained so as to minimize the residual Euclidean norm. For the 3×3 case, φ_k must minimize the norm of both block vectors in (3.9), whereas in the 2×2 case, only the vector $\varphi_k(K_{2,\text{reg}}P_{2,\mathcal{C}}^{-1})f_2$ needs to be minimized in norm. Following the discussion in [35], we expect a similar behavior of the two residual norms if the minimizing polynomial for the 2×2 problem is small at one, which may be the case whenever one is enclosed in the spectral region of $K_{2,\text{reg}}P_{2,\mathcal{C}}^{-1}$. Clearly, the two problems become equivalent as soon as $f_z = 0$, and their residual norms are very close if f_z is small; we refer the reader to [35] for these and other considerations on the spectral properties of constraint-preconditioned Krylov subspace methods.

Finally, we observe that the spectral analysis of [31, §3] ensures that solving with $P_{3,\mathcal{C}}$ may remain a well conditioned problem throughout the IP iterations and at the limit point unlike with $P_{2,\mathcal{C}}$. Additional comments are postponed to the end of this section.

The second proof uses similar arguments.

THEOREM 3.2. *Suppose that H is symmetric positive semidefinite, X and Z are diagonal with positive entries, and $K_{2,\mathrm{reg}}$, $\widehat{K}_{3,\mathrm{reg}}$ and $K_{3,\mathrm{reg}}$ in (2.13), (2.9) and (2.11) are nonsingular. Let $P_{2,\mathcal{T}}$, $\widehat{P}_{3,\mathcal{T}}$ and $P_{3,\mathcal{T}}$ be the preconditioners (3.4), (3.5) and (3.6) respectively with $\kappa = 1$, and consider systems (2.10), (2.8), and (2.12). Then*

i) $\theta \in \Lambda(P_{3,\mathcal{T}}^{-1} K_{3,\mathrm{reg}})$ if and only if either $\theta = 1$ or $\theta \in \Lambda(P_{2,\mathcal{T}}^{-1} K_{2,\mathrm{reg}})$.

ii) Solving $P_{3,\mathcal{T}}^{-1} K_{3,\mathrm{reg}} \Delta_3 = P_{3,\mathcal{T}}^{-1} f_3$ is equivalent to solving $P_{2,\mathcal{T}}^{-1} K_{2,\mathrm{reg}} \Delta_2 = P_{2,\mathcal{T}}^{-1} f_2$ and recovering Δz from the third equation in (2.10).

iii) $\theta \in \Lambda(\widehat{P}_{3,\mathcal{T}}^{-1} \widehat{K}_{3,\mathrm{reg}})$ if and only if either $\theta = 1$ or $\theta \in \Lambda(P_{2,\mathcal{T}}^{-1} K_{2,\mathrm{reg}})$.

iv) The same feature in (ii) holds with $\widehat{K}_{3,\mathrm{reg}}$ and $\widehat{P}_{3,\mathcal{T}}$.

Proof. To characterize the spectrum of $P_{3,\mathcal{T}}^{-1} K_{3,\mathrm{reg}}$, we first observe that for L in (2.16),

$$L^{-T} P_{3,\mathcal{T}} L^{-1} = \begin{bmatrix} \multicolumn{2}{c}{P_{2,\mathcal{T}}} & 0 \\ & & 0 \\ Z^{\frac{1}{2}} & 0 & -X \end{bmatrix}, \tag{3.10}$$

and that the eigenvalue problem $K_{3,\mathrm{reg}} u = \theta P_{3,\mathcal{T}} u$, $u \in \mathbb{R}^{2n+m}$ can be written as

$$L^T \begin{bmatrix} \multicolumn{2}{c}{K_{2,\mathrm{reg}}} & 0 \\ & & 0 \\ 0 & 0 & -X \end{bmatrix} Lu = \theta L^T L^{-T} P_{3,\mathcal{T}} L^{-1} Lu.$$

Thus, by using (3.10)

$$\begin{bmatrix} \multicolumn{2}{c}{K_{2,\mathrm{reg}}} & 0 \\ & & 0 \\ 0 & 0 & -X \end{bmatrix} Lu = \theta \begin{bmatrix} \multicolumn{2}{c}{P_{2,\mathcal{T}}} & 0 \\ & & 0 \\ Z^{\frac{1}{2}} & 0 & -X \end{bmatrix} Lu. \tag{3.11}$$

With $\hat{u} = Lu = (\hat{u}_1, \hat{u}_2)$ where $\hat{u}_1 \in \mathbb{R}^{n+m}$, $\hat{u}_2 \in \mathbb{R}^n$, the first block row gives the eigenproblem $K_{2,reg} \hat{u}_1 = \theta P_{2,\mathcal{T}} \hat{u}_1$ while the second block row gives $-X\hat{u}_2 = \theta([Z^{\frac{1}{2}}, 0]\hat{u}_1 - X\hat{u}_2)$. Therefore, all eigenvalues of $(K_{2,reg}, P_{2,\mathcal{T}})$ are also eigenvalues of the unreduced problem with $\hat{u}_2$ given by the second block row. The remaining eigenvalues are obtained for $\hat{u}_1 = 0$, which gives $\theta = 1$. Note that if $\theta = 1$ is also an eigenvalue of $(K_{2,reg}, P_{2,\mathcal{T}})$, then the matrix pencil in (3.11) is not diagonalizable.

We now prove item *ii)*. By multiplying from the left by L in (2.16) and using (2.17), we can write the preconditioned system $P_{3,\mathcal{T}}^{-1} K_{3,\mathrm{reg}} \Delta_3 = P_{3,\mathcal{T}}^{-1} f_3$ as

$$L P_{3,\mathcal{T}}^{-1} L^T \begin{bmatrix} \multicolumn{2}{c}{K_{2,\mathrm{reg}}} & 0 \\ & & 0 \\ 0 & 0 & -X \end{bmatrix} L \Delta_3 = L P_{3,\mathcal{T}}^{-1} L^T L^{-T} f_3.$$

Since $LP_{3,\mathfrak{I}}^{-1}L^T = (L^{-T}P_{3,\mathfrak{I}}L^{-1})^{-1}$ and $L^{-T}f_3 = (f_2, Z^{-\frac{1}{2}}f_z)$, from (3.10) it follows that the system can be rewritten as

$$\begin{bmatrix} P_{2,\mathfrak{I}} & & 0 \\ & & 0 \\ Z^{\frac{1}{2}} & 0 & -X \end{bmatrix}^{-1} \begin{bmatrix} K_{2,\text{reg}} & & 0 \\ & & 0 \\ 0 & 0 & -X \end{bmatrix} L\Delta_3 = \begin{bmatrix} P_{2,\mathfrak{I}} & & 0 \\ & & 0 \\ Z^{\frac{1}{2}} & 0 & -X \end{bmatrix}^{-1} \begin{bmatrix} f_2 \\ Z^{-\frac{1}{2}}f_z \end{bmatrix}.$$

By (3.7) the first block equation coincides with the reduced preconditioned system while the third block equation is equivalent to the third equation in (2.10).

The proof of (*iii*) and (*iv*) parallels the above arguments and is based on (3.2) and on the equality

$$\widehat{L}_1^{-T}\widehat{P}_{3,\mathfrak{I}}\widehat{L}_2^{-1} = \begin{bmatrix} P_{2,\mathfrak{I}} & & 0 \\ & & 0 \\ Z & 0 & -X \end{bmatrix}. \qquad \square$$

Let now suppose that the 3×3 and 2×2 systems are solved by a Krylov subspace solver coupled with the triangular augmented preconditioner analyzed above. By (2.15) and

$$\widehat{P}_{3,\mathfrak{I}} = \widehat{L}_1^T \begin{bmatrix} P_{2,\mathfrak{I}} & & 0 \\ & & 0 \\ Z & 0 & -X \end{bmatrix} \widehat{L}_2,$$

we get

$$\begin{aligned}
\widehat{K}_{3,\text{reg}}\widehat{P}_{3,\mathfrak{I}}^{-1} &= \widehat{L}_1^T \begin{bmatrix} K_{2,\text{reg}} & & 0 \\ & & 0 \\ 0 & 0 & -X \end{bmatrix} \widehat{L}_2\widehat{L}_2^{-1} \begin{bmatrix} P_{2,\mathfrak{I}} & & 0 \\ & & 0 \\ Z & 0 & -X \end{bmatrix}^{-1} \widehat{L}_1^{-T} \\
&= \widehat{L}_1^T \begin{bmatrix} K_{2,\text{reg}} & & 0 \\ & & 0 \\ 0 & 0 & -X \end{bmatrix} \widehat{L}_2\widehat{L}_2^{-1} \begin{bmatrix} P_{2,\mathfrak{I}}^{-1} & 0 \\ & 0 \\ [X^{-1}Z \;\; 0]P_{2,\mathfrak{I}}^{-1} & -X^{-1} \end{bmatrix} \widehat{L}_1^{-T} \\
&= \widehat{L}_1^T \begin{bmatrix} K_{2,\text{reg}}P_{2,\mathfrak{I}}^{-1} & 0 \\ & 0 \\ -[Z \;\; 0]P_{2,\mathfrak{I}}^{-1} & I_n \end{bmatrix} \widehat{L}_1^{-T} =: \widehat{L}_1^T M \widehat{L}_1^{-T}.
\end{aligned}$$

Therefore, for any polynomial φ_k of degree not greater than k and setting $B = -[Z, 0]P_{2,\mathfrak{I}}^{-1}$, we have

$$\varphi_k(\widehat{K}_{3,\text{reg}}\widehat{P}_{3,\mathfrak{I}}^{-1}) = \widehat{L}_1^T\varphi_k(M)\widehat{L}_1^{-T} \tag{3.12}$$

$$= \widehat{L}_1^T \begin{bmatrix} \varphi_k(K_{2,\text{reg}}P_{2,\mathfrak{I}}^{-1}) & 0 \\ B\psi_{k-1}(K_{2,\text{reg}}P_{2,\mathfrak{I}}^{-1}) & \varphi_k(1)I \end{bmatrix} \widehat{L}_1^{-T}. \tag{3.13}$$

We thus find that

$$\varphi_k(\widehat{K}_{3,\text{reg}}\widehat{P}_{3,\mathfrak{I}}^{-1})\widehat{f} = \widehat{L}_1^T \begin{bmatrix} \varphi_k(K_{2,\text{reg}}P_{2,\mathfrak{I}}^{-1})f_2 \\ B\psi_{k-1}(K_{2,\text{reg}}P_{2,\mathfrak{I}}^{-1})f_2 + \varphi_k(1)f_z. \end{bmatrix}$$

Here ψ_{k-1} is as defined in [35, formula (2.7)], and the derivation is the same as in [27, Theorem 1.21] and [35, Lemma 2.2], except that it is lower block triangular instead of upper block triangular. Unlike the case discussed above for the constraint preconditioner, now the two residual minimizations do not become equivalent for $f_z = 0$, and the residual vectors are not expected to be close for small values of f_z.

The above theorem is valid as long as $\kappa = 1$ and the weight is the one employed so far. It does not further hold if the $(3,3)$ block is different from X or if augmentation is performed using only the Schur complement $J^T R_d^{-1} J$. From the previous theorem, the spectrum of the preconditioned matrices is fully characterized by the eigenvalues for $P_{2,\mathfrak{T}}^{-1} K_{2,\mathrm{reg}}$ and we refer to the results in [9, 24, 34]. As for the conditioning of matrices $P_{2,\mathfrak{T}}$ in (3.4), $\widehat{P}_{3,\mathfrak{T}}$ in (3.5), and $P_{3,\mathfrak{T}}$ in (3.6), we observe that the (1,1) block of the preconditioners may be badly affected by a small regularizing parameter δ and possibly by $X^{-1}Z$.

If we assume that R_d is inexpensive to invert, the only significant effort in the application of $P_{3,\mathfrak{T}}$, $\widehat{P}_{3,\mathfrak{T}}$ and $P_{2,\mathfrak{T}}$ is represented by the solution of a system with coefficient matrix

$$M := H + \rho I_n + X^{-1} Z + J^T R_d^{-1} J.$$

This solution is likely to be too expensive for the preconditioning strategy to be effective. Thus, in practice M should be replaced with a suitable approximation $\widetilde{M}$. The equivalence established in Theorem 3.2 between reduced and unreduced formulations still holds in this case. This is formalized in the following corollary, where for simplicity we consider only the symmetrized case.

Corollary 3.3. *Let $\widetilde{M}$ be an invertible matrix, and consider*

$$\widetilde{P}_{2,\mathfrak{T}} := \begin{bmatrix} \widetilde{M} & J^T \\ 0 & -R_d \end{bmatrix}, \qquad \widetilde{P}_{3,\mathfrak{T}} := \begin{bmatrix} \begin{matrix} \widetilde{P}_{2,\mathfrak{T}} \end{matrix} & \begin{matrix} -Z^{1/2} \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & -X \end{bmatrix}.$$

Then items (i) and (ii) of Theorem 3.2 still hold if $P_{3,\mathfrak{T}}$ is replaced with $\widetilde{P}_{3,\mathfrak{T}}$ and $P_{2,\mathfrak{T}}$ is replaced with $\widetilde{P}_{2,\mathfrak{T}}$.

Proof. We observe that relation (3.10) still holds for $\widetilde{P}_{3,\mathfrak{T}}$ and $\widetilde{P}_{2,\mathfrak{T}}$, i.e.

$$L^{-T} \widetilde{P}_{3,\mathfrak{T}} L^{-1} = \begin{bmatrix} \begin{matrix} \widetilde{P}_{2,\mathfrak{T}} \end{matrix} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} Z^{\frac{1}{2}} & 0 \end{matrix} & -X \end{bmatrix}.$$

The rest of the proof is similar to the one of Theorem 3.2. □

Summarizing our results, for the preconditioners considered, the spectral properties of the preconditioned 2×2 and 3×3 matrices are the same, and the computational performance of an iterative solver, at least in terms of number of iterations, is expected to be similar for the reduced and unreduced systems (the similarity matrices L and L_2 do not affect the first two blocks). On the other hand, due to the potential well-conditioning of the preconditioner in (3.3) in the limit of the IP process, $P_{3,\mathfrak{C}}$ seems to genuinely exploit the unreduced formulation and may justify the larger memory allocations and computational costs per iteration required with the unreduced formulation. An experimental analysis on constraint and augmented preconditioners is performed in §4.

3.2. Further relations between the 2×2 and 3×3 preconditioned systems. Further relationships between systems (2.8), (2.10) and (2.12) preconditioned by augmented diagonal and triangular preconditioners are shown in this section. The triangular preconditioners are $P_{2,\mathcal{T}}$ in (3.4), $\widehat{P}_{3,\mathcal{T}}$ in (3.5), and $P_{3,\mathcal{T}}$ in (3.6), and here we are interested in the case $\kappa \neq 1$. The diagonal preconditioners of type $P_{\mathcal{D}}$ in (2.18) are defined with the same augmentation, thus,

$$P_{2,\mathcal{D}} = \begin{bmatrix} H + \rho I_n + X^{-1}Z + J^T R_d^{-1} J & 0 \\ 0 & R_d \end{bmatrix}, \tag{3.14}$$

$$P_{3,\mathcal{D}} = \begin{bmatrix} H + \rho I_n + X^{-1}Z + J^T R_d^{-1} J & 0 & 0 \\ 0 & R_d & 0 \\ 0 & 0 & X \end{bmatrix} = \begin{bmatrix} \multicolumn{2}{c}{P_{2,\mathcal{D}}} & 0 \\ & & 0 \\ 0 & 0 & X \end{bmatrix}. \tag{3.15}$$

Observe that $P_{2,\mathcal{D}}$, $P_{3,\mathcal{D}}$ are positive definite.

We remark that in the regularized case, W is equal to the $(2,2)$ block of the matrices and this is an "optimal" choice for the diagonal preconditioner in terms of spectral distribution [31, §4.3].

We now show that, upon elimination of Δz, the 3×3 preconditioned system reduces to the 2×2 preconditioned system; note that here this property holds for any real κ in the block triangular preconditioner.

THEOREM 3.4. *Suppose that H is symmetric and positive semidefinite, X and Z are diagonal with positive entries, $K_{2,\mathrm{reg}}$ and $K_{3,\mathrm{reg}}$ given in (2.13) and (2.11) are nonsingular. Let $P_{2,\mathcal{T}}$ and $P_{3,\mathcal{T}}$ be the preconditioners in (3.4) and (3.6) respectively, with $\kappa \in \mathbb{R}$. Then, after the same elimination of Δz as in $K_{3,\mathrm{reg}}\Delta_3 = f_3$, the system $P_{3,\mathcal{T}}^{-1} K_{3,\mathrm{reg}} \Delta_3 = P_{3,\mathcal{T}}^{-1} f_3$ reduces to $P_{2,\mathcal{T}}^{-1} K_{2,\mathrm{reg}} \Delta_2 = P_{2,\mathcal{T}}^{-1} f_2$.*

The same feature holds with $P_{2,\mathcal{D}}$, $P_{3,\mathcal{D}}$ in (3.14), (3.15).

Proof. We first observe that

$$P_{3,\mathcal{T}}^{-1} = \begin{bmatrix} P_{2,\mathcal{T}}^{-1} & -\kappa P_{2,\mathcal{T}}^{-1} \begin{bmatrix} X^{-1} Z^{\frac{1}{2}} \\ 0 \end{bmatrix} \\ 0 \quad 0 & -X^{-1} \end{bmatrix}.$$

Then, by (2.17) the preconditioned system $P_{3,\mathcal{T}}^{-1} K_{3,\mathrm{reg}} \Delta_3 = P_{3,\mathcal{T}}^{-1} f_3$ can be written as

$$P_{3,\mathcal{T}}^{-1} L^T \begin{bmatrix} K_{2,reg} & 0 \\ 0 & -X \end{bmatrix} L \Delta_3 = P_{3,\mathcal{T}}^{-1} L^T L^{-T} f_3,$$

and takes the form

$$\begin{bmatrix} P_{2,\mathcal{T}}^{-1} K_{2,\mathrm{reg}} & -(1-\kappa) P_{2,\mathcal{T}}^{-1} \begin{bmatrix} Z^{\frac{1}{2}} \\ 0 \end{bmatrix} \\ 0 \quad 0 & I_m \end{bmatrix} L \Delta_3 = \begin{bmatrix} P_{2,\mathcal{T}}^{-1} & (1-\kappa) P_{2,\mathcal{T}}^{-1} \begin{bmatrix} X^{-1} Z^{\frac{1}{2}} \\ 0 \end{bmatrix} \\ 0 \quad 0 & -X^{-1} \end{bmatrix} L^{-T} f_3.$$

Finally, by (3.7) and (3.8) it readily follows that by back substitution of Δz the first block equation coincides with the reduced preconditioned system.

The claim for the diagonal preconditioners $P_{2,\mathcal{D}}$, $P_{3,\mathcal{D}}$ in (3.14), (3.15) can be proved by repeating the above arguments. □

The previous result is still valid if we consider the systems (2.12) and (2.8) and either the preconditioners $P_{2,\mathcal{T}}$ and $\widehat{P}_{3,\mathcal{T}}$ in (3.4), (3.5) or $P_{2,\mathcal{D}}$, $P_{3,\mathcal{D}}$ in (3.14), (3.15); the proof is omitted.

4. Qualitatively different preconditioning strategies. Our previous results show that the reduced and unreduced formulations are closely related, also when a large class of preconditioning strategies is used. To be able to investigate their true potential in a comparative manner, we need to select the free parameters of these acceleration strategies and to apply the preconditioners in an ad-hoc way.

In this section, we experiment with an IP method for problem (2.1). Our aim is to compare the performance of the preconditioners of the form $P_{\mathfrak{D}}$, $P_{\mathfrak{I}}$, $T_{\mathcal{W}}$, $P_{\mathfrak{C}}$ in (2.18)–(2.21) and to assess whether the unreduced formulation can be advantageous in terms of conditioning and execution time relative to the reduced one. We put emphasis on the solution of the whole sequence of linear systems generated with each formulation.

The numerical experiments were conducted using MATLAB R2015a on a 12x Intel Core i7-5820K, 3.30GHz, 64 GB of RAM processor. Elapsed times were measured by the `tic` and `toc` MATLAB commands.

The datasets used are eight convex QP problems from the CUTEr collection [23] (CVXQP1–CVXQP3, STCQP1_a–STCQP2, GOULDQP3_a–GOULDQP3_b), and four problems from the Maros and Mészáros collection [30] (QSHIP12_s, QSHIP12_l, LASER and QSCTAP3). Possible linear inequality constraints have been reformulated as equality constraints by using a slack variable. The features of these problems are summarized in Table 4.1. In all the considered datasets H is not a diagonal matrix and the number of its nonzero elements is generally quite large, so as to justify the use of an iterative solver. The last four data sets have medium dimensions and the matrices H and J are very sparse. In spite of their relatively small dimensions, these problems may be representative of the spectral features of larger matrices stemming from similar application problems, thus offering some insights into conditioning issues. The considered data from the CUTEr collection are the largest ones in the data set with H non-diagonal.

TABLE 4.1
Test problems: values of n, m, nonzeros in H and in J

Problem	n	m	nnz(H)	nnz(J)
CVXQP1	10000	5000	69968	14998
CVXQP2	10000	7500	69968	7499
CVXQP3	10000	7500	69968	22497
STCQP1_a	8193	4095	106473	28865
STCQP1_b	16385	8190	229325	61425
STCQP2	16385	8190	229325	61425
GOULDQP3_a	9999	4999	29993	14997
GOULDQP3_b	19999	9999	59993	29997
QSHIP12_s	2869	1151	33764	8284
QSHIP12_l	5533	1151	122433	16276
LASER	3002	2000	5430	8000
QSCTAP3	3340	1480	1908	9734

The QP problems were solved by the MATLAB code PDCO (Primal Dual interior method for optimization with Convex Objectives) developed by Michael Saunders and available in [33]. In PDCO, nonnegativity constraints are replaced by the log barrier function and a sequence of convex barrier subproblems is generated; for each barrier subproblem, one step of Newton's method is applied to the KKT conditions with perturbed complementarity conditions. In order to improve solver stability in PDCO it is suggested to regularize (2.1) as

$$\min_{x,r}\ c^T x + \frac{1}{2}x^T H x + \frac{1}{2}\|D_1 x\|^2 + \frac{1}{2}\|r\|^2 \quad \text{subject to} \quad Jx + D_2 r = b,\ x \geq 0,$$

where D_1 and D_2 are positive definite diagonal matrices specified by the user. Therefore, with $D_1 = \sqrt{\rho} I_n$ and $D_2 = \sqrt{\delta} I_m$ for positive ρ and δ, the unreduced and reduced coefficient matrices of the systems to be solved take the form $K_{2,\text{reg}}$, $\widehat{K}_{3,\text{reg}}$ and $K_{3,\text{reg}}$ respectively.

We applied regularization only when needed, that is we set $\rho = \delta = 0$ everywhere except in the following cases. In problems STCQP1_A, STCQP1_B, QSHIP12_S, QSHIP12_L we used $\delta > 0$, otherwise $\widehat{K}_3$ is singular at every iteration due to the rank deficiency of the matrix J (see Theorem 2.1). In problems CVXQP1, CVXQP2, CVXQP3 we used $\delta > 0$ because otherwise $\widehat{K}_3$ is singular in the limit, since the Linear Independence constraint Qualification is not satisfied (see Theorem 2.2). For these cases, we set $\delta = 10^{-6}$. Moreover, to overcome failures of PDCO in the unregularized setting, we set $\delta = 10^{-6}$ for problem LASER and $\rho = 10^{-6}$ for problem QSCTAP3.

PDCO allows one to work with two alternative linear system formulations: a reduced form with matrix $K_{2,\text{reg}}$ and a condensed formulation where the coefficient matrix is a Schur complement. Symmetric indefinite systems are solved by a direct solver whereas for condensed systems both direct and iterative solvers are available. We modified PDCO so that the unreduced regularized formulation with matrices $\widehat{K}_{3,\text{reg}}$, $K_{3,\text{reg}}$ could also be explicitly formed. For the 3×3 formulation, the unsymmetric form (2.9) will be used when indefinite preconditioners will be employed. Indeed, the process of symmetrization (2.6) requires the application of matrix R, which becomes increasingly ill-conditioned along the IP process and may also have a detrimental effect on the scaling of the right-hand side of the linear systems. Summarizing, the systems in the sequence were solved by using: i) MATLAB's sparse direct solver (function "\"); ii) MINRES [17, 32] coupled with a diagonal augmented preconditioners $P_{\mathcal{D}}$ in (2.18) for the symmetric matrices $K_{3,\text{reg}}$, $K_{2,\text{reg}}$; iii) GMRES [36] coupled with the indefinite constraint preconditioner $P_{\mathcal{C}}$ (see (2.21)) and the augmented triangular preconditioners $P_{\mathcal{T}}$ (see (2.19) and $T_{\mathcal{W}}$ (see (2.20)) for the matrices $\widehat{K}_{3,\text{reg}}$, $K_{2,\text{reg}}$. We do not report the results obtained with $P_{\mathcal{T}}$ as they were slightly worse than those obtained with the other triangular preconditioner $T_{\mathcal{W}}$.

We are aware that the sparse direct solver is a compiled code whereas the iterative solvers are Matlab implementations (interpreted) and thus slower; nonetheless, we will show runs where the iterative methods can largely overcome this disadvantage.

4.1. Implementation issues in preconditioning. The distinguishing feature of the preconditioner $\widehat{P}_{3,\mathcal{C}}$ given in (3.2) with respect to $P_{2,\mathcal{C}}$ in (3.1) is that the former is invertible and possibly well conditioned throughout the IP iterations as long as $K_{3,\text{reg}}$ is invertible at the solution. The application of $\widehat{P}_{3,\mathcal{C}}$ requires solves that can be performed either via a sparse complete or incomplete LU factorization, or by the following standard a-priori block decomposition Letting $D = \text{diag}(H + \rho I_n)$, as long as D is invertible, we have

$$\widehat{P}_{3,\mathcal{C}} = \begin{bmatrix} I_n & 0 & 0 \\ JD^{-1} & I_m & 0 \\ -ZD^{-1} & 0 & I_n \end{bmatrix} \begin{bmatrix} D & 0 & 0 \\ 0 & & \\ 0 & -S_1 & \end{bmatrix} \begin{bmatrix} I_n & D^{-1}J^T & -D^{-1} \\ 0 & I_m & 0 \\ 0 & 0 & I_n \end{bmatrix}, \tag{4.1}$$

with

$$S_1 = \begin{bmatrix} J \\ -Z \end{bmatrix} D^{-1} \begin{bmatrix} J^T & -I_n \end{bmatrix} + \begin{bmatrix} \delta I_m & 0 \\ 0 & X \end{bmatrix} = \begin{bmatrix} JD^{-1}J^T + \delta I_m & -JD^{-1} \\ -ZD^{-1}J^T & ZD^{-1} + X \end{bmatrix},$$

and, since $ZD^{-1} + X$ is diagonal, it holds

$$S_1 = \begin{bmatrix} I_m & -J(Z+DX)^{-1} \\ 0 & I_n \end{bmatrix} \begin{bmatrix} S_2 & 0 \\ 0 & ZD^{-1} + X \end{bmatrix} \begin{bmatrix} I_m & 0 \\ -(Z+DX)^{-1}ZJ^T & I_n \end{bmatrix},$$

with

$$S_2 = JX\left(Z + DX\right)^{-1} J^T + \delta I_m. \tag{4.2}$$

This matrix S_2 is the same Schur complement one obtains by block factorizing the 2×2 formulation, that is

$$P_{2,\mathcal{C}} = \begin{bmatrix} I_n & 0 \\ J\left(D + X^{-1}Z\right)^{-1} & I_m \end{bmatrix} \begin{bmatrix} D + X^{-1}Z & 0 \\ 0 & -S_2 \end{bmatrix} \begin{bmatrix} I_n & \left(D + X^{-1}Z\right)^{-1} J^T \\ 0 & I_m \end{bmatrix}. \tag{4.3}$$

It holds that $JX\left(Z+DX\right)^{-1} J^T = JD^{-1}(D^{-1}X^{-1}Z + I)^{-1}J^T$, and the (diagonal) entries of $(D^{-1}X^{-1}Z + I)^{-1}$ belong to the interval $(0, 1)$ with the smallest value being of the order of $\min_i x_i$ if strict complementarity holds at the solution. Thus, the eigenvalues of S_2 belong to the interval $[\delta, \delta + \|JD^{-1}J^T\|]$ throughout the IP iterations and the lower bound is possibly achieved as $\min_i x_i$ tends to zero. As a consequence, the condition number of S_2 may be large, e.g., in the case when H is positive semidefinite and the regularization parameters δ and ρ are small. This ill-conditioning correctly reflects that of $P_{2,\mathcal{C}}$, whereas in the 3×3 case S_2 may be much more ill-conditioned than the original preconditioner.

As for the augmented preconditioners $P_{\mathcal{D}}$, $P_{\mathcal{T}}$, $T_{\mathcal{W}}$ in (2.18)–(2.20), their effectiveness is very sensitive to the choice of the weight matrix W; from a computational point of view, it is convenient to choose W diagonal so that it is inexpensive to invert. We know that an optimal choice of W in terms of eigenvalue distribution of the preconditioned matrices is $W = C$ [31, §4.3], [38, Remark 2.1]. For our purposes, however, this choice is not interesting, since we have shown in Theorems 3.2 and 3.4, and Corollary 3.3 that this choice makes the unreduced and reduced formulations behave very similarly. Moreover, we notice that $W = C$ is optimal only in the presence of regularization in the $(2, 2)$ block.

We have used a diagonal augmented preconditioner $P_{\mathcal{D}}$ of the form (2.18), and a triangular augmented preconditioner $T_{\mathcal{W}}$ of the form (2.20). For the definition of $T_{\mathcal{W}}$, following [38], we set $\kappa = 2$ in (2.20) and consequently

$$T_{\mathcal{W}} = \begin{bmatrix} E + F^TW^{-1}G & 2F^T \\ 0 & -(W + C) \end{bmatrix}. \tag{4.4}$$

As for the unreduced case (2.9) we let

$$E = H + \rho I_n, \quad F = \begin{bmatrix} J \\ -I_n \end{bmatrix}, \quad G = \begin{bmatrix} J \\ -Z \end{bmatrix}, \quad C = \begin{bmatrix} \delta I_m & 0 \\ 0 & X \end{bmatrix},$$

while in the reduced case (2.13) we have

$$E = H + \rho I_n + X^{-1}Z, \quad F = G = J, \quad C = \delta I_m.$$

Approximations such as $W = \gamma I_{n+m}$ with γ equal to either the arithmetic mean, geometric mean or median of $\mathrm{diag}(C)$ actually led to slow convergence of preconditioned MINRES and GMRES. In the spirit of the proposal in [34], for the unreduced systems we obtained an effective choice of W by setting

$$W = \begin{bmatrix} \gamma_1 I_m & 0 \\ 0 & \gamma_2 I_n \end{bmatrix}, \tag{4.5}$$

where

$$\gamma_1 = \frac{\alpha}{\|H + \rho I_n\|_F}, \qquad \gamma_2 = \gamma_1 \cdot \mathrm{mean}(z), \tag{4.6}$$

α is a positive parameter, and $\|\cdot\|_F$ is the Frobenius norm. In all our experiments we set $\alpha = 10^{-2}$. For the reduced system $K_{2,\mathrm{reg}}$, we instead set

$$W = \gamma I_m, \qquad \gamma = \frac{1}{\|H + \rho I_n + X^{-1}Z\|_F}. \tag{4.7}$$

This choice of W, in both the reduced and unreduced cases, makes the preconditioner robust with respect to the scaling of the variables. Specifically, consider the unreduced formulation. When the variables x and z are scaled by some positive parameters s_x and s_z, respectively, the resulting coefficient matrix in (2.10) takes the form $\widetilde{K}_{3,reg} = T^{1/2} K_{3,\mathrm{reg}} T^{1/2}$, where T is the diagonal matrix

$$T = \begin{bmatrix} \frac{s_x}{s_z} & 0 & 0 \\ 0 & \frac{s_z}{s_x} & 0 \\ 0 & 0 & \frac{1}{s_x} \end{bmatrix}.$$

The augmented block diagonal preconditioner (3.15) constructed for the scaled system, with W chosen as in (4.5), has the form $\widetilde{P}_{\mathcal{D}} = T^{1/2} P_{\mathcal{D}} T^{1/2}$. Thus, the eigenvalues of the preconditioned scaled matrix $\widetilde{P}_{\mathcal{D}}^{-1} \widetilde{K}_{3,reg}$ are the same as those of the preconditioned unscaled matrix $P_{\mathcal{D}}^{-1} K_{3,\mathrm{reg}}$. Similar relations hold for preconditioners $P_{\mathcal{T}}$ and $T_{\mathcal{W}}$, and also for the reduced formulation.

The choice (4.5) for W exploits the partitioning of the original matrix by giving different weights at the two diagonal blocks. If the problem is well scaled, the values of γ_1 and γ_2 are expected to be of moderate size and bounded away from zero; this feature may mitigate ill-conditioning in the (1,1) block of the preconditioner and the following numerical experiments support this claim. On the other hand, we remark that γ in (4.7) is expected to tend to 0.

4.2. Selected numerical results. We report selected numerical results on the inexact implementation of the IP method in PDCO and the comparison of different linear system formulations and linear algebra solvers. Once again, our main aim is to report on our numerical experience while comparing the reduced and unreduced preconditioned formulations. Nonetheless, we report and discuss the performance of different approaches, such as that of direct solvers, which should be a reference when trying to attack the solution of this class of problems by means of iterative methods.

The issue of how much inaccuracy is allowed in the solution of the linear systems while preserving the convergence properties of IP methods has been investigated in several papers [3,4,7,8,16,22,29]. The norm of the residual in the linear systems can be directly linked to the barrier parameter $\mu = x^T z/n$ in (2.4) whereas the convergence of the IP method is driven by the barrier term tending to zero. This setting implies an increasing accuracy in the solution of the linear system as the IP iterate approaches the solution. The resulting IP method can be interpreted as an Inexact Newton method [12] and its convergence can be analyzed in such framework [4]. On the basis of the mentioned literature we solve iteratively the unreduced system (2.8) and monitor the unpreconditioned system residual until

$$\|\widehat{K}_{3,\mathrm{reg}}\,\widehat{\Delta}_3 - \widehat{f}_3\| \le \eta\mu, \tag{4.8}$$

with η being a forcing term in $(0,1)$. Analogously for system (2.10), the control on inexactness is $\|K_{3,\mathrm{reg}}\,\Delta_3 - f_3\| \le \eta\mu$. For the 2×2 formulation, we solve iteratively (2.12) until the residual satisfies

$$\|K_{2,\mathrm{reg}}\Delta_2 - f_2\| \le \eta\mu. \tag{4.9}$$

Successively, the step Δz is recovered from the third block equation in (2.8). .

The following results were obtained by applying MINRES and GMRES with stopping criterion (4.9) and η equal to 10^{-2}. In the whole section, the total execution times for the iterative solution of the systems include the time needed to form the preconditioner and its factorization and the time for the iterative solver. PDCO was stopped with accuracy on feasibility (`FeaTol`) equal to 10^{-6} and complementarity tolerance (`OptTol`) equal to 10^{-6}.

In PDCO, variable scaling can be applied if estimates on the largest components of the solution x and z are available, and this may improve the code performance. We assumed no a priori knowledge on the solution and thus no scaling was implemented.

Table 4.2

Conditioning data of $K_{3,\mathrm{reg}}$, $K_{2,\mathrm{reg}}$, expressed in the form $10^{\min - \max}$. For $P_{\mathcal{D}}$: conditioning of the (1,1) block Cholesky factor. For $P_{\mathcal{C}}$: conditioning of L and U factors (unreduced case) and of the Cholesky factor of the Schur complement in the reduced case.

	$K_{3,\mathrm{reg}}$				$K_{2,\mathrm{reg}}$		
	Backslash	$P_{\mathcal{C}}$		$P_{\mathcal{D}}$	Backslash	$P_{\mathcal{C}}$	$P_{\mathcal{D}}$
	$\kappa_e(K_{3,\mathrm{reg}})$	$\kappa_e(L)$	$\kappa_e(U)$	$\kappa_e(L)$	$\kappa_e(K_{2,\mathrm{reg}})$	$\kappa_e(L)$	$\kappa_e(L)$
Problem	min-max	min-max	min-max	min-max	min-max	min-max	min-max
CVXQP1	11-12	7-8	15-17	3-3	11-22	2-4	2-6
CVXQP2	11-11	5-6	11-13	2-3	11-21	2-3	2-6
CVXQP3	11-12	7-8	16-17	2-3	12-22	3-4	3-5
STCQP1_a	10-10	5-5	12-12	3-4	10-19	4-5	3-6
STCQP1_b	10-10	6-6	13-13	4-4	10-19	4-5	3-6
STCQP2	7-7	2-3	5-5	4-4	7-15	1-2	3-8
GOULDQP3_a	2-7	1-4	1-6	3-4	1-11	0-4	2-5
GOULDQP3_b	2-7	1-4	1-6	3-4	1-10	0-4	2-5
QSHIP12_s	12-12	2-4	11-14	2-3	10-22	2-8	3-8
QSHIP12_l	11-13	2-4	11-14	3-5	10-23	2-7	3-8
LASER	10-13	0-3	10-14	0-4	8-25	0-7	1-10
QSCTAP3	4-9	3-6	7-12	6-9	4-12	2-4	5-9

We start by analyzing the conditioning of matrices $K_{3,\mathrm{reg}}$, $K_{2,\mathrm{reg}}$, and the corresponding preconditioners. Condition numbers $\kappa_e(\cdot)$ were estimated by the MATLAB function `condest` and only the order of magnitude exponent is reported. In the unreduced formulation, we decided to avoid using the factored form of $P_{\mathcal{C}}$ in (4.1) since the matrix D may not be invertible, and the use of matrix S_2 in (4.2) would introduce unnecessary ill-conditioning. Thus, preconditioner $P_{\mathcal{C}}$ was applied using the factorization $\Pi P_{\mathcal{C}}\, Q = LU$. Column reordering (matrix Q) reduces fill-in in the sparse case and is commonly more time and memory efficient than just row pivoting via the permutation matrix Π, although we observed that the condition number of the factors L and U may be considerably higher than the conditioning of $P_{\mathcal{C}}$. However, we should say that the IP method did not display any anomalous behavior related to the conditioning.

In Table 4.2 we display the observed condition numbers, for both (unreduced and reduced) formulations. We report the minimum and maximum condition numbers of the system matrix and of the Cholesky factor L of the (1,1) block of $P_{\mathcal{D}}$ during the IP iterations. For the unreduced case, where $P_{\mathcal{C}}$ is applied via direct factorization, we also report the conditioning of its L and U factors. For the reduced case, where $P_{\mathcal{C}}$ is applied via the Schur complement factorization (4.3), we report the conditioning of the Cholesky factor L of S_2. Since the condition number of $K_{3,\mathrm{reg}}$ and $K_{2,\mathrm{reg}}$ did not vary significantly in runs with different linear algebra solvers, we report only the values obtained with the MATLAB's direct solver. Similarly, we do not show the condition number of the Cholesky factors of the (1,1) block of $T_{\mathcal{W}}$ as they are very similar to those of $P_{\mathcal{D}}$.

As expected, the maximum value attained by the conditioning of $K_{3,\mathrm{reg}}$ is several orders of magnitude smaller than that of $K_{2,\mathrm{reg}}$. In the unreduced formulation the condition number of the Cholesky factor L of the (1,1) block of $P_{\mathcal{D}}$ is at most a couple of orders of magnitude smaller than the condition number of the Cholesky factor for $K_{2,\mathrm{reg}}$.

Next, we present the performance of PDCO with the discussed strategies for the linear algebra phase. As for the augmented preconditioners, in the first eight problems $P_{\mathcal{D}}$ and $T_{\mathcal{W}}$ were approximated by replacing $E + F^T W^{-1} G$ with its incomplete Cholesky factors computed with threshold 10^{-3}; larger thresholds led to a breakdown of the factorization. In the remaining problems, the exact Cholesky factors were computed as the incomplete factorization failed with threshold down to 10^{-4}.

Interestingly, for all problems the number of IP iterations was insensitive to the linear algebra solver used in the inner phase, for both formulations. This seems to imply that with the chosen stopping criterion for the inner approximate solution, the selection criterion of the inner solver should be based on performance, also when comparing the reduced and unreduced formulations. Moreover, the iterative solution of the inner systems does not degrade the overall performance of the IP method compared with the sparse direct solver.

Table 4.3

Unreduced system. Number of linear systems solved (outer iterations), number of inner iterations required by the iterative solvers.

$K_{3,\mathrm{reg}}$				
		$P_{\mathcal{C}}$-GMRES	$P_{\mathcal{D}}$-MINRES	$T_{\mathcal{W}}$-GMRES
Problem	Outer Iter	Inner Iter min/max/avg	Inner Iter min/max/avg	Inner Iter min/max/avg
CVXQP1	21	9/22/18.9	32/114/ 72.0	15/ 34/28.5
CVXQP2	22	20/23/22.2	51/ 90/ 64.2	21/ 27/24.7
CVXQP3	23	9/28/22.5	34/118/ 88.1	16/ 42/36.3
STCQP1_a	17	2/ 9/ 4.1	22/ 30/ 26.5	17/ 22/18.5
STCQP1_b	17	3/10/ 4.3	29/ 38/ 33.3	20/ 25/22.1
STCQP2	18	2/15/ 4.6	12/ 26/ 20.2	8/ 14/11.4
GOULDQP3_a	13	7/10/ 7.8	36/241/115.9	14/136/57.3
GOULDQP3_b	14	7/11/ 8.0	38/526/202.2	15/208/88.6
QSHIP12_s	56	3/ 5/ 4.2	57/111/ 91.0	21/ 48/34.7
QSHIP12_l	61	3/ 5/ 4.3	50/161/109.4	19/ 94/46.1
LASER	31	2/ 8/ 4.9	19/131/ 57.1	14/ 90/35.3
QSCTAP3	31	4/ 7/ 6.1	14/ 30/ 22.4	5/ 8/ 5.6

Table 4.4
Unreduced systems. Number of linear systems solved (outer iterations) and total execution times (secs) for solving the sequence of unreduced systems. (Timings of Backslash are included for completeness.)

$K_{3,\mathrm{reg}}$					
		Backslash	$P_{\mathcal{C}}$-GMRES	$P_{\mathcal{D}}$-MINRES	$T_{\mathcal{W}}$-GMRES
Problem	Outer Iter	Total time	Total time	Total time	Total time
CVXQP1	21	156.93	5.54	4.09	2.99
CVXQP2	22	63.37	1.80	2.98	2.12
CVXQP3	23	275.24	21.14	6.80	5.10
STCQP1_a	17	29.30	3.32	1.33	1.37
STCQP1_b	17	137.31	15.86	3.50	3.43
STCQP2	18	2.22	0.79	2.16	1.89
GOULDQP3_a	13	0.52	0.58	2.40	10.44
GOULDQP3_b	14	1.13	1.28	7.61	44.21
QSHIP12_s	56	2.08	1.20	9.85	10.74
QSHIP12_l	61	7.59	2.73	72.32	51.55
LASER	31	0.44	0.59	1.65	4.38
QSCTAP3	31	0.91	1.11	2.55	1.22

Table 4.5
Reduced systems. Number of linear systems solved (outer iterations), number of inner iterations required by the iterative solvers.

$K_{2,\mathrm{reg}}$				
		$P_{\mathcal{C}}$-GMRES	$P_{\mathcal{D}}$-MINRES	$T_{\mathcal{W}}$-GMRES
Problem	Outer Iter	Inner Iter min/max/avg	Inner Iter min/max/avg	Inner Iter min/max/avg
CVXQP1	21	9/22/18.8	31/ 91/68.6	14/ 33/27.3
CVXQP2	22	20/23/22.1	35/ 52/43.1	19/ 27/24.0
CVXQP3	23	9/28/22.5	36/109/84.9	16/ 40/35.5
STCQP1_a	17	2/ 9/ 4.1	19/ 28/22.1	12/ 24/14.3
STCQP1_b	17	3/10/ 4.3	26/ 36/29.1	15/ 31/18.1
STCQP2	18	2/15/ 4.6	7/ 13/ 8.7	8/ 12/ 8.6
GOULDQP3_a	13	7/10/ 7.8	9/107/50.9	7/ 60/30.9
GOULDQP3_b	14	7/11/ 8.0	10/190/86.4	7/102/46.7
QSHIP12_s	56	3/ 5/ 4.2	10/ 22/18.8	2/ 11/ 3.6
QSHIP12_l	61	3/ 5/ 4.3	14/ 23/19.7	2/ 11/ 3.5
LASER	31	2/ 8/ 4.9	6/ 23/11.8	2/ 6/ 2.8
QSCTAP3	31	4/ 7/ 6.1	3/ 8/ 4.3	3/ 4/ 3.4

Table 4.3 shows the number of linear systems solved (outer iterations) and the minimum, maximum and average (min/max/avg) number of inner iterations required by the iterative solvers $P_{\mathcal{C}}$-GMRES, $P_{\mathcal{D}}$-MINRES and $T_{\mathcal{W}}$-GMRES in the 3×3 case. Table 4.4 displays the total execution time for solving the sequence of systems with Matlab's direct solver, $P_{\mathcal{C}}$-GMRES, $P_{\mathcal{D}}$-MINRES and $T_{\mathcal{W}}$-GMRES. For the sake of completeness, timings of the Matlab's direct solver are also reported.

Comparisons with sparse direct solvers should in general be performed on more challenging, larger and possibly denser problems. Nonetheless, even on these contrived examples, the performance of the iterative solvers is in general satisfactory. Analogous statistics for the solution of the sequences of reduced systems are given in Tables 4.5 and 4.6.

Tables 4.3 and 4.5 show that in general preconditioned GMRES and MINRES behave somewhat similarly in the two formulations in terms of number of iterations, except that for the GOULDQP3 data, where the reduced system requires significantly fewer iterations. Closeness of the residual norms of $P_{\mathcal{C}}$-GMRES for both the 3×3 and 2×2 systems was experimentally verified in the whole IP process, see the discussion in Section 3.

Tables 4.4 and 4.6 report timings for the unreduced and reduced systems, respectively. All vectors have smaller dimensions in the reduced formulation, and this shows up both in the memory requirements and in the elapsed times, apart from the two runs on problems STCQP1_a and STCQP2_a with preconditioner $P_{\mathcal{C}}$, where the factorized Schur complement appears to be denser than for the other cases, and thus expensive to solve with.

A few comments on performance within each formulation can also be made. In the 3×3 formulation, runs with $P_{\mathcal{C}}$-GMRES show significant CPU time reductions over the direct solver in most of the cases; see Table 4.4. Augmented preconditioners outperform $P_{\mathcal{C}}$ in problems STCQP1_a and STCQP1_b, while there is clear computational disadvantage from their application in the last six problems. Concerning the reduced formulation, the computational times in Table 4.6 are low but iterative solvers still produce significant savings in the first six problems.

In light of our theoretical analysis and numerical experience, we can conclude that the better spectral properties of the unreduced formulation and possibly of the corresponding preconditioners, do not provide sufficient benefits in terms of robustness and efficiency. The step computed through iterative solvers applied to the 2×2 systems are accurate enough to guarantee the success of the IP method, and the smaller dimension provides computational and time savings in the IP procedure.

Table 4.6

Reduced systems. Number of linear systems solved (outer iterations) and total execution times (secs) for solving the sequence of unreduced systems. (Timings of Backslash are included for completeness.)

$K_{2,\mathrm{reg}}$					
		Backslash	$P_{\mathcal{C}}$-GMRES	$P_{\mathcal{D}}$-MINRES	$T_{\mathcal{W}}$-GMRES
Problem	Outer Iter	Total time	Total time	Total time	Total time
CVXQP1	21	10.10	1.31	3.48	2.18
CVXQP2	22	7.24	1.27	1.80	1.59
CVXQP3	23	15.15	3.03	6.03	4.12
STCQP1_a	17	2.57	7.78	1.03	1.04
STCQP1_b	17	8.92	31.58	2.87	2.75
STCQP2	18	6.56	0.55	1.34	1.48
GOULDQP3_a	13	0.23	0.23	0.88	1.86
GOULDQP3_b	14	0.50	0.45	2.47	7.67
QSHIP12_s	56	0.63	0.55	3.18	1.76
QSHIP12_l	61	2.59	0.99	17.27	6.78
LASER	31	0.19	0.31	0.37	0.25
QSCTAP3	31	0.25	0.33	0.94	1.17

5. Conclusions. The high sparsity structure of KKT systems arising in the numerical solution of quadratic programming problems by means of interior point methods encourages the use of reduction strategies before solving the systems. For stability purposes, however, the unreduced formulation may be appealing. Following previous analysis in [18, 19, 25, 31], we have explored these two formulations when preconditioned iterative solvers are applied in the inner phase.

We have shown that the application of certain classes of preconditioners to the two formulations may result in either mathematically equivalent solution procedures, or to strictly related sequences of approximations. Moreover, our experimental analysis illustrates that once specifically designed preconditioners are used, and as long as implementations with the unreduced and reduced formulations are successful, the CPU time performance seems to favor the latter, in view of the smaller dimensions.

Acknowledgments. We wish to thank Michael Saunders for his many comments and suggestions on an earlier draft of this manuscript.

REFERENCES

[1] A. Altman and J. Gondzio. Regularized symmetric indefinite systems in interior point methods for linear and quadratic optimization. *Optim. Meth. and Softw.*, 11:275–302, 1999.

[2] O. Axelsson. Preconditioners for regularized saddle point matrices. *J. Numer. Math.*, 19(2):91–112, 2011.

[3] V. Baryamureeba and T. Steihaug. On the convergence of an inexact primal-dual interior point method for linear programming. In I. Lirkov, S. Margenov, and J. Wasniewski, editors, *Proceedings of the 5th International Conference on Large-Scale Scientific Computing*, Lecture Notes in Computer Science 3743, pages 629–637, Berlin, 2006. Springer-Verlag.

[4] S. Bellavia. An inexact interior point method. *J. Optim. Theory App.*, 96:109–121, 1998.

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

[6] M. Benzi and M. A. Olshanskii. An augmented Lagrangian-based approach to the Oseen problem. *SIAM J. Sci. Comput.*, 28:2095–2113, 2006.

[7] L. Bergamaschi, J. Gondzio, and G. Zilli. Preconditioning indefinite systems in interior point methods for optimization. *Comput. Optim. Appl.*, 28:149–171, 2004.

[8] S. Cafieri, M. D'Apuzzo, V. De Simone, and D. di Serafino. Stopping criteria for inner iterations in inexact potential reduction methods: a computational study. *Comput. Optim. Appl.*, 36:165–193, 2007.

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

[10] J. Castro and J. Cuesta. Quadratic regularizations in an interior-point method for primal block-angular problems. *Math. Prog.*, 130:415–445, 2011.

[11] 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. *Comput. Optim. Appl.*, 45:283–310, 2010.

[12] R.S. Dembo, S.C. Eisenstat, and T. Steihaug. Inexact newton methods. *SIAM J. Numer. Anal.*, 19:400–408, 1982.

[13] H.S. Dollar. Constraint-style preconditioners for regularized saddle-point systems. *SIAM J. Matrix Anal. Appl.*, 29:672–684, 2007.

[14] H.S. Dollar, N.I.M. Gould, W.H.A. Schilders, and A.J. Wathen. Implicit-factorization preconditioning and iterative solvers for regularized saddle-point systems. *SIAM J. Matrix Anal. Appl.*, 28:170–189, 2006.

[15] H.S. Dollar, N.I.M. Gould, W.H.A. Schilders, and A.J. Wathen. Using constraint preconditioners with regularized saddle-point problems. *Comput. Optim. Appl.*, 36:249–270, 2007.

[16] C. Durazzi and V. Ruggiero. Indefinitely preconditioned conjugate gradient method for large sparse equality and inequality constrained quadratic problems. *Numer. Linear Algebra Appl.*, 10:673–688, 2003.

[17] H.C. Elman, D.J. Silvester, and A.J. Wathen. *Finite Elements and Fast Iterative Solvers.* Oxford University Press, 2005.

[18] A. Forsgren. Inertia-controlling factorizations for optimization algorithms. *Appl. Numer. Math.*, 43:91–107, 2002.

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

[20] A. Forsgren, P.E. Gill, and J.R. Shinnerl. Stability of symmetric ill-conditioned systems arising in interior methods for constrained optimization. *SIAM J. Optim.*, 17:187–211, 1996.

[21] M.P. Friedlander and D. Orban. A primal-dual regularization interior-point method for convex quadratic programs. *Math. Prog. Comp.*, 4:71–107, 2012.

[22] J. Gondzio. Convergence analysis of an inexact feasible interior point method for convex quadratic programming. *SIAM J. Optim.*, 23:1510–1527, 2013.

[23] N.I.M. Gould, D. Orban, and Ph. L. Toint. Cuter (and sifdec), a contrained and unconstrained testing environment, revisited,. *ACM Trans. Math. Softw.*, 29:373–394, 2003.

[24] N.I.M. Gould and V. Simoncini. Spectral analysis of saddle point matrices with indefinite leading blocks. *SIAM J. Matrix Anal. Appl.*, 31:1152–1171, 2009.

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

[26] C. Greif and D. Schötzau. Preconditioners for saddle point linear systems with highly singular (1, 1) blocks. *Electron. Trans. Numer. Anal.*, 22:114–121, 2006.

[27] N. J. Higham. *Functions of matrices: Theory and Computation.* The Johns Hopkins University Press, 2008.

[28] C. Keller, N.I.M. Gould, and A.J. Wathen. Constraint preconditioning for indefinite linear

systems. *SIAM J. Matrix Anal. Appl.*, 21:1300–1317, 2000.

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

[30] I. Maros and C. Meszaros. A repository of convex quadratic programming problems. *Optim. Meth. and Softw.*, 11&12:671–681, 1999.

[31] B. Morini, V. Simoncini, and M. Tani. Spectral estimates for unreduced symmetric KKT systems arising from interior point methods. *Numer. Linear Algebra Appl.*, 2016.

[32] C. C. Paige and M. A. Saunders. Solution of sparse indefinite systems of linear equations. *SIAM J. Numer. Anal.*, 12:617–629, 1975.

[33] PDCO: Primal-Dual interior method for Convex Objectives. `http://stanford.edu/group/software/pdco`.

[34] T. Rees and C. Greif. A preconditioner for linear systems arising from interior point optimization methods. *SIAM J. Sci. Comput.*, 29:1992–2007, 2007.

[35] M. Rozložník and V. Simoncini. Krylov subspace methods for saddle point problem with indefinite preconditioning. *SIAM J. Matrix Anal. Appl.*, 24(2):368–391, 2002.

[36] Y. Saad and M. H. Schultz. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. and Stat. Comput.*, 7:856–869, 1986.

[37] 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. SIAM, Philadelphia, 1994.

[38] S.Q. Shen, T.Z. Huang, and J.S. Zhang. Augmentation block triangular preconditioners for regularized saddle point problems. *SIAM J. Matrix Anal. Appl.*, 33:721–741, 2012.

[39] S.J. Wright. Stability of linear equations solvers in interior-point methods. *SIAM J. Matrix Anal. Appl.*, 16:1287 – 1307, 1995.

[40] S.J. Wright. *Primal-dual Interior-point Methods*. SIAM, Philadelphia, USA, 1997.