A structured modified Newton approach for solving systems of nonlinear equations arising in interior-point methods for quadratic programming

David Ek* Anders Forsgren*

September 18, 2020

Abstract

The focus in this work is interior-point methods for quadratic optimization problems with linear inequality constraints where the system of nonlinear equations that arise are solved with Newton-like methods. In particular, the concern is the system of linear equations to be solved at each iteration. Newton systems give high quality solutions but there is an interest in designing modified Newton systems which are computationally less expensive but normally at the expense of lower quality solutions. We propose a structured modified Newton approach where the Jacobian of each Newton system is approximated by a Jacobian at a previous point plus a sequence of low-rank update matrices. The updates are such that the Jacobian approximation preserves its sparsity structure and in consequence may be viewed as a Jacobian evaluated at a different point.

We introduce the theoretical setting, motivate the approach by theoretical results applicable in the asymptotic region and give numerical results for a set of convex quadratic programs. In an attempt to improve performance we also motivate and construct two heuristics to be added to the method.

Keywords. interior-point methods, modified Newton methods, quasi-Newton methods, low-rank updates

1. Introduction

This work is intended for quadratic optimization problems with linear inequality constraints on the form

$$\begin{aligned} \text{minimize} \quad & \tfrac{1}{2}x^T H x + c^T x \\ \text{subject to} \quad & Ax \geq b, \end{aligned} \tag{IQP}$$

where $H \in \mathbb{R}^{n\times n}$, $A \in \mathbb{R}^{m\times n}$, $x \in \mathbb{R}^n$, $c \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$. We consider classical primal-dual interior-point methods which means solving or approximately solving a sequence of systems of nonlinear equations. Application of Newton's method on the first order optimality conditions of (IQP) gives an *unreduced* unsymmetric block 3-by-3 system of linear equations with dimension $n + 2m$ to be solved at each iteration. This system can be put on an equivalent form which contains a *reduced* symmetric block 2-by-2 system of dimension $n + m$ or a *condensed* system

*Optimization and Systems Theory, Department of Mathematics, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden (daviek@kth.se,andersf@kth.se).

of dimension n, see Section 3.3. Both these systems typically become increasingly ill-conditioned as the iterates converge whereas the unreduced system may stay well conditioned throughout, see [15, 19] for analysis of the spectral properties of systems arising in interior-point methods for convex quadratic problems with linear constraints on standard form. The sparsity structure of the unreduced system is maintained in the reduced system. However, the condensed system is typically dense if A contains dense rows. The most computationally expensive part of an interior-point iteration is the solution of the Newton systems that arise, see e.g. [8, 10] for details on the solution of the systems and [20] for a comparison of the solution of unreduced and reduced systems. We are particularly interested in the convergence towards a solution when the Jacobian in each Newton system is replaced by an approximation which ideally makes the system computationally less expensive to solve. We propose Jacobian approximations composed of a Jacobian at a previous point, whose factorization is assumed to be known, plus a sequence of low-rank update matrices. This may be interpreted as solving the sequence of systems of nonlinear equations with a modified Newton method. Quasi-Newton approaches have been studied by Gondzio and Sobral [14] where Broyden's rank-1 updates are performed on the Jacobian approximation. The block-sparsity pattern of the true Jacobian is maintained for feasible iterates however in general it is typically lost. In contrast we consider low-rank update matrices which capture the sparsity pattern of the true Jacobian for all iterates. In consequence the modified Newton approach may be interpreted in the framework of previous work on e.g. effects of finite-precision arithmetic, stability, convergence and solution techniques for interior-point methods [5, 8, 9, 15, 18–20, 25–28].

The updates and the theory are given for the unreduced Jacobian but we also discuss how analogous updates can be made on both the reduced and the condensed Jacobian. The modified Newton approach is also compatible with certain regularization strategies, see e.g. [1, 11, 23], although it is outside the scope of this initial study.

The work is meant to contribute to the theoretical and numerical understanding of modified Newton methods for nonlinear systems of equations arising in primal-dual interior-point methods. We envisage the use of the modified Newton approach as an accelerator to a Newton approach. E.g., when these can be run in parallel and the modified Newton approach may utilize factorizations from the Newton approach when it is appropriate. To give an indication of the potential we show numerical simulations on benchmark problems from the repository of convex quadratic programming problems by Maros and Mészáros [17].

The manuscript is organized as follows; Section 2 contains a brief background to primal-dual interior-point methods and an introduction to the theoretical framework; in Section 3 we propose a modified Newton approach and discuss how it relates to some previous work on interior-point methods; Section 4 contains a description of the implementation along with two heuristics and a re-factorization strategy; in Section 5 we give numerical results on convex quadratic optimization problems; finally in Section 6 we give some concluding remarks.

1.1. Notation

Throughout, $\rho(M)$ will be used to denote the spectral radius of a matrix M and $|\mathcal{S}|$ will be used to denote the cardinality of a set $\mathcal{S}$. The notion "$\cdot$" is defined as the component-wise operator, "$\succ 0$" as positive definite, and "$\wedge$" as the logical and. Quantities associated with Newton's method will throughout be labeled with "$\hat{}$". Vector subscript and superscript denote component index and iteration index respectively. The only exception is e_i which denotes the ith unit vector of appropriate dimension. All norms considered are of type 2-norm unless otherwise stated.

2. Background

The theoretical setting is analogous to the setting in a previous work of ours on bound-constrained nonlinear problems [7]. For completeness, we review the background adapted to quadratic problems with linear inequality constraints in this section. Our interest is focused on the situation as primal-dual interior-point methods converge to a local minimizer $x^* \in \mathbb{R}^n$ with its corresponding multipliers $\lambda^* \in \mathbb{R}^m$ and slack variables $s^* \in \mathbb{R}^m$. Specifically we assume that the iterates of the method converge to a vector $\left(x^{*T}, \lambda^{*T}, s^{*T}\right)^T \triangleq (x^*, \lambda^*, s^*)$ that satisfies

$$
\begin{aligned}
Hx^* + c - A^T\lambda^* &= 0, && \text{(stationarity)} && \text{(2.1a)}\\
Ax^* - s - b &= 0, && \text{(feasibility)} && \text{(2.1b)}\\
s^* &\geq 0, && \text{(non-negativity of slack variables)} && \text{(2.1c)}\\
\lambda^* &\geq 0, && \text{(non-negativity of multipliers)} && \text{(2.1d)}\\
s^* \cdot \lambda^* &= 0, && \text{(complementarity)} && \text{(2.1e)}\\
Z(x^*)^T H Z(x^*) &\succ 0, && && \text{(2.1f)}\\
s^* + \lambda^* &> 0, && \text{(strict complementarity)} && \text{(2.1g)}
\end{aligned}
$$

with $Z(x^*)$ denoting a matrix whose columns span the nullspace of the Jacobian corresponding to the active constraints with a strictly positive multiplier, λ^*. The first-order necessary optimality conditions for a local minimizer of (IQP) are given by (2.1a)-(2.1e). The first-order conditions together with (2.1f) constitute the second-order sufficient conditions (2.1a)-(2.1f) [16]. In the theoretical results, we also assume that (x^*, λ^*, s^*) satisfies (2.1g). To simplify the notation, let z denote the triplet (x, λ, s). For a given barrier parameter $\mu \in \mathbb{R}$, we are interested in the function $F_\mu : \mathbb{R}^{n+2m} \to \mathbb{R}^{n+2m}$ given by

$$
F_\mu(z) = \begin{bmatrix} Hx + c - A^T\lambda \\ Ax - s - b \\ \Lambda S e - \mu e \end{bmatrix}, \text{ with } z = (x, \lambda, s), \tag{2.2}
$$

where $S = \operatorname{diag}(s)$, $\Lambda = \operatorname{diag}(\lambda)$ and e is a vector of ones of appropriate size. The first-order necessary optimality conditions (2.1a)-(2.1e) are satisfied by a vector z with $s \geq 0$ and $\lambda \geq 0$ that fulfill $F_\mu(z) = 0$ for $\mu = 0$. In primal-dual interior point methods $F_\mu(z) = 0$ is solved or approximately solved for a decreasing sequence of

$\mu > 0$ while preserving $s > 0$ and $\lambda > 0$. Application of Newton's method means solving a sequence of systems of linear equations on the form

$$F'(z)\Delta\hat{z} = -F_\mu(z), \tag{2.3}$$

where $\Delta\hat{z} = (\Delta\hat{x}, \Delta\hat{\lambda}, \Delta\hat{s})$ and $F' : \mathbb{R}^{n+2m} \to \mathbb{R}^{(n+2m)\times(n+2m)}$ is the Jacobian of F_μ, defined by

$$F'(z) = \begin{bmatrix} H & -A^T & \\ A & & -I \\ & S & \Lambda \end{bmatrix}. \tag{2.4}$$

The subscript μ has been omitted since F' is independent of the barrier parameter. To improve efficiency many methods seek approximate solutions for each μ. The reduction of μ is typically determined by some specified measure of improvement. A natural measure in our framework is $\|F_\mu(z)\|$ as the exact solution satisfies $\|F_\mu(z)\| = 0$. In the theoretical framework we consider the basic condition $\|F_\mu(z)\| < C\mu$, for some constant $C > 0$, see, e.g., [21, Ch. 17, p. 572]. The additional assumption that all vectors z satisfy $s > 0$ and $\lambda > 0$ is made throughout.

In the remaining part of this section we give some definitions and provide the details for the theoretical framework.

Definition 2.1. (Order-notation) *Let $\alpha, \gamma \in \mathbb{R}$ be two positive related quantities. If there exists a constant $C_1 > 0$ such that $\gamma \geq C_1\alpha$ for sufficiently small α, then $\gamma = \Omega(\alpha)$. Similarly, if there exists a constant $C_1 > 0$ such that $\gamma \leq C_1\alpha$ for sufficiently small α, then $\gamma = \mathcal{O}(\alpha)$. If there exist constants $C_1, C_2 > 0$ such that $C_1\alpha \leq \gamma \leq C_2\alpha$ for sufficiently small α then, $\gamma = \Theta(\alpha)$.*

Definition 2.2. (Neighborhood) *For a given $\delta > 0$, let the neighborhood around z^* be defined by $\mathcal{B}(z^*, \delta) = \{z : \|z - z^*\| < \delta\}$.*

Assumption 2.1. (Strict local minimizer) *The vector z^* satisfies (2.1), i.e., second-order sufficient optimality conditions and strict complementarity hold.*

The first of the following two lemmas provides the existence of a neighborhood where the Jacobian is nonsingular. The second lemma gives the existence of a Lipschitz continuous barrier-trajectory z^μ in the neighborhood where the Jacobian is nonsingular. The results are well known and can be found in e.g. the work of Ortega and Rheinboldt [22]. See also Byrd, Liu and Nocedal [4] for the corresponding results in a setting similar to the one considered here.

Lemma 2.1. *Under Assumption 2.1 there exists $\delta > 0$ such that $F'(z)$ is continuous and nonsingular for $z \in \mathcal{B}(z^*, \delta)$ and*

$$\|F'(z)^{-1}\| \leq M,$$

for some constant $M > 0$.

Proof. See [22, p. 46]. ∎

Lemma 2.2. *Let Assumption 2.1 hold and let $\mathcal{B}(z^*, \delta)$ be defined by Lemma 2.1. Then there exists $\hat{\mu} > 0$ such that for each $0 < \mu \leq \hat{\mu}$ there is a Lipschitz continuous function $z^\mu \in \mathcal{B}(z^*, \delta)$ that satisfies $F_\mu(z^\mu) = 0$ and*

$$\|z^\mu - z^*\| \leq C_3 \mu,$$

where $C_3 = \inf_{z \in \mathcal{B}(z^*, \delta)} \|F'(z)^{-1} \frac{\partial F_\mu(z)}{\partial \mu}\|$.

Proof. The result follows from the implicit function theorem, see e.g. [22, p. 128]. ∎

The following lemma provides a relation between the distance of vectors z to the barrier-trajectory and the quantity $\|F_\mu(z)\|$, when the distance is sufficiently small. A corresponding result is also given by Byrd, Liu and Nocedal [4].

Lemma 2.3. *Under Assumption 2.1 let $\mathcal{B}(z^*, \delta)$ and $\hat{\mu}$ be defined by Lemma 2.1 and Lemma 2.2 respectively. For $0 < \mu \leq \hat{\mu}$ and z sufficiently close to $z^\mu \in \mathcal{B}(z^*, \delta)$ there exist constants $C_4, C_5 > 0$ such that*

$$C_4 \|z - z^\mu\| \leq \|F_\mu(z)\| \leq C_5 \|z - z^\mu\|.$$

Proof. See [4, p. 43]. ∎

The next lemma provides a bound on the Newton direction, $\Delta\hat{z}$, for z sufficiently close to the barrier trajectory.

Lemma 2.4. *Under Assumption 2.1 let $\mathcal{B}(z^*, \delta)$ and $\hat{\mu}$ be defined by Lemma 2.1 and Lemma 2.2 respectively. For $0 < \mu \leq \hat{\mu}$ and $z \in \mathcal{B}(z^*, \delta)$, let $\Delta\hat{z}$ be the solution of (2.3) with $\mu^+ = \sigma\mu$, where $0 < \sigma < 1$. If z is sufficiently close to $z^\mu \in \mathcal{B}(z^*, \delta)$ such that $\|F_\mu(z)\| = \mathcal{O}(\mu)$ then*

$$\|\Delta\hat{z}\| = \mathcal{O}(\mu).$$

Proof. Analogous to [7, Lemma 5]. ∎

3. A structured modified Newton approach

In order to describe the approach and its ideas in a simple setting we first consider one iteration ahead which we denote by "+". For a given $\mu > 0$, consider the interior-point iterate $z^+ \in \mathcal{B}(z^*, \delta)$ defined by $z^+ = z + \Delta\hat{z}$, where $z \in \mathcal{B}(z^*, \delta)$ and $\Delta\hat{z}$ satisfies (2.3) with $\mu^+ = \sigma\mu$, $0 < \sigma < 1$. Since $\Delta\hat{z}$ has been computed with (2.3) we assume that a factorization of $F'(z)$ is known. Instead of performing another Newton step $\Delta\hat{z}^+$ at z^+ for some $\mu^{++} = \sigma^+\mu^+$, $0 < \sigma^+ \leq 1$, which requires the solution of (2.3) with μ^{++}, we would like to compute an approximate solution, which is computationally less expensive, from

$$B^+\Delta z^+ = -F_{\mu^{++}}(z^+), \quad \text{where } B^+ = F'(z) + U, \tag{3.1}$$

and U is some low-rank update matrix. A natural question is then how to choose the update matrix U. Gondzio and Sobral [14] consider rank-1 update matrices such that B^+ is closest to B in Frobenious norm where B^+ satisfies the secant condition. They show that the block sparsity of the Jacobian is maintained for feasible iterates, however the sparsity of some individual blocks is typically lost. As we will see, the loss of sparsity in the individual blocks is a consequence of the secant condition requirement. In contrast we consider update matrices of rank r such that B^+ is closest to the real Jacobian $F'(z^+)$ in both Frobenious and 2-norm. In addition the sparsity of the true Jacobian is maintained, however there is no requirement for B^+ to fulfill the secant condition. To further motivate the choice of update matrix we also give some results in the asymptotic region of our setting.

Before we propose an update matrix U; note that the change in the Jacobian at $z + \Delta z$ is linear by the definition of F' in (2.4)

$$F'(z + \Delta z) = F'(z) + \Delta F'(\Delta z),$$

where

$$\Delta F'(\Delta z) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \Delta S & \Delta \Lambda \end{pmatrix}, \tag{3.2}$$

with $\Delta \Lambda = \operatorname{diag}(\Delta \lambda)$ and $\Delta S = \operatorname{diag}(\Delta s)$. The Jacobian approximation error with B^+ as in (3.1) is

$$E = F'(z^+) - B^+ = \Delta F'(\Delta \hat{z}) - U. \tag{3.3}$$

The Eckart-Young-Mirsky theorem then provides the best update matrix U of a given rank r in terms of the measure $\|.\|_2$ and $\|.\|_F$. In Proposition 3.1 below we give an expression for U and show that the resulting Jacobian approximation B^+ may be viewed as a Jacobian evaluated at a point $\bar{z}^+ = (\bar{x}^+, \bar{\lambda}^+, \bar{s}^+)$.

Proposition 3.1. *For $z = (x, \lambda, s)$ and $\Delta z = (\Delta x, \Delta \lambda, \Delta s)$, let $F'(z)$ and $\Delta F'(\Delta z)$ be defined by (2.4) and (3.2) respectively, and let $z^+ = z + \Delta z$. For a given rank r, $0 < r \leq m$, let $\mathcal{U}_r$ be a set that contains the indices of the r largest quantities of $\sqrt{(\Delta \lambda_i)^2 + (\Delta s_i)^2}$, $i = 1, \ldots, m$. The optimal solution of*

$$\begin{aligned} \underset{U \in \mathbb{R}^{(2m+n)\times(2m+n)}}{\text{minimize}} \quad & \|F'(z^+) - B^+\| \\ \text{subject to} \quad & B^+ = F'(z) + U, \\ & \operatorname{rank}(U) \leq r, \end{aligned}$$

where $\|.\|$ is either of type $\|.\|_2$ or $\|.\|_F$, is

$$U_* = \sum_{i \in \mathcal{U}_r} e_{n+m+i} \left((s_i^+ - s_i) e_{n+i} + (\lambda_i^+ - \lambda_i) e_{m+n+i} \right)^T.$$

In consequence, it holds that

$$B^+ = F'(\bar{z}^+), \text{ with } (\bar{x}_i^+, \bar{\lambda}_i^+, \bar{s}_i^+) = \begin{cases} (x_i^+, \lambda_i^+, s_i^+) & i \in \mathcal{U}_r, \\ (x_i^+, \lambda_i, s_i) & i \in \{1, \ldots, m\} \setminus \mathcal{U}_r. \end{cases}$$

Proof. Note that $\|F'(z^+) - B^+\| = \|\Delta F'(\Delta z) - U\|$ by (3.3). The result then follows from the Eckart-Young-Mirsky theorem, stated in Theorem A.1, together with Lemma A.1. The last part of the proposition follows directly from performing the update. ∎

Proposition 3.1 shows that each rank-1 term of the sum in U_* added to $F'(z)$ is equivalent to the update of a component-pair, (λ, s), in the Λ and S blocks of the Jacobian approximation. The essence is that adding the rank-r update matrix U_* to $F'(z)$ is equivalent to updating pairs (λ_i, s_i) to $\left(\lambda_i^+, s_i^+\right)$, $i \in \mathcal{U}_r$, and that the Jacobian approximation at z^+ may be viewed as a Jacobian evaluated at $\bar{z}^+$. In particular, $r = m$ gives $\bar{z}^+ = z^+$ and $B^+ = F'(z^+)$.

To further motivate the update matrix U_* of Proposition 3.1 we give a bound on the search direction error $\|\Delta z^+ - \Delta\hat{z}^+\|$ in the asymptotic region. Namely, where Δz^+ satisfies (3.1) with update matrix U_* of rank r as given in Proposition 3.1. In the derivation, the inverse of B^+ is expressed as a Neumann series which requires $\rho(F'(z^+)^{-1}E) < 1$. We first show that the update matrix U_* is sound in regards to the reduction of an upper bound of $\rho(F'(z^+)^{-1}E)$. Thereafter, in Lemma 3.1 we show that $\|F'(z^+)^{-1}E\|$, and in consequence also $\rho(F'(z^+)^{-1}E)$, is bounded above by a constant times μ in the asymptotic region. This gives the existence of a region where $\rho(F'(z^+)^{-1}E) < 1$.

By assumption $z^+ \in \mathcal{B}(z^*, \delta)$ and hence by Lemma 2.1 there exists a constant $M > 0$ such that $\|F'(z^+)^{-1}\| \leq M$, and

$$\rho(F'(z^+)^{-1}E) \leq \|F'(z^+)^{-1}E\| \leq M\sigma_{max}(\Delta F'(\Delta\hat{z}) - U). \tag{3.4}$$

Lemma A.1 shows that the singular values of $\Delta F'(\Delta\hat{z})$ are given by $\sqrt{(\Delta\hat{\lambda}_i)^2 + (\Delta\hat{s}_i)^2}$, $i = 1, \ldots, m$. The largest reduction of the upper bound in (3.4) is obtained with the rank-r update matrix U_* of Proposition 3.1 which gives

$$\begin{aligned}\rho(F'(z^+)^{-1}E) &\leq M \max_{i=r+1,\ldots,m} \sqrt{(\Delta\hat{\lambda}_i)^2 + (\Delta\hat{s}_i)^2} \\ &= M\sqrt{(\Delta\hat{\lambda}_{r+1})^2 + (\Delta\hat{s}_{r+1})^2},\end{aligned} \tag{3.5}$$

where the indices $i = 1, \ldots, m$ are ordered such that $\sqrt{(\Delta\hat{\lambda}_i)^2 + (\Delta\hat{s}_i)^2}$ are in descending order. Thus motivating the choice of update matrix in regards to the reduction of the upper bound of the spectral radius.

Lemma 3.1. *Under Assumption 2.1 let $\mathcal{B}(z^*, \delta)$ and $\hat{\mu}$ be defined by Lemma 2.1 and Lemma 2.2 respectively. For $0 < \mu \leq \hat{\mu}$ and $z \in \mathcal{B}(z^*, \delta)$, define $z^+ = z + \Delta\hat{z}$ where $\Delta\hat{z}$ is the solution of (2.3) with $\mu^+ = \sigma\mu$, where $0 < \sigma < 1$. Moreover, let U be an update matrix of rank r, $0 \leq r < m$, given by U_* of Proposition 3.1 and define $E = \Delta F'(\Delta\hat{z}) - U$. If z is sufficiently close to $z^\mu \in \mathcal{B}(z^*, \delta)$ such that $\|F_\mu(z)\| = \mathcal{O}(\mu)$ and $z^+ \in \mathcal{B}(z^*, \delta)$, then*

$$\|F'(z^+)^{-1}E\| \leq MC^{(r+1)}\mu, \tag{3.6}$$

where M is defined by Lemma 2.1 and $C^{(r+1)} > 0$ is a constant such that $\sqrt{(\Delta\hat{\lambda}_{r+1})^2 + (\Delta\hat{s}_{r+1})^2} \leq C^{(r+1)}\mu$ with $\sqrt{(\Delta\hat{\lambda}_i)^2 + (\Delta\hat{s}_i)^2}$, $i = 1, \ldots, m$, ordered in descending order.

Proof. Since $\Delta\hat{z}$ is the solution to (2.3) with $\mu^+ = \sigma\mu$, $0 < \sigma < 1$ it follows from Lemma 2.4 that there exists a constant $C > 0$ such that $\|\Delta\hat{z}\| \leq C\mu$. Hence there exist constants $C^{(i)} > 0$, $i = 1, \ldots, m$, such that

$$\sqrt{(\Delta\hat{\lambda}_i)^2 + (\Delta\hat{s}_i)^2} \leq C^{(i)}\mu, \quad i = 1, \ldots, m. \tag{3.7}$$

A combination of (3.5) and (3.7) gives the result. ∎

The bound in (3.6) of Lemma 3.1 shows that $\|F'(z^+)^{-1}E\|$, and by (3.4) also $\rho(F'(z)^{-1}E)$, will be less than unity for sufficiently small μ. This is also true when U is a zero matrix, i.e. for a simplified-Newton strategy. However the essence of the result is that U_* of Proposition 3.1 is the update matrix of rank r that gives the largest decrease in the upper bound, since $C^{(1)} \geq ... \geq C^{(m)}$, and hence increases the region. In particular (3.6) gives an explicit relation of how much small μ must be, depending on M (or $\|F'(z^+)^{-1}\|$) and $C^{(r+1)}$, for $\rho(F'(z^+)^{-1}E) < 1$. By (3.6) it holds that $\rho(F'(z^+)^{-1}E) < 1$ if $\mu < \frac{1}{\|F'(z^+)^{-1}\|C^{(r+1)}}$. Or equivalently a condition on $\|F'(z^+)^{-1}\|$ depending on μ that is $\|F'(z^+)^{-1}\| < \frac{1}{C^{(r+1)}\mu}$.

Next we give a bound on the search direction error at z^+ with the modified Newton equation (3.1) relative to the Newton equation (2.3) with μ^{++}. It is shown that the error in the asymptotic region is bounded by a constant times μ^3 when $\mu^{++} = \mu^+$ and a constant times μ^2 when $\mu^{++} < \mu^+$. As can be anticipated, the bound is tighter when μ is not decreased in the corresponding iteration.

Theorem 3.1. *Under Assumption 2.1 let $\mathcal{B}(z^*, \delta)$ and $\hat{\mu}$ be defined by Lemma 2.1 and Lemma 2.2 respectively. For $0 < \mu \leq \hat{\mu}$ let $z \in \mathcal{B}(z^*, \delta)$ be sufficiently close to $z^\mu \in \mathcal{B}(z^*, \delta)$ such that $\|F_\mu(z)\| = \mathcal{O}(\mu)$ and define $z^+ = z + \Delta\hat{z}$ where $\Delta\hat{z}$ is the solution of (2.3) with $\mu^+ = \sigma\mu$, $0 < \sigma < 1$. Moreover, let Δz^+ be defined by (3.1) with $\mu^{++} = \sigma^+\mu^+$, $0 < \sigma^+ \leq 1$, and update matrix U of rank r, $0 \leq r < m$, given by U_* of Proposition 3.1. If $z^+ \in \mathcal{B}(z^*, \delta)$ then there exist $\bar{\mu}$, with $\bar{\mu} \leq \hat{\mu}$, such that for $0 < \mu \leq \bar{\mu}$*

$$\left\|\Delta\hat{z}^+ - \Delta z^+\right\| = \begin{cases} \mathcal{O}(\mu^2) & 0 < \sigma^+ < 1, \\ \mathcal{O}(\mu^3) & \sigma^+ = 1, \end{cases} \tag{3.8}$$

where $\Delta\hat{z}^+$ is the Newton step at $\hat{z}^+$, given by $F'(z^+)\Delta\hat{z}^+ = -F_{\mu^{++}}(z^+)$.

Proof. The Jacobian approximation can by construction be written as

$$B^+ = F'(z^+)(I - F'(z^+)^{-1}E)$$

since $B^+ = F'(z) + U - \Delta F'(\Delta\hat{z}) + \Delta F'(\Delta\hat{z}) = F'(z^+) + U - \Delta F'(\Delta\hat{z}) = F'(z^+) - E = F'(z^+)(I - F'(z^+)^{-1}E)$. By Lemma 3.1 it holds that $\|F'(z^+)^{-1}E\| \leq MC^{(r+1)}\mu$ and

hence there exist $\bar{\mu}$, with $\bar{\mu} \leq \hat{\mu}$ such that for $0 < \mu \leq \bar{\mu}$ it holds that $MC^{(r+1)}\mu < 1$. In consequence, for $0 < \mu \leq \bar{\mu}$, $(B^+)^{-1}$ can be written with von Neumann series as

$$(B^+)^{-1} = \left(I - F'(z^+)^{-1}E\right)^{-1} F'(z^+)^{-1} = \sum_{j=0}^{\infty} \left(F'(z^+)^{-1}E\right)^j F'(z^+)^{-1}. \tag{3.9}$$

The error with respect to the Newton step $\Delta\hat{z}^+$ can with (3.9) be written as

$$\begin{aligned}\Delta\hat{z}^+ - \Delta z^+ &= \Delta\hat{z}^+ + (B^+)^{-1}F_{\mu^{++}}(z^+) = \left(I - \left(I - F'(z^+)^{-1}E\right)^{-1}\right)\Delta\hat{z}^+ \\ &= \left(I - \sum_{j=0}^{\infty}\left(F'(z^+)^{-1}E\right)^j\right)\Delta\hat{z}^+ = -\sum_{j=1}^{\infty}\left(F'(z^+)^{-1}E\right)^j\Delta\hat{z}^+. \end{aligned} \tag{3.10}$$

Taking 2-norm on both sides of (3.10) and making use of norm inequalities give

$$\left\|\Delta\hat{z}^+ - \Delta z^+\right\| \leq \sum_{j=1}^{\infty}\left\|F'(z^+)^{-1}E\right\|^j \left\|\Delta\hat{z}^+\right\|. \tag{3.11}$$

The sum in (3.11) is a geometric series which is convergent since $\|F'(z^+)^{-1}E\| < 1$ for $\mu \leq \bar{\mu}$ and hence

$$\left\|\Delta\hat{z}^+ - \Delta z^+\right\| \leq \frac{\left\|F'(z^+)^{-1}E\right\|}{1 - \|F'(z^+)^{-1}E\|}\left\|\Delta\hat{z}^+\right\|. \tag{3.12}$$

Recall that $\Delta\hat{z}^+$ is the solution of $F'(z^+)\Delta\hat{z}^+ = -F_{\mu^{++}}(z^+)$ where $z^+ \in \mathcal{B}(z^*, \delta)$ and hence

$$\begin{aligned}\|\Delta\hat{z}^+\| &= \|F'(z^+)^{-1}F_{\mu^{++}}(z^+)\| = \|F'(z^+)^{-1}\left(F_{\mu^{++}}(z^+) - F_{\mu^{++}}(z^{\mu^{++}})\right)\| \\ &\leq ML_{F'}\|z^+ - z^{\mu^{++}}\| = ML_{F'}\|z + \Delta\hat{z} - z^{\mu^{++}}\| \\ &= ML_{F'}\|z - F'(z)^{-1}F_{\mu^+}(z) - z^{\mu^{++}}\|. \end{aligned} \tag{3.13}$$

Addition and subtraction of z^{μ^+} plus addition of $F'(z)^{-1}F_{\mu^+}(z^{\mu^+}) = 0$ inside the last factor of (3.13) give

$$\begin{aligned}\|\Delta\hat{z}^+\| &\leq ML_{F'}\|F'(z)^{-1}(F_{\mu^+}(z^{\mu^+}) - F_{\mu^+}(z) - F'(z)(z^{\mu^+} - z)) + z^{\mu^+} - z^{\mu^{++}}\| \\ &\leq ML_{F'}\left(\|F'(z)^{-1}(F_{\mu^+}(z^{\mu^+}) - F_{\mu^+}(z) - F'(z)(z^{\mu^+} - z))\| + \|z^{\mu^+} - z^{\mu^{++}}\|\right) \\ &\leq ML_{F'}\left(\frac{ML_{F'}}{2}\|z - z^{\mu^+}\|^2 + \|z^{\mu^+} - z^{\mu^{++}}\|\right). \end{aligned} \tag{3.14}$$

Addition and subtraction of z^{μ} in the last factor of the first term of (3.14) yield

$$\begin{aligned}\|\Delta\hat{z}^+\| &\leq ML_{F'}\left(\frac{ML_{F'}}{2}\|z - z^{\mu} + z^{\mu} - z^{\mu^+}\|^2 + \|z^{\mu^+} - z^{\mu^{++}}\|\right) \\ &\leq \frac{M^2L_{F'}^2}{2}\left(\|z - z^{\mu}\|^2 + 2\|z - z^{\mu}\|\|z^{\mu} - z^{\mu^+}\| + \|z^{\mu} - z^{\mu^+}\|^2\right) \\ &\quad + ML_{F'}\|z^{\mu^+} - z^{\mu^{++}}\|. \end{aligned}$$

By assumption there exist a constant $C > 0$ such that $\|F_\mu(z)\| \leq C\mu$, from Lemma 2.3 and Lemma 2.2 it follows that

$$\|\Delta\hat{z}^+\| \leq \frac{M^2L_{F'}^2}{2}\left(\frac{C^2}{C_4^2} + 2\frac{C}{C_4}C_3(1-\sigma) + C_3^2(1-\sigma)^2\right)\mu^2 + ML_{F'}C_3(1-\sigma^+)\sigma\mu.$$

Thus there exist a constant $\hat{C} > 0$ such that

$$\left\|\Delta\hat{z}^+\right\| \leq \begin{cases} \hat{C}\mu & 0 < \sigma^+ < 1, \\ \hat{C}\mu^2 & \sigma^+ = 1. \end{cases} \tag{3.15}$$

Insertion of (3.15) into (3.12), making use of Lemma 3.1, gives

$$\left\|\Delta\hat{z}^+ - \Delta z^+\right\| \leq \begin{cases} \frac{MC^{(r+1)}}{1-MC^{(r+1)}\mu}\hat{C}\mu^2 & 0 < \sigma^+ < 1, \\ \frac{MC^{(r+1)}}{1-MC^{(r+1)}\mu}\hat{C}\mu^3 & \sigma^+ = 1, \end{cases}$$

i.e. $\|\Delta\hat{z}^+ - \Delta z^+\| = \mathcal{O}(\mu^3)$ or $\|\Delta\hat{z}^+ - \Delta z^+\| = \mathcal{O}(\mu^2)$, depending on if $\mu^{++} = \mu^+$ or not. ∎

Similarly as for Proposition 3.1, the result of Theorem 3.1 is also true for an update matrix U of rank 0 asymptotically. The essence is again that the largest decrease in the upper bound in the proof is obtained with the update matrix U_* of Proposition 3.1. Hence increases the region for when the bounds are valid, and in consequence when the proposed modified Newton approach is a viable alternative.

3.1. At a general iteration

In this subsection we give the analogous result of Proposition 3.1 at a general iteration k, $k \geq 1$, with a variable step size α^k. Consider the sequence $\{z^i\}_{i=0}^k$ generated by $z^{i+1} = z^i + \alpha^i\Delta z^i$, $i = 0, \ldots, k-1$, where α^i is the step size. Suppose that each Δz^i satisfies

$$B^i\Delta z^i = -F_{\mu^i}(z^i), \quad \text{with } B^i = \begin{cases} F'(z^0) & i = 0, \\ B^{i-1} + U^i & i = 1, \ldots, k-1, \end{cases} \tag{3.16}$$

for some update matrices U^i of rank r^i and $\mu^i > 0$. If at $k = 1$, the update matrix is chosen as the optimal solution of the optimization problem of Proposition 3.1 then $B^1 = F'(\bar{z}_1)$, for some $\bar{z}_1$. Inductively, at an iteration k, $k \geq 1$, for a given a given rank r^k, $0 < r^k \leq m$, we wish to choose U^k as the optimal solution to

$$\begin{array}{ll} \underset{U\in\mathbb{R}^{(2m+n)\times(2m+n)}}{\text{minimize}} & \|F'(z^k) - B^k\| \\ \text{subject to} & B^k = B^{k-1} + U, \quad B^{k-1} = F'(\bar{z}^{k-1}), \\ & \text{rank}\,(U) \leq r^k, \end{array} \tag{3.17}$$

where $\bar{z}_{k-1}$ is given and $\|.\|$ is either of type $\|.\|_2$ or $\|.\|_F$. The optimal solution of (3.17), the update of $\bar{z}^k$ from $\bar{z}^{k-1}$ and the resulting optimal B^k are shown in Proposition 3.2 This is analogous to the update from z^0 to $\bar{z}^1$ given in Proposition 3.1, so

that the essence is that low-rank update matrices, defined by the solution of (3.17), is equivalent to updating the information corresponding to the r^k largest quantities $\sqrt{(\lambda_i^k - \bar{\lambda}_i^{k-1})^2 + (s_i^k - \bar{s}_i^{k-1})^2}$. In essence, the r^k largest deviations from the Newton step is corrected. In particular, $r^k = m$ gives $B^k = F'(z^k)$.

Proposition 3.2. *For iteration k, $k \geq 1$, consider optimization problem (3.17) for a given $\bar{z}^{k-1}$ and rank r^k, $0 < r^k \leq m$. The optimal solution U^k is given by*

$$U_*^k = \sum_{i \in \mathcal{U}_{r^k}} e_{n+m+i} \left((s_i^k - \bar{s}_i^{k-1}) e_{n+i} + (\lambda_i^k - \bar{\lambda}_i^{k-1}) e_{m+n+i} \right)^T,$$

where $\mathcal{U}_{r^k}$ is a set that contains the indices of the r^k largest quantities of $\sqrt{(\lambda_i^k - \bar{\lambda}_i^{k-1})^2 + (s_i^k - \bar{s}_i^{k-1})^2}$, $i = 1, \ldots, m$. In consequence, it holds that

$$B^k = F'(\bar{z}^k), \text{ with } \bar{z}^k = (\bar{x}^k, \bar{\lambda}^k, \bar{s}^k) = \begin{cases} (x_i^k, \lambda_i^k, s_i^k) & i \in \mathcal{U}_{r^k}, \\ (x_i^k, \bar{\lambda}_i^{k-1}, \bar{s}_i^{k-1}) & i \in \{1, \ldots, m\} \setminus \mathcal{U}_{r^k}. \end{cases}$$

Proof. The proof is analogous to that of Proposition 3.1 with $z = \bar{z}^{k-1}$, $z^+ = z^k$ and $\Delta z = z^k - \bar{z}^{k-1}$. ∎

A direct consequence of the property $B^k = F'(\bar{z}^k)$ is that iterates that become primal-dual feasible, i.e., satisfies the first two block equations of (2.3), will remain so.

3.2. Convergence

In this section we give two notes on the effects on convergence in the context of inexact Newton methods.

Convergence towards the barrier-trajectory

In the framework of Dembo, Eisenstat and Steihaug [6] steps of inexact Newton methods may be viewed on the form

$$F'(z^k)\Delta z^k = -F_\mu(z^k) - q^k, \text{ where } \|q^k\|/\|F_\mu(z^k)\| \leq \eta^k. \tag{3.18}$$

The authors show that the sequence of iterates $z^k + \Delta z^k$ converges to z^* for any sufficiently good initial guess z^0 if $\eta^k < 1$ uniformly, and also that the convergence is linear. Given that the iterates converge they also show that the convergence is superlinear if and only if $\|q^k\| = o(\|F_\mu(z^k)\|)$, as $k \to \infty$.

The modified Newton approach can be put onto the form of (3.18) under the assumption that $(I - E^k F'(z^k)^{-1})$ is nonsingular, where $E^k = B^k - F'(z^k)$. A straightforward calculation shows that (3.16) at iteration k, with $\mu^k = \mu$, can be written as

$$F'(z^k)\Delta z^k = -F_\mu(z^k) + \left(I - (I - E^k F'(z^k)^{-1})^{-1} \right) F_\mu(z^k). \tag{3.19}$$

Identification of terms in (3.18) and (3.19) gives

$$q^k = \left(I - (I - E^k F'(z^k)^{-1})^{-1}\right) F_\mu(z^k). \tag{3.20}$$

If in addition $\|E^k F'(z^k)^{-1}\| < 1$ then $q^k = \sum_{j=1}^{\infty}(E^k F'(z^k)^{-1})^j F_\mu(z^k)$. Analogous arguments as those for (3.10) to (3.12) give

$$\|q^k\| \leq \frac{\|E^k F'(z^k)^{-1}\|}{1 - \|E^k F'(z^k)^{-1}\|}\|F_\mu(z^k)\|.$$

Local convergence towards the barrier trajectory then follows if $\frac{\|E^k F'(z^k)^{-1}\|}{1-\|E^k F'(z^k)^{-1}\|} < 1$ and the convergence is superlinear if in addition $\frac{\|E^k F'(z^k)^{-1}\|}{1-\|E^k F'(z^k)^{-1}\|} \to 0$ as $k \to \infty$. From the result of Lemma 3.1 it follows that there exist a region where superlinear convergence is not lost, provided that Jacobian approximation is re-factorized every other iteration. Indeed, this results also holds for simplified-Newton however again the essence is that the region for when this is true increases with the proposed update matrix.

Convergence towards an optimal solution

Interior-point methods may also be interpreted as an inexact Newton method with steps on the form

$$F'(z^k)\Delta z^k + F_0(z^k) - \mu^k \tilde{e} = r^k, \tag{3.21}$$

where $\tilde{e} = (0^T, 0^T, e^T)^T$ and $r^k \in \mathbb{R}^{n+2m}$ is a residual vector, see, e.g., Bellavia [3] and Armand, Benoit and Dussault [2]. The term r^k may thus be put in relation to how accurately the systems is solved. Note that (3.19) is equivalent to

$$F'(z^k)\Delta z^k + F_0(z^k) - \mu^k \tilde{e} = \left(I - (I - E^k F'(z^k)^{-1})^{-1}\right)\left(F_0(z^k) - \mu^k \tilde{e}\right). \tag{3.22}$$

Identification of terms in (3.21) and (3.22) gives $r^k = q^k$ with q^k as in (3.20). General conditions for convergence may thus be obtained from the work by Bellavia [3] and Armand, Benoit and Dussault [2].

As primal-dual feasibility is maintained with the modified Newton approach it may also be interpreted in the framework of Gondzio [13] for convergence analysis on convex quadratic problems.

3.3. Reduced systems, regularization and solvability

The ideas presented so far have been given on the unreduced unsymmetric block 3-by-3 system (2.3). An analogous update can be made to both the reduced symmetric 2-by-2 system and the condensed system discussed below. Under the assumption that Λ is nonsingular (2.3) can be reformulated as the *reduced* system

$$\begin{pmatrix} H & A^T \\ A & -\Lambda^{-1}S \end{pmatrix}\begin{pmatrix} \Delta\hat{x} \\ -\Delta\hat{\lambda} \end{pmatrix} = -\begin{pmatrix} Hx + c - A^T\lambda \\ Ax - b - \mu\Lambda^{-1}e \end{pmatrix}, \tag{3.23}$$

together with $\Delta s = -(s - \mu\Lambda^{-1}e) - \Lambda^{-1}S\Delta\lambda$. If in addition S is nonsingular then a Schur complement reduction of $\Lambda^{-1}S$ in (3.23) gives the *condensed* system

$$(H + A^T S^{-1}\Lambda A)\Delta\hat{x} = -(Hx + c - A^T\lambda) - A^T S^{-1}(\Lambda(Ax - b) - \mu e). \qquad (3.24)$$

As mentioned, the proposed rank-r update matrix of Proposition 3.2 is equivalent to updating r component pairs (λ, s). Although the terms are nonlinear in (3.23) and (3.24), the analogous update can be performed by subtraction of old information and addition of new, which does not affect the rank of the update. In order to have simple notation, we have chosen to formulate our problem on the form (IQP), with inequality constraints only. The analogous results hold for quadratic optimization problems on standard form, as considered in [1, 11, 14, 18, 19]. However the update will then be on the diagonal of the H-matrix of the symmetric block 2-by-2 indefinite system. The proposed modified Newton approach may also be interpreted in the framework of previous work on stability, effects of finite-precision arithmetic and spectral properties of the arising systems, e.g. [5, 8, 9, 15, 18–20, 25–28]. Note also that the approach does not affect the solvability of (2.3) for convex problems since the positive definiteness of the Schur complement of $-\Lambda^{-1}S$ in (3.23) is maintained for $\lambda \geq 0$ and $s \geq 0$. The inertia of the Jacobian approximation is thus the same as the inertia of the true Jacobian.

Moreover, the proposed approach is also compatible with regularized methods for linearly constrained quadratic optimization problems, e.g. [1, 11, 23], as long as the scaling of the regularization is not changed at iterations where the Jacobian approximation is updated by a low-rank matrix. The scaling of the regularization may be changed at a re-factorization step, e.g. on the form suggested in (4.7) of Section 4.

4. Implementation

All numerical simulations were performed in `matlab` on benchmark problems from the repository of convex quadratic programming problems by Maros and Mészáros [17]. In general these problems contain both linear equality and linear inequality constraints. However, in order not to complicate the description of the implementation with further technical details we choose to give the description for problems on the same form as in previous sections. Note however that some of the parameters will depend on quantities related to the format of benchmark problems.

4.1. Basic method

The modified Newton approach may be incorporated into existing interior-point solvers. However, in this initial study we are interested in understanding the fundamental behavior as primal-dual interior-point methods converge. In particular when the search direction is obtained from a modified Newton equation on the form (3.16) with a rank-r update matrix with structure as U_*^k of Proposition 3.2, relative to a Newton equation (2.3). In order not to risk combining effects of the proposed update procedure with effects from other features in more advanced methods we

choose to implement the modified Newton approach in a simple primal-dual interior point framework. The methods that we consider are defined by its particular search direction update in combination with Algorithm 4.1 below. In all numerical experiments the method `Newton` will refer to Algorithm 4.1 where the search direction at iteration k, $\Delta z^k = (\Delta x^k, \Delta\lambda^k, \Delta s^k)$, satisfies (2.3). The method `mN-r(U)` will refer to Algorithm 4.1 where the search direction satisfies (3.16) with an update matrix of rank $r^k = r$ given by U_*^k of Proposition 3.2, i.e. `r(U)` = `r(`r`)`. Although the rank of the update matrices can be varied between the iterations, this initial study is limited to update matrices of constant rank in order to keep the comparisons clean.

Algorithm 4.1 Simple interior-point method for convex (IQP).

$k \leftarrow 0$, $\mu^k \leftarrow$ inital μ,
$(x^k, \lambda^k, s^k) \leftarrow$ Point such that $\lambda^k > 0, s^k > 0$ and $\|F_{\mu^k/\sigma}(x^k, \lambda^k, s^k)\| < \mu^k/\sigma$.
While $\|F_0(z^k)\| > \epsilon$ **do**
  $(\Delta x^k, \Delta\lambda^k, \Delta s^k) \leftarrow$ search direction
  $(\alpha_P^k, \alpha_D^k) \leftarrow \Big(\min\{1, 0.98\alpha_P^{max,k}\}, \min\{1, 0.98\alpha_D^{max,k}\}\Big)$
  $(x^{k+1}, \lambda^{k+1}, s^{k+1}) \leftarrow (x^k + \alpha_P^k \Delta x^k, \lambda^k + \alpha_D^k \Delta\lambda^k, s^k + \alpha_P^k \Delta s^k)$
  If $\|F_{\mu^k}(x^{k+1}, \lambda^{k+1}, s^{k+1})\| < \mu^k$
    $\mu^{k+1} \leftarrow \sigma\mu^k$
  Else
    $\mu^{k+1} \leftarrow \mu^k$
  End
  $k \leftarrow k + 1$
End

In Algorithm 4.1 at iteration k, $\alpha_P^{max,k}$ and $\alpha_D^{max,k}$ are the maximum feasible step sizes for s^k along Δs^k and λ^k along $\Delta\lambda^k$ respectively.

4.2. Benchmark problems

The problems were pre-processed and put on an equivalent form with n x-variables, m_{in} inequality constraints and m_{eq} equality constraints. The total number of variables in the primal-dual formulation is thus $N = n + m_{eq} + 2m_{in}$ variables, see Appendix for a formulation and description of the systems that arise. A trivial equality constraint that fixed a variable at any of its bounds was removed from the problem along with the variable. A problem was accepted if $m_{in} \geq 4$ and in addition, if `Newton` converged from a given initial solution. Due to the simplicity of `Newton` convergence was not achieved for some problems due to reasons as, non-trivial equality constraints fixing variables at its bounds, singular Jacobians caused by linearly dependent equality constraints, etc. Moreover, we were not able to run `CONT-300`, `BOYD1` and `BOYD2` due to memory restrictions. These conditions reduced the benchmark set, $\mathcal{P}$, to 90 problems (out of 138). The problems were divided into the three subsets: small, $\mathcal{S}$, medium, $\mathcal{M}$, and large, $\mathcal{L}$. The sets were defined as follows: $\mathcal{S} = \{p \in \mathcal{P} : N < 500\}$, $\mathcal{M} = \{p \in \mathcal{P} : 500 \leq N < 10000\}$ and $\mathcal{L} = \{p \in \mathcal{P} : N \geq 10000\}$. In consequence $|\mathcal{S}| = 25$, $|\mathcal{M}| = 37$ and $|\mathcal{L}| = 28$. The

specific problems of each group and details on their individual sizes can be found in Appendix.

4.3. Heuristics

Initial numerical experiments with `mN-r(U)` have shown that convergence may slow down due to small step sizes α_P^k and α_D^k. Small step sizes can be caused by few components in the modified Newton direction direction which differ considerably from the Newton direction. We first show some numerical evidence of this behavior and suggest a partial explanation on which we base two heuristics. The effectiveness of the heuristics is then illustrated and finally a re-factorization strategy is included in the modified Newton approach. Step sizes and convergence, in terms of the measure $\|F_\mu\|$ with $\mu = 0$, for `Newton` and `mN-r(`r`)` with $r = [0,\ 2,\ 4]$ are shown in the left-hand side of Figure 1 . The results are for benchmark problem `qafiro` with parameters $\mu^0 = 1$, $\sigma = 0.1$ and $\epsilon = 10^{-6}$. The right-hand side of the figure shows the inverse of the limiting step sizes and the relative error in the search direction at the iteration marked by the red circle of `mN-r(2)`, hence large spikes imply small step sizes. Moreover, the figure only contains negative components of the modified Newton direction. The result for $r = 0$ is given to illustrate that low-rank updates can indeed make a difference compared to a simplified-Newton approach for which the theoretical results are still valid, although in a smaller region

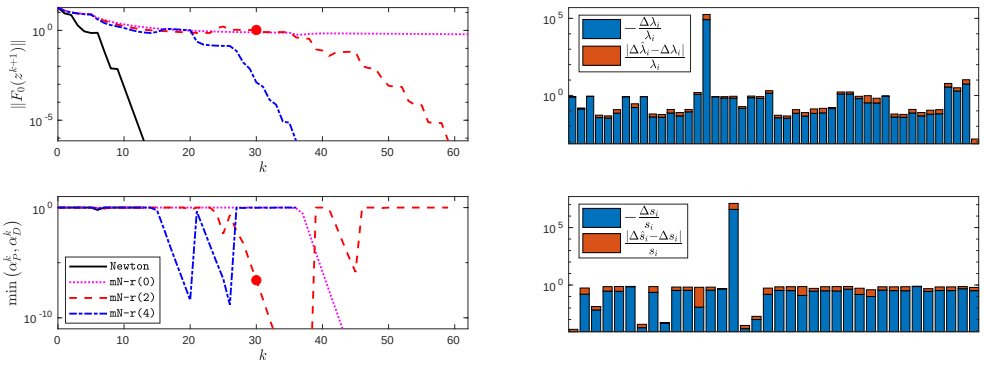

Figure 1: The left-hand side shows step sizes and convergence on benchmark problem `qafiro`. The right-hand side shows the inverse of the limiting step sizes and the relative error in the search direction for negative components of the modified Newton direction at the iteration marked by the red circle of `mN-r(2)`.

The results of Figure 1 indicate that convergence may slow down with the low-rank modified Newton approach due to small step sizes which are caused by large relative errors in certain components of the search direction. The results are similar to those shown by Gondzio and Sobral [14] for quasi-Newton approaches, hence indicating that the proposed modified Newton approach suffer from the same phenomenon as quasi-Newton approaches. In theory, zero steps are not harmful for the modified Newton approach, as long as Newton's method makes progress from

this point since after m/r iterations with zero steps the Jacobian approximation will indeed be the true Jacobian at that point. In practice however close to zero steps have negative effects on the convergence. In consequence we would like to understand what causes these steps and how to avoid them.

A potential reason for this behavior can be observed in an approximation of (2.3) for z sufficiently close to a solution z^* in our theoretical framework. The approximation is obtained by performing similar analysis as given by Wright [24] for a primal interior-point method on a nonlinear problem. This observation also provides some motivation to the choice of heuristics presented below.

Suppose that z is sufficiently close to z^*. Under the assumption that Λ and S are nonsingular the Newton system of (2.3) can be written on the condensed form of (3.24), which is equivalent to

$$\begin{aligned}(H + A_{\mathcal{A}}^T S_{\mathcal{A}}^{-1} \Lambda_{\mathcal{A}} A_{\mathcal{A}} + A_{\mathcal{I}}^T S_{\mathcal{I}}^{-1} \Lambda_{\mathcal{I}} A_{\mathcal{I}}) \Delta\hat{x} = &-(Hx + c - A_{\mathcal{A}}^T \lambda_{\mathcal{A}} - A_{\mathcal{I}}^T \lambda_{\mathcal{I}}) \\ &- A_{\mathcal{A}}^T S_{\mathcal{A}}^{-1} \Lambda_{\mathcal{A}} (Ax - b)_{\mathcal{A}} + \mu A_{\mathcal{A}}^T S_{\mathcal{A}}^{-1} e \\ &- A_{\mathcal{I}}^T S_{\mathcal{I}}^{-1} \Lambda_{\mathcal{I}} (Ax - b)_{\mathcal{I}} + \mu A_{\mathcal{I}}^T S_{\mathcal{I}}^{-1} e,\end{aligned} \tag{4.1}$$

with $\mathcal{A} = \{i \in \{1, \ldots, m\} : (Ax^* - b)_i = 0\}$, and $\mathcal{I} = \{i \in \{1, \ldots, m\} : (Ax^* - b)_i > 0\}$, i.e. $\mathcal{A}$ and $\mathcal{I}$ denotes the set of active and inactive constraints respectively. Moreover, if x is feasible then we may let $s = Ax - b$, in consequence (4.1) may be simplified to

$$\begin{aligned}(H + A_{\mathcal{A}}^T S_{\mathcal{A}}^{-1} \Lambda_{\mathcal{A}} A_{\mathcal{A}} + A_{\mathcal{I}}^T S_{\mathcal{I}}^{-1} \Lambda_{\mathcal{I}} A_{\mathcal{I}}) \Delta\hat{x} = &-(Hx + c) + \mu A_{\mathcal{A}}^T S_{\mathcal{A}}^{-1} e \\ &+ \mu A_{\mathcal{I}}^T S_{\mathcal{I}}^{-1} e.\end{aligned} \tag{4.2}$$

By first-order optimality conditions $Hx^* + c = A_{\mathcal{A}}^T \lambda_{\mathcal{A}}^*$ and hence $Hx + c = A_{\mathcal{A}}^T \lambda_{\mathcal{A}}^* + H(x - x^*)$ which gives

$$\begin{aligned}(H + A_{\mathcal{A}}^T S_{\mathcal{A}}^{-1} \Lambda_{\mathcal{A}} A_{\mathcal{A}} + A_{\mathcal{I}}^T S_{\mathcal{I}}^{-1} \Lambda_{\mathcal{I}} A_{\mathcal{I}}) \Delta\hat{x} = &-A_{\mathcal{A}}^T \lambda_{\mathcal{A}}^* - H(x - x^*) + \mu A_{\mathcal{A}}^T S_{\mathcal{A}}^{-1} e \\ &+ \mu A_{\mathcal{I}}^T S_{\mathcal{I}}^{-1} e,\end{aligned} \tag{4.3}$$

If in addition, $A_{\mathcal{A}} \Delta\hat{x}$ is sufficiently large, i.e, $\Delta\hat{x}$ is not in or almost in the null-space of $A_{\mathcal{A}}$ (if it is then the the search direction will not cause limiting steps). Since z is sufficiently close to z^* there exist a constant $\delta_x > 0$ such that $\|x - x^*\| < \delta_x$. In consequence, at sufficiently small μ, (4.4) may be approximated by

$$A_{\mathcal{A}}^T S_{\mathcal{A}}^{-1} \Lambda_{\mathcal{A}} A_{\mathcal{A}} \Delta\hat{x} = -A_{\mathcal{A}}^T \lambda_{\mathcal{A}}^* + \mu A_{\mathcal{A}}^T S_{\mathcal{A}}^{-1} e \tag{4.4}$$

By assumption $A_{\mathcal{A}}$ has full row rank and hence $S_{\mathcal{A}}^{-1} \Lambda_{\mathcal{A}} A_{\mathcal{A}} \Delta\hat{x} = -\lambda_{\mathcal{A}}^* + \mu S_{\mathcal{A}}^{-1} e$. Component-wise this means that

$$\frac{(\lambda_{\mathcal{A}})_i}{(s_{\mathcal{A}})_i} (A_{\mathcal{A}})_i^T \Delta\hat{x} = -(\lambda_{\mathcal{A}}^*)_i + \frac{\mu}{(s_{\mathcal{A}})_i}, \qquad i = 1, \ldots, |\mathcal{A}|, \tag{4.5}$$

or equivalently

$$\frac{(\lambda_{\mathcal{A}})_i}{(s_{\mathcal{A}})_i} \|(A_{\mathcal{A}})_i\| \|\Delta\hat{x}\| \cos\hat{\theta}_i = -(\lambda_{\mathcal{A}}^*)_i + \frac{\mu}{(s_{\mathcal{A}})_i}, \qquad i = 1, \ldots, |\mathcal{A}|, \tag{4.6}$$

where $\hat{\theta}_i$ is the angle between $\Delta\hat{x}$ and $(A_{\mathcal{A}})_i$, $i = 1, \ldots, |\mathcal{A}|$, which is the ith row of $A_{\mathcal{A}}$. Equation (4.5) and (4.6) show that if some of the coefficients in the blocks Λ and S of the Jacobian are inaccurate then this affects the quotient $\frac{(\lambda_{\mathcal{A}})_i}{(s_{\mathcal{A}})_i}$ and in consequence the inner product of the search direction and the rows of A corresponding to active constraints. Hence it affects the direction and/or the magnitude of the direction. Both of them can cause small step sizes by themselves. In the proposed modified Newton approach, depending on the rank of the update matrix, some of the factors in the Jacobian approximation, i.e. some components of $\frac{(\lambda_{\mathcal{A}})_i}{(s_{\mathcal{A}})_i}$, contain old information. As mentioned, each rank-1 term of the sum in U_* of Proposition 3.1 is equivalent to the update of information in the Jacobian approximation corresponding to one component-pair (λ, s). Although U_* gives the largest reduction in the upper bound of Theorem 3.1 it might not be good enough with respect to step sizes.

In light of the discussion above and the results of Figure 1 we construct two heuristics in an attempt to decrease negative effects on convergence cased by small step sizes. Both consist of a potential replacement of indices in the set $\mathcal{U}_r$ of Proposition 3.2, $\mathcal{U}_r$ since $r^k = r$, and will hence not change the rank of the update matrix between iterations. This was done to obtain a fair comparison in the study of the heuristics. The first heuristic can at most change two indices in $\mathcal{U}_r$ whereas the second is more flexible and can at most change $r^H \leq r$ of them.

Heuristic `H1`

The idea of the first heuristic is to ensure that information corresponding to component-pairs (λ, s) is updated if either limited the step size in the previous iteration. At iteration k, $k \geq 1$, the last one or two indices in $\mathcal{U}_r$ of Proposition 3.2 is replaced by

$$\hat{i}_1 = \underset{i:\Delta\lambda_i^{k-1}<0}{\operatorname{argmin}} \frac{\lambda_i^{k-1}}{-\Delta\lambda_i^{k-1}}, \quad \text{and} \quad \hat{i}_2 = \underset{i:\Delta s_i^{k-1}<0}{\operatorname{argmin}} \frac{s_i^{k-1}}{-\Delta s_i^{k-1}},$$

if $\min_{i:\Delta\lambda_i^{k-1}<0} \frac{\lambda_i^{k-1}}{-\Delta\lambda_i^{k-1}} < 1 \wedge \hat{i}_1 \notin \mathcal{U}_r$ and/or $\min_{i:\Delta s_i^{k-1}<0} \frac{s_i^{k-1}}{-\Delta s_i^{k-1}} < 1 \wedge \hat{i}_2 \notin \mathcal{U}_r$ respectively.

Heuristic `H2`

The principle of the second heuristic is directly based on the observation in the analysis above. Similarly as in `H1` the idea is to ensure that certain component-pairs (λ, s) is updated. In particular, the components with the largest relative error in the coefficients of the Jacobian approximation according to (4.5) and (4.6). However, the set of active constraints at the solution is unknown, instead all components which could have limited the step size in the previous iteration are considered in the selection. At iteration k, $k \geq 1$, at most r^H indices in $\mathcal{U}_r$ of Proposition 3.2 are replaced by the indices corresponding to the, at most, r^H largest quantities of

$$\frac{|s_i^k/\lambda_i^k - \bar{s}_i^k/\bar{\lambda}_i^k|}{s_i^k/\lambda_i^k}, \quad i \in \mathcal{H}^k$$

where $\mathcal{H}^k = \{i : \Delta\lambda_i^{k-1} < 0 \wedge \frac{\lambda_i^{k-1}}{-\Delta\lambda_i^{k-1}} < 1\} \cup \{i : \Delta s_i^{k-1} < 0 \wedge \frac{s_i^{k-1}}{-\Delta s_i^{k-1}} < 1\}$.

Heuristic test

To demonstrate the impact of heuristic `H1` and `H2` we show results in Figure 2 which are analogous to those in the left-hand side of Figure 1. The methods `mN-r(`r`)-H1` and `mN-r(`r`)-H2` denotes `mN-r(`r`)`, $r = [2, 4]$, combined with heuristic `H1` and `H2` respectively. In addition, Table 1 shows the average of the sum $(\alpha_P^k + \alpha_D^k)/2$ for a subset of the benchmark problems. The problems in the subset are sorted by increasing size and contains problems from each set $\mathcal{S}$, $\mathcal{M}$ and $\mathcal{L}$.

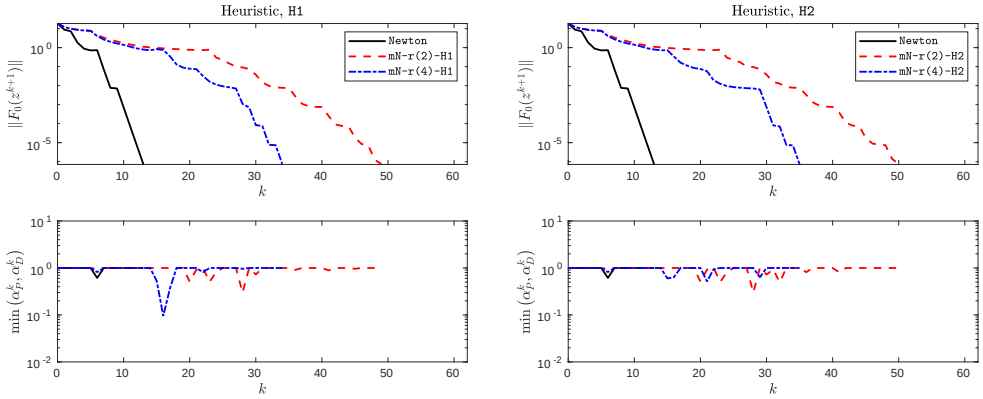

Figure 2: Step sizes and convergence for `mN-r(`r`)`, $r = [2, 4]$ combined with heuristic `H1` and `H2` on benchmark problem `qafiro`.

The results of Figure 2 show that `mN-r(`r`)-H1` and `mN-r(`r`)-H2`, $r = [2, 4]$, use larger step sizes, and converges in fewer iterations, compared to `mN-r(`r`)` in Figure 1. Hence showing that the heuristics `H1` and `H2` have the intended effect on benchmark problem `qafiro`.

Table 1: Average of the sum $(\alpha_P^k + \alpha_D^k)/2$ for a subset of the benchmark problems.

	Newton	mN-r(2)			mN-r(4)		
			H1	H2		H1	H2
hs118	0.964	0.370	**0.844**	0.814	0.513	0.817	**0.818**
qafiro	0.986	0.762	**0.973**	0.966	0.846	0.957	**0.971**
primal1	0.965	0.677	0.945	**0.956**	0.817	0.936	**0.957**
dualc8	1.000	0.584	0.909	**0.911**	0.845	0.875	**0.916**
laser	0.994	0.072	**0.883**	0.723	0.111	0.865	**0.890**
yao	0.958	0.683	**0.859**	0.444	0.732	**0.753**	0.462
stcqp2	0.999	0.705	**0.971**	0.916	0.699	0.957	**0.976**
ubh1	1.000	0.695	0.995	**0.998**	0.699	0.995	**0.999**
aug2dqp	0.996	0.832	0.952	**0.970**	0.589	**0.997**	0.994

The results in Table 1 indicate that `H1` and `H2` have the intended effect on more benchmark problems but also that they are not effective on all problems. Part of the reason is due to that the choice of methods is restricted to updates of rank 2 and 4 but for some problems there are many components which limit the step size. For instance `H1` can at most take two of these and `H2` can, at most, take

as many as the maximum rank of the update. The results for problem `yao` is an example where `H2` does worse than without heuristic. The heuristic replaced indices in $\mathcal{U}_r$ which caused low quality in the search direction; indicating that it might be beneficial to also update the information that is suggested in Proposition 3.2. Numerical experiments has further shown small step sizes can be avoided by allowing update matrices of varying rank. In particular where the rank is determined by the components which potentially limit the step size. However, numerical experiments have also shown that avoiding small step sizes is not sufficient to obtain increased convergence speed. Similarly as in the results of Gondizo and Sobral for quasi-Newton approaches [14], numerical experiments have shown that it is occasionally important to use the true Jacobian in order to improve convergence. In light of this, we will limit the allowed steps before the Jacobian is re-factorized and in consequence limit the total rank change of the Jacobian. In the following numerical simulations the Jacobian approximation update includes a re-factorization strategy on the form

$$B^k = \begin{cases} F'(z^k) & k = 0, l+1, 2l+2, 3l+3, \dots, \\ B^{k-1} + U_*^k & k \neq 0, l+1, 2l+2, 3l+3, \dots, \end{cases} \tag{4.7}$$

where U_*^k is defined in Proposition 3.2. Note that (4.7) in combination with `H1` or `H2` may affect the set $\mathcal{U}_r$ of Proposition 3.2. Other re-factorization strategies could be considered in practice such as re-factor if a particular step is considered bad instead of accepting all steps. Alternatively, other dynamical approaches where the rank of the update matrix is increased if a particular step is deemed bad for some reason.

5. Numerical results

In this section we give results on the form of number of iterations and factorizations for different versions of Algorithm 4.1. The results are meant to give an initial indication of the performance of the proposed modified Newton approach in a simple primal-dual interior-point framework. The results are for the methods `Newton`, `mN-r(`r`)`, `mN-r(`r`)-H1` and `mN-r(`r`)-H2`, with $r = [2, 16]$, described in Section 4. In essence the methods differ in how the search direction is computed. The direction at iteration k satisfies (2.3) in `Newton` and (3.16) in the `mN`-methods. In contrast to Section 4 here the `mN`-methods also include the re-factorization strategy described in (4.7). Due to the large variety in number of inequality constraints and problem sizes the parameter l of (4.7) were defined as the closest integer of $l_{\mathcal{S}}$, $l_{\mathcal{M}}$ and $l_{\mathcal{L}}$ for $p \in \mathcal{S}$, $p \in \mathcal{M}$ and $p \in \mathcal{L}$ respectively, see Table 2 for the specific values. The computational cost of a re-factorization of the unreduced, reduced and condensed system all depends on the sparsity structure given by the specific problem. We therefore choose $l_{\mathcal{S}}$, $l_{\mathcal{M}}$ and $l_{\mathcal{L}}$ such that they relate to the full rank change corresponding to a new factorization. The values of Table 2 were chosen such that a low-rank update is performed as long as the total rank change on the latest factorization is not larger than a factor of 1/2, 1/10 and 1/100 for the small, medium and large problems respectively. Moreover, the parameter σ of Algorithm 4.1 was set to $\sigma = 0.1$ and the tolerance in the stopping criterion, ϵ, was set to 10^{-6} for the small and medium sized problems and to 10^{-5} for the large sized problems.

Table 2: Re-factorization parameter for the different problem sizes.

$l_{\mathcal{S}}$	$l_{\mathcal{M}}$	$l_{\mathcal{L}}$
$rm_{in}/2$	$rm_{in}/10$	$rm_{in}/100$

Results are first shown for problems in the set $\mathcal{S}$, Table 3-4, thereafter for problems in $\mathcal{M}$, Table 5-7, and finally for problems in $\mathcal{L}$, Table 8-10. The results are for three different regions depending on μ^0, namely $\mu^0 = [1,\ 10^{-3},\ 10^{-6}]$. The intention is to illustrate the performance of the modified Newton approach both close to a solution as well as in a larger region where the theoretical results are not expected to hold. The results corresponding to $r = 16$ for problems in $\mathcal{S}$ are omitted due to similarity of the performance caused by the re-factorization strategy. In all tables the initial factorization of $B^0 = F'(z^0)$ is counted as one factorization. In essence, "1" in the factorization column, `F`, means that no re-factorization was performed. Moreover, "-" denotes that the method failed to converge within a maximum number of iterations.

Table 3: Number of factorizations and iterations for problems in $\mathcal{S}$ with $\mu^0 = 1$.

	Newton	mN-r(2)		mN-r(2)-H1		mN-r(2)-H2	
	F/It	F	It	F	It	F	It
cvxqp1_s	13	2	57	1	43	1	41
cvxqp2_s	14	2	54	2	54	2	54
cvxqp3_s	17	2	74	2	59	2	58
dual1	21	5	188	4	132	3	128
dual2	19	4	150	3	109	3	110
dual3	20	4	168	3	127	3	126
dual4	18	3	91	3	82	3	80
dualc1	10	1	23	1	23	1	23
dualc2	9	1	13	1	13	1	13
dualc5	11	1	15	1	15	1	15
qafiro	14	3	34	3	34	3	33
hs118	15	3	46	3	39	3	42
hs268	14	10	19	10	19	10	19
hs53	8	3	8	3	8	3	8
hs76	11	5	14	5	13	5	14
lotschd	11	4	14	4	15	4	14
primal1	24	5	97	4	83	4	85
primalc1	11	1	24	1	22	1	21
primalc2	10	1	17	1	17	1	17
qadlittl	11	2	39	2	37	2	37
qisrael	20	6	403	3	162	3	163
qpcblend	27	6	150	5	118	5	118
qscagr7	13	2	62	2	51	2	51
qshare2b	21	4	158	3	99	3	99
s268	14	10	19	10	19	10	19

Table 4: Number of factorizations and iterations for problems in $\mathcal{S}$ with $\mu^0 = 10^{-3}$ to the left and $\mu^0 = 10^{-6}$ to the right.

	Newton	mN-r(2)		mN-r(2)-H1		mN-r(2)-H2	
	F/It	F	It	F	It	F	It
cvxqp1_s	5	1	5	1	5	1	5
cvxqp2_s	5	1	5	1	5	1	5
cvxqp3_s	6	1	7	1	7	1	7
dual1	12	3	86	2	51	2	51
dual2	10	2	50	2	50	2	49
dual3	11	2	58	2	57	2	57
dual4	9	1	37	1	32	1	31
dualc1	5	1	5	1	5	1	5
dualc2	5	1	5	1	5	1	5
dualc5	5	1	5	1	5	1	5
qafiro	4	1	4	1	4	1	4
hs118	4	1	4	1	4	1	4
hs268	7	5	9	5	9	5	9
hs53	4	2	4	2	4	2	4
hs76	4	2	4	2	4	2	4
lotschd	4	2	4	2	4	2	4
primal1	12	3	45	2	37	2	37
primalc1	5	1	5	1	5	1	5
primalc2	5	1	5	1	5	1	5
qadlittl	5	1	5	1	5	1	5
qisrael	7	1	16	1	16	1	16
qpcblend	15	4	92	3	59	3	60
qscagr7	5	1	5	1	5	1	5
qshare2b	7	1	16	1	15	1	14
s268	7	5	9	5	9	5	9

	Newton	mN-r(2)		mN-r(2)-H1		mN-r(2)-H2	
	F/It	F	It	F	It	F	It
cvxqp1_s	3	1	3	1	3	1	3
cvxqp2_s	3	1	3	1	3	1	3
cvxqp3_s	3	1	3	1	3	1	3
dual1	6	2	43	1	19	1	21
dual2	5	1	8	1	8	1	8
dual3	5	1	13	1	11	1	11
dual4	5	1	6	1	6	1	6
dualc1	3	1	3	1	3	1	3
dualc2	3	1	3	1	3	1	3
dualc5	3	1	3	1	3	1	3
qafiro	2	1	2	1	2	1	2
hs118	2	1	2	1	2	1	2
hs268	3	2	3	2	3	2	3
hs53	2	1	2	1	2	1	2
hs76	2	1	2	1	2	1	2
lotschd	2	1	2	1	2	1	2
primal1	5	2	22	1	12	1	12
primalc1	3	1	3	1	3	1	3
primalc2	3	1	3	1	3	1	3
qadlittl	3	1	3	1	3	1	3
qisrael	3	1	3	1	3	1	3
qpcblend	9	2	45	2	32	2	31
qscagr7	3	1	3	1	3	1	3
qshare2b	3	1	3	1	3	1	3
s268	3	2	3	2	3	2	3

Table 5: Number of factorizations and iterations for problems in $\mathcal{M}$ with $\mu^0 = 1$.

	Newton	mN-r(2)		mN-r(2)-H1		mN-r(2)-H2		mN-r(16)		mN-r(16)-H1		mN-r(16)-H2	
	F/It	F	It	F	It	F	It	F	It	F	It	F	It
cvxqp1_m	15	2	140	2	112	2	111	3	29	3	29	3	29
cvxqp2_m	13	2	103	2	102	2	102	2	24	2	24	2	24
cvxqp3_m	16	3	204	2	127	2	121	3	37	3	33	3	34
dualc8	10	2	30	2	30	2	30	3	10	3	10	3	10
gouldqp2	15	4	213	3	168	3	194	4	32	4	31	4	30
gouldqp3	19	4	215	3	206	4	213	5	40	5	40	5	40
ksip	29	6	281	6	268	6	258	11	72	10	68	9	61
laser	18	4	305	2	163	2	178	4	46	4	43	4	42
primal2	23	8	42	8	42	8	42	23	23	23	23	23	23
primal3	24	9	52	8	44	8	44	24	24	24	24	24	24
primal4	21	8	36	8	37	8	37	21	21	21	21	21	21
primalc5	13	3	30	3	31	3	34	8	15	7	13	7	13
primalc8	10	1	17	1	17	1	17	3	10	3	10	3	10
q25fv47	24	10	910	3	215	11	950	13	151	6	67	9	98
qgrow15	16	3	134	3	129	3	129	5	37	5	34	5	33
qgrow22	20	5	402	3	192	3	192	6	69	5	57	5	57
qgrow7	16	4	91	3	65	3	67	6	23	6	21	6	21
qshell	10	1	33	1	25	1	25	2	15	2	15	2	15
qpcstair	25	7	163	5	109	5	109	11	42	10	39	10	39
qcapri	24	10	278	5	123	5	122	12	46	11	42	9	34
qsctap1	25	10	319	6	181	6	181	11	53	10	47	9	42
qsctap2	24	-	-	4	383	7	759	12	191	7	99	6	85
qsctap3	25	-	-	5	673	6	842	12	232	7	129	7	127
qsc205	23	5	74	5	73	5	73	11	31	11	31	11	31
qscagr25	12	2	64	2	39	2	39	4	19	4	17	4	18
qscsd1	21	6	206	4	124	4	134	9	43	9	41	8	37
qscsd6	26	8	487	4	257	8	478	10	87	9	73	8	71
qscsd8	22	6	701	5	553	4	442	6	102	7	108	5	79
qshare1b	20	7	85	4	48	6	68	14	26	13	24	14	26
values	26	6	108	5	103	5	99	13	36	12	35	12	35
aug3dcqp	26	3	434	3	424	3	430	5	103	5	103	5	102
aug3dqp	27	3	575	4	582	5	837	6	126	6	126	6	129
stadat1	13	2	323	1	191	2	303	2	60	2	49	2	49
stadat2	26	6	1533	4	1114	5	1244	7	231	6	191	6	191
mosarqp1	23	5	651	3	422	4	514	6	110	5	87	5	87
mosarqp2	22	4	229	3	164	3	166	5	46	5	46	5	44
yao	19	3	300	3	204	3	205	5	59	4	47	4	47

Table 6: Number of factorizations and iterations for problems in $\mathcal{M}$ with $\mu^0 = 10^{-3}$.

	Newton	mN-r(2)		mN-r(2)-H1		mN-r(2)-H2		mN-r(16)		mN-r(16)-H1		mN-r(16)-H2	
	F/It	F	It	F	It	F	It	F	It	F	It	F	It
cvxqp1_m	8	1	12	1	12	1	12	1	9	1	9	1	9
cvxqp2_m	6	1	6	1	6	1	6	1	6	1	6	1	6
cvxqp3_m	8	1	21	1	20	1	21	2	13	1	12	2	13
dualc8	6	1	6	1	6	1	6	2	6	2	6	2	6
gouldqp2	12	3	151	3	144	3	153	7	54	4	27	4	27
gouldqp3	14	3	145	2	130	2	131	4	28	4	28	4	28
ksip	20	4	185	4	174	5	207	8	53	7	44	7	44
laser	9	1	83	1	32	1	33	2	16	2	15	2	15
primal2	14	5	24	4	21	4	21	12	14	12	14	12	14
primal3	15	5	25	5	25	5	25	12	15	12	15	12	15
primal4	11	4	15	4	15	4	15	9	11	9	11	9	11
primalc5	6	1	6	1	6	1	6	4	6	4	6	4	6
primalc8	6	1	6	1	6	1	6	2	6	2	6	2	6
q25fv47	10	2	111	2	96	7	587	3	25	2	19	3	25
qgrow15	8	1	14	1	13	1	13	2	9	2	9	2	9
qgrow22	11	1	52	1	48	1	48	2	20	2	20	2	20
qgrow7	8	1	12	1	12	1	12	3	8	3	8	3	8
qshell	6	1	6	1	6	1	6	1	6	1	6	1	6
qpcstair	16	4	82	3	63	3	63	6	22	6	22	6	22
qcapri	9	2	37	2	32	2	32	3	10	3	10	3	10
qsctap1	9	2	35	2	34	2	34	3	12	3	12	3	12
qsctap2	8	1	47	1	17	1	17	2	17	2	16	2	16
qsctap3	9	1	43	1	21	1	22	2	22	2	21	2	21
qsc205	14	3	39	3	39	3	39	6	17	6	17	6	17
qscagr25	6	1	6	1	6	1	6	2	6	2	6	2	6
qscsd1	11	2	56	2	44	2	59	3	14	3	13	3	14
qscsd6	16	5	283	3	138	3	138	6	51	5	37	4	35
qscsd8	11	2	151	1	69	1	74	2	29	2	24	2	24
qshare1b	7	2	15	2	14	2	14	4	8	4	8	4	8
values	17	4	69	4	66	3	54	7	21	7	21	7	21
aug3dcqp	16	2	208	2	196	2	198	2	36	2	36	2	36
aug3dqp	16	2	234	2	238	2	255	3	63	3	63	3	63
stadat1	6	1	6	1	6	1	6	1	6	1	6	1	6
stadat2	20	4	1057	3	697	3	655	6	193	4	131	3	96
mosarqp1	14	3	328	2	186	2	193	3	47	3	45	3	45
mosarqp2	13	2	77	2	77	2	77	2	18	2	17	2	17
yao	9	2	107	1	25	1	25	2	19	2	16	2	16

Table 7: Number of factorizations and iterations for problems in $\mathcal{M}$ with $\mu^0 = 10^{-6}$.

	Newton	mN-r(2)		mN-r(2)-H1		mN-r(2)-H2		mN-r(16)		mN-r(16)-H1		mN-r(16)-H2	
	F/It	F	It	F	It	F	It	F	It	F	It	F	It
cvxqp1_m	3	1	3	1	3	1	3	1	3	1	3	1	3
cvxqp2_m	3	1	3	1	3	1	3	1	3	1	3	1	3
cvxqp3_m	3	1	3	1	3	1	3	1	3	1	3	1	3
dualc8	3	1	3	1	3	1	3	1	3	1	3	1	3
gouldqp2	8	2	80	2	73	2	82	3	19	2	17	3	19
gouldqp3	5	1	15	1	14	1	16	2	10	2	9	2	10
ksip	10	2	77	2	55	2	54	3	15	2	13	3	14
laser	3	1	3	1	3	1	3	1	3	1	3	1	3
primal2	4	2	6	2	6	2	6	3	4	3	4	3	4
primal3	5	2	7	2	8	2	8	4	5	4	5	4	5
primal4	3	1	4	1	4	1	4	2	3	2	3	2	3
primalc5	3	1	3	1	3	1	3	2	3	2	3	2	3
primalc8	3	1	3	1	3	1	3	1	3	1	3	1	3
q25fv47	3	1	3	1	3	1	3	1	3	1	3	1	3
qgrow15	3	1	3	1	3	1	3	1	3	1	3	1	3
qgrow22	4	1	8	1	8	1	8	1	8	1	8	1	8
qgrow7	3	1	3	1	3	1	3	1	3	1	3	1	3
qshell	3	1	3	1	3	1	3	1	3	1	3	1	3
qpcstair	7	2	27	2	27	2	27	3	9	3	9	3	9
qcapri	3	1	3	1	3	1	3	1	3	1	3	1	3
qsctap1	3	1	3	1	3	1	3	1	3	1	3	1	3
qsctap2	3	1	3	1	3	1	3	1	3	1	3	1	3
qsctap3	3	1	3	1	3	1	3	1	3	1	3	1	3
qsc205	6	1	11	1	11	1	11	2	6	2	6	2	6
qscagr25	3	1	3	1	3	1	3	1	3	1	3	1	3
qscsd1	3	1	3	1	3	1	3	1	3	1	3	1	3
qscsd6	5	1	10	1	8	1	9	2	9	1	8	2	9
qscsd8	4	1	4	1	4	1	4	1	4	1	4	1	4
qshare1b	3	1	3	1	3	1	3	2	3	2	3	2	3
values	8	2	24	2	23	2	23	3	9	3	9	3	9
aug3dcqp	7	1	18	1	18	1	18	1	9	1	9	1	9
aug3dqp	7	1	53	1	53	1	53	2	26	2	26	2	26
stadat1	3	1	3	1	3	1	3	1	3	1	3	1	3
stadat2	6	1	32	1	17	1	18	1	15	1	15	1	15
mosarqp1	5	1	8	1	9	1	8	1	7	1	7	1	7
mosarqp2	5	1	10	1	8	1	8	1	5	1	5	1	5
yao	3	1	3	1	3	1	3	1	3	1	3	1	3

Table 8: Number of factorizations and iterations for problems in $\mathcal{L}$ with $\mu^0 = 1$.

	Newton	mN-r(2)		mN-r(2)-H1		mN-r(2)-H2		mN-r(16)		mN-r(16)-H1		mN-r(16)-H2	
	F/It	F	It	F	It	F	It	F	It	F	It	F	It
aug2dcqp	20	3	205	2	163	2	162	4	46	4	46	4	46
aug2dqp	25	3	305	3	244	3	244	6	69	6	69	6	68
cont-050	17	4	89	4	85	4	85	8	28	7	24	7	25
cont-100	19	4	322	4	312	4	316	7	85	7	81	7	81
cont-101	20	5	439	4	312	4	312	7	79	5	62	6	68
cont-200	20	4	1291	3	815	3	819	5	208	4	158	4	160
cont-201	20	5	1625	4	1216	3	1205	5	208	4	166	4	168
stcqp2	18	3	85	3	85	3	87	5	27	5	27	5	27
stadat3	26	7	400	6	306	6	306	11	82	10	77	11	82
cvxqp1_l	15	2	104	2	104	2	104	3	28	3	29	3	28
cvxqp2_l	18	2	153	2	114	2	114	3	31	3	31	3	31
cvxqp3_l	18	2	141	2	129	2	131	4	49	4	49	4	49
exdata	25	10	344	7	252	7	241	11	51	11	51	11	51
hues-mod	25	3	132	3	132	3	132	7	44	7	44	7	44
huestis	21	2	78	2	78	2	78	4	26	4	26	4	26
liswet1	25	3	146	3	146	3	146	7	44	7	44	7	44
liswet2	33	6	256	5	219	5	220	11	72	9	61	11	72
liswet3	32	5	230	5	207	5	206	11	72	9	58	9	58
liswet4	33	7	306	5	233	6	255	11	76	9	59	9	59
liswet5	32	6	255	6	257	5	250	11	73	9	58	9	59
liswet6	32	5	234	5	205	5	205	11	72	9	58	9	58
liswet7	20	3	105	3	105	3	105	6	35	6	35	6	35
liswet8	33	9	408	10	479	8	359	15	102	13	84	12	79
liswet9	35	11	513	8	403	8	359	15	101	14	93	10	66
liswet10	33	8	362	6	278	8	402	13	89	14	96	13	84
liswet11	33	9	414	9	421	9	428	15	100	16	106	13	87
liswet12	30	8	361	7	319	9	415	15	100	11	74	10	68
ubh1	9	2	66	2	66	2	66	2	14	2	14	2	14

Table 9: Number of factorizations and iterations for problems in $\mathcal{L}$ with $\mu^0 = 10^{-3}$.

	Newton	mN-r(2)		mN-r(2)-H1		mN-r(2)-H2		mN-r(16)		mN-r(16)-H1		mN-r(16)-H2	
	F/It	F	It	F	It	F	It	F	It	F	It	F	It
aug2dcqp	11	1	65	1	66	1	65	2	19	2	19	2	19
aug2dqp	16	2	142	2	142	2	142	4	39	4	39	4	39
cont-050	13	3	61	3	58	3	58	6	20	5	17	5	17
cont-100	15	3	219	3	209	3	213	5	59	5	55	4	42
cont-101	16	4	336	3	209	3	210	5	63	4	39	3	37
cont-200	17	4	1219	3	815	3	815	4	157	3	107	3	108
cont-201	16	4	1220	3	811	3	811	4	157	4	154	3	115
stcqp2	9	1	29	1	28	1	29	2	11	2	11	2	11
stadat3	21	7	399	5	253	4	184	10	74	7	53	9	66
cvxqp1_l	7	1	12	1	12	1	12	1	9	1	9	1	9
cvxqp2_l	10	1	22	1	22	1	22	2	14	2	14	2	14
cvxqp3_l	10	1	22	1	22	1	22	2	17	2	17	2	17
exdata	14	5	154	3	84	3	84	5	21	5	21	5	22
hues-mod	16	2	60	2	60	2	60	4	21	4	21	4	21
huestis	12	1	40	1	40	1	40	2	13	2	13	2	13
liswet1	13	2	56	2	56	2	56	4	22	4	22	4	22
liswet2	21	5	205	3	118	4	154	8	51	6	37	6	37
liswet3	20	5	206	3	121	4	159	7	44	6	37	6	37
liswet4	21	5	210	5	210	6	268	8	55	6	38	6	38
liswet5	20	5	206	3	118	5	219	7	45	6	37	6	38
liswet6	20	4	189	4	170	5	205	7	44	6	37	6	37
liswet7	9	1	23	1	23	1	23	3	14	3	14	3	14
liswet8	21	7	336	6	271	6	255	9	58	10	63	8	49
liswet9	22	9	411	4	200	4	155	13	87	7	46	7	44
liswet10	21	8	360	5	210	5	213	11	70	13	85	7	46
liswet11	21	8	359	7	315	8	357	11	74	10	63	9	61
liswet12	18	9	413	4	176	4	169	9	61	6	40	5	33
ubh1	6	1	6	1	6	1	6	1	6	1	6	1	6

Table 10: Number of factorizations and iterations for problems in $\mathcal{L}$ with $\mu^0 = 10^{-6}$.

	Newton	mN-r(2)		mN-r(2)-H1		mN-r(2)-H2		mN-r(16)		mN-r(16)-H1		mN-r(16)-H2	
	F/It	F	It	F	It	F	It	F	It	F	It	F	It
aug2dcqp	5	1	5	1	5	1	5	1	5	1	5	1	5
aug2dqp	7	1	35	1	35	1	35	2	15	2	15	2	15
cont-050	3	1	3	1	3	1	3	1	3	1	3	1	3
cont-100	6	1	87	1	87	1	87	2	14	2	14	2	14
cont-101	6	1	10	1	10	1	9	1	10	1	10	1	9
cont-200	8	2	407	2	406	2	406	2	53	2	52	2	52
cont-201	7	2	407	2	406	2	406	2	53	2	52	1	44
stcqp2	3	1	3	1	3	1	3	1	3	1	3	1	3
stadat3	8	2	63	1	59	1	56	2	11	2	10	2	10
cvxqp1_l	3	1	3	1	3	1	3	1	3	1	3	1	3
cvxqp2_l	4	1	4	1	4	1	4	1	4	1	4	1	4
cvxqp3_l	5	1	8	1	8	1	8	1	4	1	4	1	4
exdata	4	1	5	1	5	1	5	2	5	2	5	2	5
hues-mod	7	1	12	1	12	1	12	2	8	2	8	2	8
huestis	5	1	6	1	6	1	6	1	5	1	5	1	5
liswet1	4	1	5	1	5	1	5	1	5	1	5	1	5
liswet2	9	2	52	2	52	2	52	3	15	3	15	3	15
liswet3	8	2	53	2	53	2	53	3	16	3	16	3	16
liswet4	9	2	56	2	55	2	67	3	19	3	17	3	17
liswet5	8	2	53	2	53	1	43	3	16	3	16	3	16
liswet6	8	2	52	2	52	2	52	3	15	3	15	3	15
liswet7	3	1	3	1	3	1	3	1	3	1	3	1	3
liswet8	11	3	125	2	91	2	96	4	27	5	28	5	28
liswet9	10	3	102	1	37	1	36	4	24	3	14	3	14
liswet10	11	4	158	2	57	2	57	5	31	3	19	3	19
liswet11	11	4	186	3	133	4	153	5	28	5	29	5	28
liswet12	6	2	53	1	19	1	19	2	9	2	9	2	8
ubh1	3	1	3	1	3	1	3	1	3	1	3	1	3

The results in Table 3-10 indicate that the number of factorizations compared to those done by `Newton` may be reduced by instead performing low-rank updates, as with `mN-r(`r`)`, $r = [2, 16]$. The reduced number of factorizations is however often at the expense of performing additional iterations with low-rank updates. The total number of iterations and/or factorizations are for many problems, but not for all, further reduced with heuristics `H1` and `H2`, as shown by the results corresponding to `mN-r(`r`)-H1` and `mN-r(`r`)-H2`, $r = [2, 16]$. This behavior is most significant in the simulations with larger values of μ as shown in Table 3, Table 5 and Table 8. Moreover, the `mN`-methods in Table 7, Table 10 and the right-hand side of Table 4 show similar performance. The low-rank updates are thus less likely to cause limiting steps in the asymptotic region on the benchmark problems. Overall `mN-r(2)` fails to converge for two problems within a maximum number of iterations due to small step sizes. This is overcome with both `H1` and `H2`, as shown by the corresponding results in Table 5.

Numerical experiments have further shown that decreasing the re-factorization parameters of Table 2 decreases the number of iterations but increases the number of factorizations done by the `mN`-methods. In general, an increased rank of the update matrix reduces the number of iterations overall but due to the re-factorization strategy the methods are required to re-factorize more often.

Table 7, Table 10 and the right-hand side of Table 4 show that low-rank updates are sufficient to achieve convergence at small μ in many of the benchmark problems, even for update matrices of rank-two on large scale problems.

6. Conclusion

In this work we have proposed and motivated a structured modified Newton approach for solving systems of nonlinear equations that arise in interior-point methods for quadratic optimization problems with linear inequality constraints. In essence the Jacobian of the Newton system is approximated by a previous Jacobian plus a sequence of low-rank update matrices. The approximation maintains the sparsity pattern of the true Jacobian and may thus be viewed as a Jacobian evaluated at a different point. The modified Newton approach may in consequence be interpreted in framework of previous work on primal-dual interior-point methods, e.g. effects of finite-precision arithmetic, stability, convergence and solution techniques.

Numerical simulations have shown that small step sizes can have negative effects on the convergence with the modified Newton approach. In an attempt to decrease these negative effects we have constructed and motivated two heuristics. Further numerical simulations have shown that the two heuristics often increase the step sizes but also that this is not always sufficient to improve convergence. We have therefore also suggested a re-factorization strategy. The heuristics and re-factorization strategy that we have proposed are merely options however the framework allows for both different versions of these as well as other heuristics and/or strategies.

In addition, we have performed numerical simulations on a set of convex quadratic benchmark problems. The results indicated that the number of factorizations compared to those done by Newton's method can be reduced, often at the expense of performing more iterations with low-rank updates. The total number of iterations and/or factorizations were for many problems, but not for all, further reduced with the two heuristics. Although the theoretical results are in the asymptotic region as $\mu \to 0$ we still obtain interesting numerical results for larger values of μ.

The results of this work are meant to contribute to the theoretical and numerical understanding of modified Newton approaches for solving systems of nonlinear equations arising in primal-dual interior-point methods. In particular as the Jacobian is approximated by a previous Jacobian plus a sequence of low-rank update matrices. We envisage that the work can lead to further research on similar types of modified Newton approaches in more sophisticated interior-point solvers as well as contribute to the development of preconditioners. The general ideas show potential but there is still a gap for the approach to be a viable alternative in practice. In particular the precise way of solving the updated modified Newton systems would have to be investigated further.

A. Appendix

For completeness we state the Eckart-Young-Mirsky theorem.

Theorem A.1. (Eckart-Young-Mirsky theorem) *Let $A \in \mathbb{R}^{m\times n}$, $m \leq n$, be of rank r and denote its singular value decomposition by $A = U\Sigma V^T$ where $U \in \mathbb{R}^{m\times m}$, $\Sigma \in \mathbb{R}^{m\times n}$, and $V \in \mathbb{R}^{n\times n}$. For a given q, $0 < q \leq r$, the optimal solution of*

$$\begin{aligned} \underset{\tilde{A}\in\mathbb{R}^{m\times n}}{\text{minimize}} \quad & \|A - \tilde{A}\| \\ \text{subject to} \quad & \text{rank}\,(\tilde{A}) \leq q, \end{aligned}$$

where $\|.\|$ is either 2-norm or Frobenius norm, is

$$A_* = \sum_{i=1}^{q} \sigma_i u_i v_i^T,$$

with σ_i, u_i and v_i as the ith diagonal element of Σ, ith column of U and ith column of V respectively.

Proof. See [12, Ch. 2]. ∎

The following lemma contains the singular value decomposition of $\Delta F'(\Delta z)$ given in (3.2).

Lemma A.1. *For $\Delta z = (\Delta x, \Delta\lambda, \Delta s) \in \mathbb{R}^{(n+2m)}$ let $\Delta F'(\Delta z) \in \mathbb{R}^{(n+2m)\times(n+2m)}$ be defined by (3.2). The singular value decomposition of $\Delta F'(\Delta z)$ can then be written as*

$$\Delta F'(\Delta z) = \sum_{i\in\mathcal{V}} e_{n+m+i} \left(\Delta s_i e_{n+i} + \Delta\lambda_i e_{m+n+i}\right)^T,$$

where $\mathcal{V}$ is a set of indices, $i = 1, \ldots, m$, ordered such that $\sqrt{(\Delta\lambda_i)^2 + (\Delta s_i)^2}$ are in descending order.

Proof. The left singular vectors are the set of orthonormal eigenvectors of $(\Delta F'(\Delta z))(\Delta F'(\Delta z))^T$, i.e. vectors u such that

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & (\Delta S)^2 + (\Delta\Lambda)^2 \end{pmatrix} u = \tilde{\lambda} u. \tag{A.1}$$

The eigenvectors of $(\Delta F'(\Delta z))(\Delta F'(\Delta z))^T$ are e_i, $i = 1, \ldots, n+2m$ and the eigenpairs, with nonzero eigenvalues, are $((\Delta\lambda_i)^2 + (\Delta s_i)^2, e_{n+m+i})$, $i = 1, \ldots, m$.

Similarly, the right singular vectors are the set of orthonormal eigenvectors of $(\Delta F'(\Delta z))^T(\Delta F'(\Delta z))$, i.e. vectors v such that

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & (\Delta S)^2 & \Delta S\Delta\Lambda \\ 0 & \Delta\Lambda\Delta S & (\Delta\Lambda)^2 \end{pmatrix} v = \tilde{\lambda} v. \tag{A.2}$$

The nonzero eigenvalues of (A.2) are the same as those in (A.1). A straightforward calculation shows that the ith eigenvector $v_i = \frac{1}{\sqrt{(\Delta\lambda_i)^2+(\Delta s_i)^2}} (\Delta s_i e_{n+i} + \Delta\lambda_i e_{m+n+i})$, $i = 1, \ldots, m$, fulfill (A.2) with $\tilde{\lambda}_i = (\Delta\lambda_i)^2 + (\Delta s_i)^2$, $i = 1, \ldots, m$, and in addition that the set of vectors v_i, $i = 1, \ldots, m$, form an orthornormal set. ∎

Pre-processing of benchmark problems

As the benchmark problems in general also contain equality constraints the problems were pre-processed and put on the form

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n}{\text{minimize}} & \frac{1}{2}x^THx + c^Tx \\ \text{subject to} & A_{eq}x = b_{eq}, \\ & A_{in}x \geq b_{in}, \end{array} \tag{A.3}$$

where $H \in \mathbb{R}^{n\times n}$, $c \in \mathbb{R}^n$, $A_{eq} \in \mathbb{R}^{m_{eq}\times n}$, $b_{eq} \in \mathbb{R}^{m_{eq}}$, $A_{in} \in \mathbb{R}^{m_{in}\times n}$, and $b_{in} \in \mathbb{R}^{m_{in}}$. The first order optimality conditions of (A.3) can be stated as i) $Hx + c - A_{eq}^T\lambda_{eq} - A_{in}^T\lambda_{in} = 0$, ii) $A_{eq}x = b_{eq}$, iii) $A_{in}x - s = b_{in}$. iv) $s \cdot \lambda_{in} = 0$, v) $s \geq 0$, vi) $\lambda_{in} \geq 0$, for vectors $\lambda_{eq} \in \mathcal{R}^{m_{eq}}$, $\lambda_{in} \in \mathcal{R}^{m_{in}}$ and $s \in \mathcal{R}^{m_{in}}$. Similarly as in Section 2, define $F_\mu : \mathbb{R}^{n+m_{eq}+2m_{in}} \to \mathbb{R}^{n+m_{eq}+2m_{in}}$ by

$$F_\mu(z) = \begin{bmatrix} Hx + c - A_{eq}^T\lambda_{eq} - A_{in}^T\lambda_{in} \\ A_{eq}x - b_{eq} \\ A_{in}x - s - b_{in} \\ \Lambda_{in}Se - \mu e \end{bmatrix}, \text{ with } z = (x, \lambda_{eq}, \lambda_{in}, s), \tag{A.4}$$

Primal-dual interior point methods involves solving or approximately solving $F_\mu(z) = 0$ for a decreasing sequence of $\mu >$ while maintaining $\lambda_{in} > 0$ and $s > 0$. Application of Newton's method gives systems on the form (2.3) with $F_\mu(z)$ as in (A.4), $\Delta\hat{z} = (\Delta\hat{x}, \Delta\hat{\lambda}_{eq}, \Delta\hat{\lambda}_{in}, \Delta\hat{s})$ and $F' : \mathbb{R}^{n+m_{eq}+2m_{in}} \to \mathbb{R}^{(n+m_{eq}+2m_{in})\times(n+m_{eq}+2m_{in})}$ defined by

$$F'(z) = \begin{bmatrix} H & -A_{eq}^T & -A_{in}^T & \\ A_{eq} & & & \\ A_{in} & & & -I \\ & & S & \Lambda_{in} \end{bmatrix}. \tag{A.5}$$

Problem data

Number of x-variables, equality constraints, inequality constraints and total number of variables in the primal-dual formulation for problems in the sets $\mathcal{S}$, $\mathcal{M}$ and $\mathcal{L}$ are shown in Table 11 and Table 12 respectively.

Table 11: Details on problem size for problems $p \in \mathcal{S}$.

	n	m_{eq}	m_{in}	N
cvxqp1_s	100	50	200	350
cvxqp2_s	100	25	200	325
cvxqp3_s	100	75	200	375
dual1	85	1	170	256
dual2	96	1	192	289
dual3	111	1	222	334
dual4	75	1	150	226
dualc1	9	1	232	242
dualc2	7	1	242	250
dualc5	8	1	293	302
qafiro	32	8	51	91
hs118	15	0	59	74
hs268	5	0	5	10
hs53	5	3	10	18
hs76	4	0	7	11
lotschd	12	7	12	31
primal1	325	0	86	411
primalc1	230	0	224	454
primalc2	231	0	236	467
qadlittl	96	14	137	247
qisrael	142	0	316	458
qpcblend	83	43	114	240
qscagr7	140	84	185	409
qshare2b	79	13	162	254
s268	5	0	5	10

Table 12: Details on problem size for problems $p \in \mathcal{M}$ and $p \in \mathcal{L}$.

	n	m_{eq}	m_{in}	N
cvxqp1_m	1000	500	2000	3500
cvxqp2_m	1000	250	2000	3250
cvxqp3_m	1000	750	2000	3750
dualc8	8	1	518	527
gouldqp2	699	349	1398	2446
gouldqp3	699	349	1398	2446
ksip	20	0	1001	1021
laser	1002	0	2000	3002
primal2	649	0	97	746
primal3	745	0	112	857
primal4	1489	0	76	1565
primalc5	287	0	286	573
primalc8	520	0	511	1031
q25fv47	1571	515	1876	3962
qgrow15	645	300	1245	2190
qgrow22	946	440	1826	3212
qgrow7	301	140	581	1022
qshell	1525	534	1644	3703
qpcstair	385	209	532	1126
qcapri	337	142	583	1062
qsctap1	480	120	660	1260
qsctap2	1880	470	2500	4850
qsctap3	2480	620	3340	6440
qsc205	202	90	315	607
qscagr25	500	300	671	1471
qscsd1	760	77	760	1597
qscsd6	1350	147	1350	2847
qscsd8	2750	397	2750	5897
qshare1b	225	89	253	567
values	202	1	404	607
aug3dcqp	3873	1000	3873	8746
aug3dqp	3873	1000	3873	8746
stadat1	2001	0	5999	8000
stadat2	2001	0	5999	8000
mosarqp1	2500	0	3200	5700
mosarqp2	900	0	1500	2400
yao	2000	0	2001	4001

	n	m_{eq}	m_{in}	N
aug2dcqp	20200	10000	20200	50400
aug2dqp	20200	10000	20200	50400
cont-050	2597	2401	5194	10192
cont-100	10197	9801	20394	40392
cont-101	10197	10098	20394	40689
cont-200	40397	39601	80794	160792
cont-201	40397	40198	80794	161389
stcqp2	4097	2052	8194	14343
stadat3	4001	0	11999	16000
cvxqp1_l	10000	5000	20000	35000
cvxqp2_l	10000	2500	20000	32500
cvxqp3_l	10000	7500	20000	37500
exdata	3000	1	7500	10501
hues-mod	10000	2	10000	20002
huestis	10000	2	10000	20002
liswet1	10002	0	10000	20002
liswet2	10002	0	10000	20002
liswet3	10002	0	10000	20002
liswet4	10002	0	10000	20002
liswet5	10002	0	10000	20002
liswet6	10002	0	10000	20002
liswet7	10002	0	10000	20002
liswet8	10002	0	10000	20002
liswet9	10002	0	10000	20002
liswet10	10002	0	10000	20002
liswet11	10002	0	10000	20002
liswet12	10002	0	10000	20002
ubh1	17997	12000	12006	42003

References

[1] A. Altman and J. Gondzio. Regularized symmetric indefinite systems in interior point methods for linear and quadratic optimization. *Optim. Methods Softw.*, 11/12(1-4):275–302, 1999.

[2] P. Armand, J. Benoist, and J.-P. Dussault. Local path-following property of inexact interior methods in nonlinear programming. *Comput. Optim. Appl.*, 52(1):209–238, 2012.

[3] S. Bellavia. Inexact interior-point method. *J. Optim. Theory Appl.*, 96(1):109–121, 1998.

[4] R. H. Byrd, G. Liu, and J. Nocedal. On the local behavior of an interior point method for nonlinear programming. In *Numerical Analysis 1997*, pages 37–56. Addison Wesley Longman, 1998.

[5] M. D'Apuzzo, V. De Simone, and D. di Serafino. On mutual impact of numerical linear algebra and large-scale optimization with focus on interior point methods. *Comput. Optim. Appl.*, 45(2):283–310, 2010.

[6] R. S. Dembo, S. C. Eisenstat, and T. Steihaug. Inexact Newton methods. *SIAM J. Numer. Anal.*, 19(2):400–408, 1982.

[7] D. Ek and A. Forsgren. Approximate solution of system of equations arising in interior-point methods for bound-constrained optimization. *ArXiv e-prints*, (arXiv:2004.04057), 2020.

[8] A. Forsgren, P. E. Gill, and J. D. Griffin. Iterative solution of augmented systems arising in interior methods. *SIAM J. Optim.*, 18(2):666–690 (electronic), 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 J. Matrix Anal. Appl.*, 17(1):187–211, 1996.

[10] A. Forsgren, P. E. Gill, and M. H. Wright. Interior methods for nonlinear optimization. *SIAM Rev.*, 44(4):525–597 (2003), 2002.

[11] M. P. Friedlander and D. Orban. A primal-dual regularized interior-point method for convex quadratic programs. *Math. Program. Comput.*, 4(1):71–107, 2012.

[12] G. H. Golub and C. F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, fourth edition, 2013.

[13] J. Gondzio. Convergence analysis of an inexact feasible interior point method for convex quadratic programming. *SIAM J. Optim.*, 23(3):1510–1527, 2013.

[14] J. Gondzio and F. N. C. Sobral. Quasi-Newton approaches to interior point methods for quadratic problems. *Comput. Optim. Appl.*, 74(1):93–120, 2019.

[15] C. Greif, E. Moulding, and D. Orban. Bounds on eigenvalues of matrices arising from interior-point methods. *SIAM J. Optim.*, 24(1):49–83, 2014.

[16] I. Griva, S. Nash, and A. Sofer. *Linear and Nonlinear Optimization: Second Edition*. Society for Industrial and Applied Mathematics, 01 2009.

[17] I. Maros and C. Mészáros. A repository of convex quadratic programming problems. *Optim. Methods Softw.*, 11/12(1-4):671–681, 1999.

[18] B. Morini and V. Simoncini. Stability and accuracy of inexact interior point methods for convex quadratic programming. *J. Optim. Theory Appl.*, 175(2):450–477, 2017.

[19] B. Morini, V. Simoncini, and M. Tani. Spectral estimates for unreduced symmetric KKT systems arising from interior point methods. *Numer. Linear Algebra Appl.*, 23(5):776–800, 2016.

[20] B. Morini, V. Simoncini, and M. Tani. A comparison of reduced and unreduced kkt systems arising from interior point methods. *Comput. Optim. Appl.*, 68:1–27, 2017.

[21] J. Nocedal and S. J. Wright. *Numerical optimization*. Springer Series in Operations Research and Financial Engineering. Springer, New York, second edition, 2006.

[22] J. M. Ortega and W. C. Rheinboldt. *Iterative Solution of Nonlinear Equations in Several Variables*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2000.

[23] M. Saunders and J. Tomlin. Solving regularized linear programs using barrier methods and KKT systems. SOL Report 96-4, Systems Optimization Laboratory, Dept. of Operations Research, Stanford University, Stanford, CA94306, USA, 1996.

[24] M. H. Wright. Why a pure primal Newton barrier step may be infeasible. *SIAM J. Optim.*, 5(1):1–12, 1995.

[25] M. H. Wright. Ill-conditioning and computational error in interior methods for nonlinear programming. *SIAM J. Optim.*, 9(1):84–111, 1999.

[26] S. Wright. Stability of augmented system factorizations in interior-point methods. *SIAM J. Matrix Anal. Appl.*, 18(1):191–222, 1997.

[27] S. J. Wright. Stability of linear equations solvers in interior-point methods. *SIAM J. Matrix Anal. Appl.*, 16(4):1287–1307, 1995.

[28] S. J. Wright. Effects of finite-precision arithmetic on interior-point methods for nonlinear programming. *SIAM J. Optim.*, 12(1):36–78, 2001.