

Homogeneous Formulation of Convex Quadratic Programs for Infeasibility Detection

Raghunathan, Arvind

TR2021-150 January 11, 2022

Abstract

Convex Quadratic Programs (QPs) have come to play a central role in the computation of control action for constrained dynamical systems. Convex QPs arise in Model Predictive Control (MPC) for linear systems, switched linear systems and in sequential linearization of nonlinear systems. A number algorithms have been developed in recent years for the solution of such QPs. However not all algorithms are capable of computing an optimal solution if feasible or producing a certificate of infeasibility. In this paper, we present a novel Homogeneous QP (HQP) formulation which is obtained by embedding the original QP in a larger space. The key properties of the HQP are: (i) is always feasible, (ii) an optimal solution to QP can be readily obtained from a solution to HQP, and (iii) infeasibility of QP corresponds to a particular solution of HQP. An immediate consequence is that all the existing algorithms for QP are now also capable of robustly detecting infeasibility. In particular, we present an Infeasible Interior Point Method (IIPM) for the HQP and show polynomial iteration complexity when applied to HQP. A key distinction with prior IPM approaches is that we do not need to solve second-order cone programs. Numerical experiments on the formulation are provided using existing codes.

IEEE Conference on Decision and Control (CDC) 2021

Homogeneous Formulation of Convex Quadratic Programs for Infeasibility Detection

Arvind U Raghunathan

Abstract—Convex Quadratic Programs (QPs) have come to play a central role in the computation of control action for constrained dynamical systems. In this paper, we present a novel Homogeneous QP (HQP) formulation which is obtained by embedding the original QP in a larger space. The key properties of the HQP are: (i) is always feasible, (ii) an optimal solution to QP can be readily obtained from a solution to HQP, and (iii) infeasibility of QP corresponds to a particular solution of HQP. An immediate consequence is that all the existing algorithms for QP are now also capable of robustly detecting infeasibility. In particular, we present an Infeasible Interior Point Method (IIPM) for the HQP and show polynomial iteration complexity when applied to HQP. A key distinction with prior IPM approaches is that we do not need to solve second-order cone programs. Numerical experiments on the formulation are provided using existing codes.

I. INTRODUCTION

Optimization algorithms are widely used in the control for constrained dynamic systems. In particular, the solution of convex Quadratic Programs (QPs) is a key ingredient of optimization algorithms. Convex QPs arise for example in the Model Predictive Control of linear systems, switched linear systems and nonlinear systems [1], [2]. The past two decades have witnessed the development of a number of algorithms for the solution of convex QPs arising in the context of optimization-based control. Initial work on QP algorithms were based on solving a system of equations representing the optimality conditions such as Interior Point Method (IPM) for QP [3], active-set methods [4], [5] and more recently, IPM for Second Order Cone Programs (SOCPs) [6]. In the last decade there has been interest in the development of first-order approaches such as gradient projection [7], dual gradient projection [8], [9], and splitting methods [10], [11], [12], [13] and iterative second-order approaches [14]. Recently, there has been work on employing these methods within Branch & Bound (B&B) algorithms for the solution of Mixed Integer Quadratic Programs (MIQPs) that arise in MPC for switched systems [15], [16], [17], [18].

A critical feature required of QP solvers for real-time applications is the ability to detect infeasibility and provide a graceful handling of the same. Infeasibility handling increases in importance when solving MIQPs in a B&B setting as it is expected that QPs resulting from the fixing of a subset of binary variables will likely be infeasible. Detecting such infeasible QPs and pruning the search tree in B&B is essential for computational efficiency of the MIQP solver.

Infeasibility of a QP is determined by the identification of a ray that is feasible for the dual and makes the dual objective unbounded. First-order primal-dual algorithms are capable of detecting infeasibility. Recent results on detecting infeasibility using first-order algorithms include [19], [20], [21]. A dual QP formulation, if it can be constructed explicitly, will allow the determination of a ray of infeasibility. However, such an explicit dual construction requires strong convexity of the QP. Active-set methods typically employ a *feasibility* phase to first determine a feasible point and then proceed to an *optimization* phase which produces the optimal solution. Failure of the feasibility phase allows to pronounce a QP as being infeasible [22]. The work performed in this phase is identical to that required for the *optimization* phase. When employing inexact linear algebra the first-phase is typically not employed and as a consequence the ability of such approaches to certify infeasibility is not clear. IPM algorithms, on the other hand, are not all capable of producing a certificate of infeasibility [23], [24]. IPMs based on Homogeneous Self-Dual (HSD) embedding [25], [26] are known to produce a certificate of infeasibility for Linear Programs (LPs). This technique has been implemented in MOSEK [27] and has also been extended to handle SOCPs. To detect infeasible QPs with IPMs, the QPs must first be formulated as a SOCP to which the HSD embedding is then applied. The ECOS solver [6] does exactly this for the solution of QPs arising in MPC and can detect infeasibility.

We present a novel embedding of the QP into a space that is one dimension higher. The embedding is produced by introducing an additional variable that is nonnegative and multiplies the affine terms in the objective and constraints. The objective contribution from the introduced variable is determined so the resulting Homogeneous Quadratic Program (HQP) satisfies the following properties: (i) HQP is always feasible, (ii) an optimal solution to the QP can always be recovered from the optimal solution to the HQP if the introduced variable is positive; and (iii) QP can be certified to be infeasible if the introduced variable is zero. More importantly, the HQP formulation can be presented to any QP solver to determine the optimal solution or infer infeasibility in a robust manner.

In this paper, we focus on applying an Infeasible IPM (IIPM) to solve the HQP. We are able to derive polynomial complexity of a standard IIPM when applied to HQP by adapting the analysis of IIPM for linear programs [23, Chapter 6]. Thus, we have an IIPM with polynomial iteration complexity for inferring optimality or infeasibility of QP without having to solve SOCPs.

The paper is organized as follows. §II presents the QP and the assumptions. §III presents the embedding of QP into a HQP and shows the equivalence of the formulations. An IIPM for solving HQP and the polynomial iteration complexity are described in §IV. We present results in §V when using the HQP formulation with IPOPT. Conclusions and directions for future work are provided in §VI.

Notation. The set of reals is denoted by \mathbb{R} and the set of vectors of dimension n by \mathbb{R}^n . The set of $n \times n$ symmetric matrices is denoted by \mathbb{S}^n . For a matrix $A \in \mathbb{S}^n$, the notation $A \succ (\succsim) 0$ denotes that A is positive semidefinite (definite). Given a vector $u \in \mathbb{R}^n$, $\text{Diag}(u)$ denotes a diagonal matrix with the elements of the vector on the diagonal. The notation I_n denotes the identity matrix of size n .

II. PROBLEM FORMULATION

Consider the QP in the form

$$\min_{y \in \mathbb{R}^n} \frac{1}{2} y^T C y + c^T y \quad (1a)$$

$$\text{s.t. } E y = f \quad (1b)$$

$$y \geq 0 \quad (1c)$$

where $C \in \mathbb{S}^n$, $c \in \mathbb{R}^n$, $E \in \mathbb{R}^{m \times n}$, and $f \in \mathbb{R}^m$. Note that any QP formulation with finite lower or upper bound on each of the variables can be put into the form in (1) through a linear transformation of variables.

We make the following assumptions on QP (1).

Assumption 1: The matrix E has full row rank of m .

Assumption 2: The matrix $Z^T C Z \succ 0$ where $Z \in \mathbb{R}^{n \times (n-m)}$ is an orthonormal basis for the null-space of E , i.e. Z is a basis for $\{v \mid E v = 0\}$.

Assumption 1 is not restrictive and can be easily satisfied by removing dependent rows if necessary. Assumption 2 requires that C be positive definite when projected onto Z . This assumption readily holds for MPC formulations [19] and also for certain spectral relaxations of nonconvex MIQPs [28], [29]. Assumption 2 implies that the optimal solution to QP (1) is unique whenever QP (1) is feasible. Note that the assumptions do not preclude the infeasibility of QP (1). The basis Z is typically computed using a QR factorization of E^T [30].

We conclude this section with a statement of the conditions for optimality and infeasibility of QP (1).

A point y^* minimizes QP (1) if there exist multipliers $\nu^* \in \mathbb{R}^m$, $\xi^* \geq 0 \in \mathbb{R}^n$, satisfying the first-order optimality conditions

$$C y^* + c + E^T \nu^* - \xi^* = 0 \quad (2a)$$

$$E y^* = f \quad (2b)$$

$$0 \leq y^* \perp \xi^* \geq 0. \quad (2c)$$

The conditions (2) are necessary and sufficient for a minimizer of QP (1) under Assumption 2.

The QP (1) is infeasible if there exist $\nu^\circ \in \mathbb{R}^m$, $\xi^\circ \in \mathbb{R}^n$ satisfying

$$E^T \nu^\circ - \xi^\circ = 0 \quad (3a)$$

$$f^T \nu^\circ = -1 \quad (3b)$$

$$\xi^\circ \geq 0. \quad (3c)$$

The conditions in (3) are obtained from applying a Theorem of the Alternative [31] to the constraints in (1b)-(1c).

III. HOMOGENEOUS QP FORMULATION

Consider embedding QP (1) into a Homogeneous Quadratic Program (HQP) as

$$\min_{y \in \mathbb{R}^n, \tau \in \mathbb{R}} \frac{1}{2} y^T C y + \tau c^T y + \frac{\theta}{2} (\tau^2 - 2\tau) \quad (4a)$$

$$\text{s.t. } E y = f \tau \quad (4b)$$

$$y \geq 0, \tau \geq 0 \quad (4c)$$

where $\tau \in \mathbb{R}$ is an additional nonnegative variable and $\theta > 0$ is a parameter. Observe that the objective (4a) is obtained by multiplying the linear term $c^T y$ in (1a) with τ and appending with the term $(\theta/2)(\tau^2 - \tau)$. The equality constraints (4b) is obtained by multiplying the right-hand side of (1b) with τ . §III-A provides conditions on θ that ensures HQP is convex. §III-B shows how the parameter θ can be computed efficiently. Finally, §III-C shows that solving HQP allows to recover a solution to QP or declare infeasibility.

A. Conditions on θ

The parameter θ is chosen to satisfy two conditions:

1) $\theta > 2|\theta^*|$ where θ^* is defined as

$$\theta^* = \min_{y \in \mathbb{R}^n} \frac{1}{2} y^T C y + c^T y \quad (5)$$

$$\text{s.t. } E y = f.$$

The parameter θ^* is well-defined by Assumption 2 and can be computed as $\theta^* = \frac{1}{2} \tilde{y}^T C \tilde{y} + c^T \tilde{y}$ where \tilde{y} is obtained by solving a single linear system

$$\begin{pmatrix} C & E^T \\ E & 0 \end{pmatrix} \begin{pmatrix} \tilde{y} \\ \tilde{\nu} \end{pmatrix} = \begin{pmatrix} -c \\ f \end{pmatrix}. \quad (6)$$

Since the nonnegativity constraints (1c) are ignored in the definition of θ^* (5) it follows that

$$-\frac{1}{2} \theta < \theta^* \leq \frac{1}{2} y^T C y + c^T y \quad (7)$$

$$\forall y \text{ satisfying (1b) - (1c).}$$

2) θ is chosen large enough so that

$$\widehat{Z}^T \begin{bmatrix} C & c \\ c^T & \theta \end{bmatrix} \widehat{Z} \succ 0 \quad (8)$$

where \widehat{Z} is a basis for the null space of (4b). Such a basis can be obtained as ,

$$\widehat{Z} = \begin{bmatrix} Z & d \\ 0 & 1 \end{bmatrix} \quad (9)$$

where $d = E^T (E E^T)^{-1} f$.

The satisfaction of (8) implies that:

- 1) the Hessian of the objective in (4a) is positive definite on the null space of the equality constraints (4b).
- 2) HQP (4) is convex and the first-order optimality conditions are necessary and sufficient for a minimizer.

B. Computing θ

We will now address the computation of θ satisfying (8).

Lemma 1: Suppose Assumptions 1 and 2 hold. Then (8) holds for all θ satisfying

$$\theta > \frac{1}{\lambda_{\min}(Z^T CZ)} \|Z^T(Cd + c)\|^2 - d^T Cd - 2c^T d. \quad (10)$$

Proof: The claim follows from an application of the Schur-complement to the left-hand side of the matrix in (8). Substitute \widehat{Z} from (9) in (8) to obtain

$$\widehat{Z}^T \begin{bmatrix} C & c \\ c^T & \theta \end{bmatrix} \widehat{Z} = \begin{bmatrix} Z^T CZ & Z^T(Cd + c) \\ (Cd + c)^T Z & \theta + d^T Cd + 2c^T d \end{bmatrix}. \quad (11)$$

Since $Z^T CZ \succ 0$ (Assumption 2) the condition in (8) holds provided the Schur-complement w.r.t. $Z^T CZ$ in (11) is positive, i.e.

$$\begin{aligned} 0 < \theta + d^T Cd + 2c^T d - \\ & (Cd + c)^T Z (Z^T CZ)^{-1} Z^T (Cd + c) \\ \implies \theta > (Cd + c)^T Z (Z^T CZ)^{-1} Z^T (Cd + c) \\ & - d^T Cd - 2c^T d. \end{aligned} \quad (12)$$

Using $Z^T CZ \succeq \lambda_{\min}(Z^T CZ)I_{n-m}$ in (12) yields (10). ■

Remark 1: It is not necessary to compute $Z^T(Cd + c)$ explicitly. Since $\|ZZ^T\| \leq 1$ it is sufficient to choose

$$\theta > \frac{1}{\lambda_{\min}(Z^T CZ)} \|Cd + c\|^2 - d^T Cd - 2c^T d \quad (13)$$

in order to satisfy (8).

Remark 2: Further, it is not necessary to compute $\lambda_{\min}(Z^T CZ)$. It is sufficient to have a nontrivial lower bound estimate of $\lambda_{\min}(Z^T CZ)$. For example, if $C \succ 0$ (as in MPC applications) then $\lambda_{\min}(C) \leq \lambda_{\min}(Z^T CZ)$. In the context of MIQPs with binary variables, QPs (1) are solved at each node of the B&B tree. The QP solved at the child node is a result of fixing a binary variable to 0 or 1. Let us suppose that in the child node the first variable index is fixed to 0 or 1. Then

$$\lambda_{\min}(Z^T CZ) = \min_{u:Eu=0} \frac{u^T Cu}{u^T u} \leq \min_{u:Eu=0, u_1 \in \{0,1\}} \frac{u^T Cu}{u^T u}.$$

Hence, the lower bound estimate available at the parent node serves as a valid lower bound estimate for smallest eigenvalue of the reduced Hessian at the child node. If $0 < \alpha \leq \lambda_{\min}(Z^T CZ)$ then the choice of

$$\theta > \frac{1}{\alpha} \|Cd + c\|^2 - d^T Cd - 2c^T d \quad (14)$$

ensures satisfaction of (8).

C. Equivalence between QP and HQP

We collect some simple observations on the HQP (4).

- (O1) HQP (4) is always feasible. It is easily verified that $(y, \tau) = 0$ satisfies (4b)-(4c).
- (O2) HQP (4) has an optimal solution with optimal value less than or equal to 0. This follows directly from (O1).
- (O3) HQP (4) has a finite optimum. This follows from (8).

We now state the first-order optimality conditions for HQP (4).

A point $(\widehat{y}, \widehat{\tau})$ minimizes HQP (4) if there exist multipliers $(\widehat{\nu}, \widehat{\xi}, \widehat{\omega}) \in \mathbb{R}^{m+n+1}$ satisfying the first-order optimality conditions

$$C\widehat{y} + \widehat{\tau}c + E^T \widehat{\nu} - \widehat{\xi} = 0 \quad (15a)$$

$$\theta \widehat{\tau} - \theta + c^T \widehat{y} - f^T \widehat{\nu} - \widehat{\omega} = 0 \quad (15b)$$

$$E\widehat{y} = f\widehat{\tau} \quad (15c)$$

$$0 \leq \widehat{y} \perp \widehat{\xi} \geq 0 \quad (15d)$$

$$0 \leq \widehat{\tau} \perp \widehat{\omega} \geq 0. \quad (15e)$$

Theorem 1: Suppose θ is chosen to satisfy the conditions in §III-A. The QP (1) has an optimal solution y^* iff the HQP (4) has an optimal solution $(\widehat{y}, \widehat{\tau})$ with $\widehat{\tau} > 0$.

Proof: Consider the if part. Let $(\widehat{y}, \widehat{\tau})$ be an optimal solution of HQP (4) and let $(\widehat{\nu}, \widehat{\xi}, \widehat{\omega})$ be multipliers such that (15) holds. Then, it is easily verified that $(y^*, \nu^*, \xi^*) = (\widehat{y}/\widehat{\tau}, \widehat{\nu}/\widehat{\tau}, \widehat{\xi}/\widehat{\tau})$ satisfies the optimality conditions (2) for QP (1). This proves the if part.

Consider the only if part. Let y^* be an optimal solution of QP (1) and let (ν^*, ξ^*) be multipliers such that (2) holds. Define $\bar{\tau} = \theta/(\theta + c^T y^* - f^T \nu^*)$. If $\bar{\tau} > 0$ then it can be easily verified that $\widehat{y} = \bar{\tau} y^*$, $\widehat{\tau} = \bar{\tau}$, $\widehat{\nu} = \bar{\tau} \nu^*$, $\widehat{\xi} = \bar{\tau} \xi^*$, $\widehat{\omega} = 0$ satisfy the optimality conditions for HQP (15). To show $\bar{\tau} > 0$ we need to show that $\theta + c^T y^* - f^T \nu^* > 0$ since $\theta > 0$. Consider

$$\theta + c^T y^* - f^T \nu^* \quad (16a)$$

$$= \theta + c^T y^* - (E^T \nu^*)^T y^* \quad (16b)$$

$$= \theta + 2c^T y^* + (y^*)^T C y^* - (y^*)^T \xi^* \quad (16c)$$

$$= \theta + 2c^T y^* + (y^*)^T C y^* \quad (16d)$$

$$> 0 \quad (16e)$$

where the equality in (16b) follows by multiplying (2b) by $(\nu^*)^T$ and substituting for $f^T \nu^*$ with $(E^T \nu^*)^T y^*$. Multiplying (2a) by $(y^*)^T$ and substituting for $-(E^T \nu^*)^T y^*$ as $c^T y^* + (y^*)^T C y^* - (y^*)^T \xi^*$ yields (16c). Using the complementarity constraints (2c) in (16c) yields (16d). The final inequality follows from (7). This proves the only if part of the claim. The claim is proven. ■

We now show that infeasibility of QP (1) is equivalent to the vanishing of the optimal solution to HQP (4).

Theorem 2: Suppose θ satisfies the conditions in §III-A. The QP (1) is infeasible iff the HQP (4) has optimal solution $(\widehat{y}, \widehat{\tau}) = 0$.

Proof: Consider the only if part of the claim. Suppose there exists (ν°, ξ°) satisfying (3). It can be verified that $(\widehat{y}, \widehat{\tau}) = 0$, $(\widehat{\nu}, \widehat{\xi}, \widehat{\omega}) = (\theta \nu^\circ, \theta \xi^\circ, 0)$ satisfies (15). Hence $(\widehat{y}, \widehat{\tau}) = 0$ is an optimal solution to HQP (4). This proves the only if part of the claim.

Consider the if part of the claim. Suppose $(\widehat{y}, \widehat{\tau}) = 0$ is the optimal solution to (4) and let $(\widehat{\nu}, \widehat{\xi}, \widehat{\omega})$ be the multipliers in (15). Then $(\nu^\circ, \xi^\circ) = (\widehat{\nu}/(\theta + \widehat{\omega}), \widehat{\xi}/(\theta + \widehat{\omega}))$ can be verified to satisfy (3). This proves the if part of the claim, completing the proof. ■

IV. INFEASIBLE INTERIOR POINT METHOD (IIPM)

For the remainder of the paper, we assume that the HQP has the form

$$\min_{x \in \mathbb{R}^{n+1}} \frac{1}{2} x^T Q x + q^T x \quad (17a)$$

$$\text{s.t. } Ax = 0 \quad (17b)$$

$$x \geq 0 \quad (17c)$$

where $x = \begin{pmatrix} y \\ \tau \end{pmatrix}$, $Q = \begin{pmatrix} C & c^T \\ c & \theta \end{pmatrix}$, $q = \begin{pmatrix} 0 \\ -\theta \end{pmatrix}$, $A = \begin{pmatrix} E & -f \end{pmatrix}$. We assume the following for the HQP (17).

Assumption 3: The matrix A has full row rank of m .

Assumption 4: The matrix Q is positive definite on the null space of A , i.e. $x^T Q x > 0 \forall x \in \{v \mid Av = 0\}$.

The above HQP can be identified with original QP (1) when $f = 0$ and the Assumptions 1-2 imply satisfaction of Assumptions 3-4. If $f \neq 0$ §III shows how to cast (1) satisfying Assumptions 1-2 into a HQP of the form in (17) satisfying Assumptions 3-4.

In the rest of the section, we show how a standard IIPM for LPs can be extended to HQP (17). The IIPM also enjoys polynomial iteration complexity.

A. IIPM for HQP

In the rest of the section we employ the functions

$$r_d(x, \lambda, s) := Qx + A^T \lambda - s \quad (18a)$$

$$r_p(x, \lambda, s) := Ax \quad (18b)$$

$$r_c(x, \lambda, s) := Xs \quad (18c)$$

$$\mu(x, s) := x^T s / (n + 1) \quad (18d)$$

where $x \in \mathbb{R}^{n+1}$, and $\lambda \in \mathbb{R}^m$, $s \in \mathbb{R}^{n+1}$ are the multipliers for the equality, nonnegativity constraints (17b)-(17c), and $X = \text{Diag}(x)$. The parameter $\mu(x, s)$ is called a barrier or centrality parameter. In the following, we will suppress the arguments when the dependence is clear from the context.

A point $x^* \geq 0$ is said to minimize HQP (17) if there exist λ^* and $s^* \geq 0$ satisfying the first-order stationary conditions

$$(r_d(x, \lambda, s), r_p(x, \lambda, s), r_c(x, \lambda, s)) = (0, 0, 0). \quad (19)$$

IIPMs aim to compute (x^*, λ^*, s^*) by following the central path which is defined by $\{(x, \lambda, s) \mid (x, s) > 0, (r_d(x, \lambda, s), r_p(x, \lambda, s), r_c(x, \lambda, s)) = (0, 0, \mu(x, s)e)\}$ where $e \in \mathbb{R}^{n+1}$ is a vector of all ones. In the limit as $\mu(x, s) \rightarrow 0$ we recover a point satisfying (19). Feasible IIPMs typically assume that an initial iterate (x, λ, s) satisfying $(x, s) > 0$ and $(r_d(x, \lambda, s), r_p(x, \lambda, s)) = (0, 0)$. Such a point is not generally guaranteed to exist and can also be difficult to compute. In the context of HQP (17) such a point will not exist if the original QP (1) is infeasible (refer Theorem 2). This is our motivation for considering IIPM.

IIPMs start from an initial iterate (x^0, λ^0, s^0) satisfying $(x^0, s^0) > 0$ but $(r_d^0, r_p^0) \neq 0$ where r_d^0, r_p^0 denote $r_d(x^0, \lambda^0, s^0)$, $r_p(x^0, \lambda^0, s^0)$. The method generates a sequence of iterates $\{(x^k, \lambda^k, s^k)\}$ where the iterates are

required to lie within a neighborhood $\mathcal{N}_{-\infty}(\gamma, \beta)$ with

$$\mathcal{N}_{-\infty}(\gamma, \beta) = \left\{ (x, \lambda, s) \mid \begin{array}{l} \frac{\|(r_d, r_p)\|}{\mu} \leq \beta \frac{\|(r_d^0, r_p^0)\|}{\mu^0} \\ (x, s) > 0, Xs \geq \gamma \mu(x, s)e \end{array} \right\} \quad (20)$$

where $\beta \geq 1$ is a fixed parameter. At each iteration a search direction $(\Delta x^k, \Delta \lambda^k, \Delta s^k)$ is generated by solving

$$\begin{pmatrix} Q & A^T & -I_{n+1} \\ A & 0 & 0 \\ S^k & 0 & X^k \end{pmatrix} \begin{pmatrix} \Delta x^k \\ \Delta \lambda^k \\ \Delta s^k \end{pmatrix} = \begin{pmatrix} -r_d^k \\ -r_p^k \\ -r_c^k + \sigma^k \mu^k e \end{pmatrix} \quad (21)$$

where $\sigma^k > 0$ such that $\sigma^k \in [\sigma^{\min}, \sigma^{\max}]$ and $\mu^k = (x^k)^T s^k / (n + 1)$. The residual of the complementarity condition r_c^k is perturbed by $\sigma^k \mu^k$ to ensure that the iterates (x^{k+1}, s^{k+1}) can be guaranteed to lie within the neighborhood $\mathcal{N}_{-\infty}(\gamma, \beta)$. Given such a direction define $(x(\alpha), \lambda(\alpha), s(\alpha)) = (x^k, \lambda^k, s^k) + \alpha(\Delta x^k, \Delta \lambda^k, \Delta s^k)$. The step α^k is chosen to be the largest value such that

$$(x(\alpha^k), \lambda(\alpha^k), s(\alpha^k)) \in \mathcal{N}_{-\infty}(\gamma, \beta) \quad (22a)$$

$$\mu(\alpha^k) \leq (1 - 0.01\alpha^k)\mu^k. \quad (22b)$$

The next iterate is defined as $(x^{k+1}, \lambda^{k+1}, s^{k+1}) = (x(\alpha^k), \lambda(\alpha^k), s(\alpha^k))$. Algorithm 1 describes the IIPM.

Algorithm 1: IIPM for HQP (17)

Data: $\beta, \gamma, \sigma^{\min}, \sigma^{\max}$ with $\beta \geq 1, \gamma \in (0, 1), 0 < \sigma^{\min} < \sigma^{\max} \leq \frac{1}{2}$.

1 Choose (x^0, λ^0, s^0) with $x^0, s^0 > 0$.

2 **for** $k = 0, 1, \dots$ **do**

3 Choose $\sigma^k \in [\sigma^{\min}, \sigma^{\max}]$ and solve (21) to obtain $(\Delta x^k, \Delta \lambda^k, \Delta s^k)$.

4 Choose the largest $\alpha^k > 0$ such that (22) holds.

5 Set $(x^{k+1}, \lambda^{k+1}, s^{k+1}) = (x^k, \lambda^k, s^k) + \alpha^k(\Delta x^k, \Delta \lambda^k, \Delta s^k)$.

B. Polynomial Complexity

The analysis of the Algorithm 1 is straightforward adaptation of the analysis in [23, Chapter 6]. We will only highlight the key steps where we differ. The results are specialized to the case where the initial iterate is chosen as

$$(x^0, \lambda^0, s^0) = (\zeta e, 0, \zeta e) \quad (23)$$

where $\zeta > 0$ is scalar for which

$$\|(x^*, \lambda^*, s^*)\| \leq \zeta \quad (24)$$

for some optimal solution (x^*, λ^*, s^*) of HQP (17). Recall from the discussion in §III that such a solution always exists under the Assumptions 3-4.

From the linearity of the residuals in (18a)-(18b), the choice of step direction (21) and the method of updating the iterate at $(k + 1)$ it is easy to show that

$$(r_d^{k+1}, r_p^{k+1}) = (1 - \alpha^k)(r_d^k, r_p^k) = \nu^{k+1}(r_d^0, r_p^0)$$

where $v^{k+1} = \prod_{j=0}^k (1 - \alpha^j)$. We first state a key result that is used in subsequent lemmas.

Lemma 2: Suppose $(\bar{x}, \bar{\lambda}, \bar{s})$ be such that $r_d(\bar{x}, \bar{\lambda}, \bar{s}) = 0$ and $r_p(\bar{x}, \bar{\lambda}, \bar{s}) = 0$. Then $\bar{x}^T Q \bar{x} = \bar{x}^T \bar{s}$. Further, $\bar{x}^T \bar{s} \geq 0$.

Proof: Substitute $(\bar{x}, \bar{\lambda}, \bar{s})$ in (18a) and multiplying by \bar{x}^T obtain

$$\bar{x}^T r_d(\bar{x}, \bar{\lambda}, \bar{s}) = \bar{x}^T Q \bar{x} + (A\bar{x})^T \bar{\lambda} - \bar{x}^T \bar{s} = 0.$$

Using $A\bar{x} = 0$ in the above proves the claim. The second claim follows from Assumption 4. ■

The proof on complexity (Theorem 3) relies on 3 key lemmas. Lemma 3 first bounds $v^k \|(x^k, s^k)\|_1$. Lemma 4 provides a bound on scaled search directions. Finally, Lemma ?? ensures that there exists a uniform lower bound on α^k . We first provide a bound on $v^k(x^k, s^k)$.

Lemma 3: Suppose the initial iterate satisfies (23). Then for any iterate (x^k, λ^k, s^k)

$$\zeta v^k \|(x^k, s^k)\|_1 \leq 4\beta(n+1)\mu^k. \quad (25)$$

Proof: Define

$$\begin{aligned} (\bar{x}, \bar{\lambda}, \bar{s}) &= v^k(x^0, \lambda^0, s^0) + (1 - v^k)(x^*, \lambda^*, s^*) \\ &\quad - (x^k, \lambda^k, s^k). \end{aligned} \quad (26a)$$

Then $(\bar{x}, \bar{\lambda}, \bar{s})$ satisfy the conditions in Lemma 2 and $\bar{x}^T \bar{s} \geq 0$. Hence,

$$\begin{aligned} 0 &\leq \bar{x}^T \bar{s} \\ &= (v^k)^2 (x^0)^T s^0 + (1 - v^k)^2 (x^*)^T s^* + (x^k)^T s^k \\ &\quad + v^k(1 - v^k)((x^0)^T s^* + (x^*)^T s^0) \\ &\quad - v^k((x^0)^T s^k + (x^k)^T s^0) \\ &\quad - (1 - v^k)((x^*)^T s^k + (x^k)^T s^*). \end{aligned}$$

Using $(x^*)^T s^* = 0$ and $(x^*)^T s^k + (x^k)^T s^* \geq 0$ in the above and rearranging obtain

$$\begin{aligned} &v^k((x^0)^T s^k + (x^k)^T s^0) \\ &\leq (v^k)^2 (x^0)^T s^0 + (x^k)^T s^k \\ &\quad + v^k(1 - v^k)((x^0)^T s^* + (x^*)^T s^0). \end{aligned}$$

The above is identical to [23, eqn. (6.20)] and the arguments in [23, Lemmas 6.3-6.4] apply to yield the result. ■

We next provide a bound on the scaled search directions.

Lemma 4: Suppose the initial iterate satisfies (23). Then there is a constant $\eta > 0$ independent of n such that

$$\|(D^k)^{-1} \Delta x^k\| \leq \eta(n+1)\mu^k, \quad \|D^k \Delta s^k\| \leq \eta(n+1)\mu^k$$

where $D^k = (X^k)^{\frac{1}{2}}(S^k)^{-\frac{1}{2}}$.

Proof: Define

$$\begin{aligned} (\bar{x}, \bar{\lambda}, \bar{s}) &= (\Delta x^k, \Delta \lambda^k, \Delta s^k) + v^k(x^0, \lambda^0, s^0) \\ &\quad - v^k(x^*, \lambda^*, s^*) \end{aligned} \quad (27a)$$

Then $(\bar{x}, \bar{\lambda}, \bar{s})$ satisfy the conditions in Lemma 2 and $\bar{x}^T \bar{s} \geq 0$. From the last row in (21)

$$\begin{aligned} &S^k(\Delta x^k + v^k(x^0 - x^*)) + X^k(\Delta s^k + v^k(s^0 - s^*)) \\ &= -X^k s^k + \sigma^k \mu^k e + v^k S^k(x^0 - x^*) + v^k X^k(s^0 - s^*). \end{aligned} \quad (27b)$$

Multiplying through by $(X^k S^k)^{-\frac{1}{2}}$ and the definition of D^k obtain

$$\begin{aligned} &(D^k)^{-1}(\Delta x^k + v^k(x^0 - x^*)) + D^k(\Delta s^k + v^k(s^0 - s^*)) \\ &= -(X^k S^k)^{-\frac{1}{2}}(X^k s^k - \sigma^k \mu^k e) \\ &\quad + v^k(D^k)^{-1}(x^0 - x^*) + v^k D^k(s^0 - s^*). \end{aligned} \quad (27c)$$

Taking norms on both sides, squaring and using $\bar{x}^T \bar{s} \geq 0$ obtain

$$\begin{aligned} &\|(D^k)^{-1}(\Delta x^k + v^k(x^0 - x^*))\|^2 \\ &\quad + \|D^k(\Delta s^k + v^k(s^0 - s^*))\|^2 \\ &\leq \left\{ \begin{aligned} &\|(X^k S^k)^{-\frac{1}{2}}\| \|(X^k s^k - \sigma^k \mu^k e)\| \\ &+ v^k \|(D^k)^{-1}(x^0 - x^*)\| + v^k \|D^k(s^0 - s^*)\| \end{aligned} \right\}^2. \end{aligned} \quad (27d)$$

Starting from this inequality the arguments in the proofs of Lemmas 6.5-6.6 [23] can be followed to show the claim. ■

Lemma 6.7 [23] provides a uniform bound on α^k and holds without any change in the arguments.

The claim on polynomial complexity of Algorithm 1 follows from the Lemmas 3,4 and Lemma 6.7 [23]. The arguments are identical to those in [23, Theorem 6.2]. We state the complexity result without proof.

Theorem 3: Let $\epsilon > 0$ be given. Suppose that the initial iterate satisfies (23) and suppose that $\zeta > 0$ satisfies $\zeta^2 \leq \chi/\epsilon^\kappa$ for some constants $\chi, \kappa > 0$. Then there is an index $K = O(n^2 |\log \epsilon|)$ such that the iterates (x^k, λ^k, s^k) generated by Algorithm 1 satisfy $\mu^k \leq \epsilon$ for all $k \geq K$.

V. NUMERICAL EXPERIMENTS

We present preliminary numerical experiments with the HQP formulation using the IPM solver IPOPT [32]. IPOPT implements the standard IPM for QPs i.e. the algorithm cannot determine a certificate of infeasibility. In our experiments, we generated random instances of QP (1) with a single equality constraint with $C = I_n$, $c = e$, $E \in \mathbb{R}^{1 \times n}$, $f = -1$. The coefficients in the single equality are restricted to be nonnegative and generated at random. It is easy to deduce that there exists no $y \geq 0$ such that $Ey < 0$ for $E \geq 0$. Hence, the instances are all infeasible. We presented the QP formulation directly to IPOPT. The HQP formulation (4) is derived for each such random instance and is also presented for solution to IPOPT. We generated 10 random instances of different sizes $n \in \{10, 25, 50\}$.

In the first set of experiments, we set the IPOPT option `mehrotra_algorithm=no`. In this case, the predictor-corrector algorithm of Mehrotra is disabled. When solving the QP formulation, IPOPT always terminates after 8-10 iterations with the indication that the problem might be infeasible. When solving the HQP formulation, IPOPT always terminates with an optimal solution of 0 as specified in Theorem 2 in about 16-20 iterations.

In the second set of experiments, we enabled the predictor-corrector algorithm in IPOPT by setting `mehrotra_algorithm=yes`. In this setting IPOPT hits the iteration limit on every single instance when solving the QP formulation. The dual infeasibility blows

up to infinity in every instance. When solving the HQP formulation IPOPT always terminates with an optimal solution of 0 in about 8-10 iterations.

In the third set of experiments, we set $f = 1$. In this case the QP instances are always feasible. We verified that both the QP and HQP formulations solved the problem to optimality. Further, the solutions to QP can be recovered from the HQP could be recovered according to Theorem 1. There is no significant difference in the number of iterations taken for convergence using either formulation. However, the computational time per iteration for HQP is higher than that of QP formulation. This is likely due to the matrix in (21) being dense. This can be rectified using a tailored Schur-complement-based implementation for the step computation. We will investigate this in a future work.

VI. CONCLUSIONS & FUTURE WORK

We presented a homogeneous formulation of QP that allows for robust detection of infeasibility in QPs. We also presented an infeasible IPM for the solution of the HQP and showed polynomial iteration complexity. In the context of IPMs, the paper leaves open a number of future avenues for exploration. Firstly, the sparsity of the linear systems in the step computation of IIPM are now affected by the sparsity of c, f which can be dense. To effectively handle this a tailored linear algebra solution is required. Secondly, the linear system E may have structure such as in the case of MPC. Exploiting this through a decomposition approach such as in [3] is critical to improving the computational efficiency. Thirdly, the user of predictor-corrector steps will be crucial for improving the practical performance. We will explore these in a future work. The homogeneous formulation can be readily presented to any QP algorithm. We will also investigate the performance of other QP algorithms on the proposed formulation.

REFERENCES

- [1] F. Borrelli, A. Bemporad, and M. Morari, *Predictive Control for Linear and Hybrid Systems*. Cambridge University Press, 2017.
- [2] D. Mayne, J. Rawlings, and M. Diehl, *Model predictive control: Theory, Computation and Design*, 2nd ed. Santa Barbara: Nob Hill, LLC, 2020.
- [3] C. Rao, S. Wright, and J. Rawlings, "Application of interior-point methods to model predictive control," *Journal of Optimization Theory and Applications*, p. 723–757, 1998.
- [4] R. Bartlett and L. Biegler, "Qpschur: A dual, active-set, schur-complement method for large-scale and structured convex quadratic programming," *Optimization and Engineering*, vol. 7, p. 5–32, 2006.
- [5] H. Ferreau, H. Bock, and M. Diehl, "An online active set strategy to overcome the limitations of explicit mpc," *International Journal of Robust and Nonlinear Control*, vol. 18, no. 8, p. 816–830, 2008.
- [6] A. Domahidi, E. Chu, and S. Boyd, "Ecos: An sopc solver for embedded systems," in *2013 European Control Conference (ECC)*, 2013, pp. 3071–3076.
- [7] S. Richter, C. Jones, and M. Morari, "Computational complexity certification for real-time mpc with input constraints based on the fast gradient method," *IEEE Trans. Autom. Control*, vol. 57, no. 6, p. 1391–1403, 2012.
- [8] P. Giselsson, "Optimal preconditioning and iteration complexity bounds for gradient-based optimization in model predictive control," in *2013 American Control Conference*, 2013, pp. 358–364.
- [9] P. Patrinos and A. Bemporad, "An accelerated dual gradient-projection algorithm for embedded linear model predictive control," *IEEE Transactions on Automatic Control*, vol. 59, no. 1, p. 18–33, 2014.
- [10] A. U. Raghunathan and S. Di Cairano, "Alternating direction method of multipliers for strictly convex quadratic programs: Optimal parameter selection," in *American Control Conference*, 2014, pp. 4324–4329.
- [11] P. Giselsson and S. Boyd, "Diagonal scaling in douglas-rachford splitting and admm," in *53rd IEEE Conference on Decision and Control*, 2014, pp. 5033–5039.
- [12] P. Patrinos, L. Stella, and A. Bemporad, "Douglas-rachford splitting: Complexity estimates and accelerated variants," in *53rd IEEE Conference on Decision and Control*, 2014, pp. 4234–4239.
- [13] E. Ghadimi, A. Teixeira, I. Shames, and M. Johansson, "Optimal parameter selection for the alternating direction method of multipliers (admm): Quadratic problems," *IEEE Transactions on Automatic Control*, vol. 60, no. 3, pp. 644–658, 2015.
- [14] R. Quirynen, A. Knyazev, and S. D. Cairano, "Block structured preconditioning within an active-set method for real-time optimal control," in *2018 European Control Conference (ECC)*, 2018, pp. 1154–1159.
- [15] D. Frick, A. Domahidi, and M. Morari, "Embedded optimization for mixed logical dynamical systems," *Computers & Chemical Engineering*, vol. 72, pp. 21–33, 2015.
- [16] V. Naik and A. Bemporad, "Embedded mixed-integer quadratic optimization using accelerated dual gradient projection," in *IFAC-PapersOnLine*, vol. 50, no. 1, 2017, pp. 10723–10728.
- [17] B. Stellato, V. Naik, A. Bemporad, P. Goulart, and S. Boyd, "Embedded mixed-integer quadratic optimization using the osqp solver," in *2018 European Control Conference (ECC)*, 2018, pp. 1536–1541.
- [18] P. Hespanhol, R. Quirynen, and S. Di Cairano, "A structure exploiting branch-and-bound algorithm for mixed-integer model predictive control," in *2019 18th European Control Conference (ECC)*, 2019, pp. 2763–2768.
- [19] A. U. Raghunathan and S. Di Cairano, "Infeasibility detection in alternating direction method of multipliers for convex quadratic programs," in *53rd IEEE Conference on Decision and Control*, 2014, pp. 5819–5824.
- [20] G. Banjac, P. Goulart, B. Stellato, and S. Boyd, "Infeasibility detection in the alternating direction method of multipliers for convex optimization," *Journal of Optimization Theory and Applications*, vol. 183, no. 2, pp. 490–519, 2019. [Online]. Available: <https://doi.org/10.1007/s10957-019-01575-y>
- [21] D. Liao-McPherson and I. Kolmanovskiy, "Fbstab: A proximally stabilized semismooth algorithm for convex quadratic programming," *Automatica*, vol. 113, p. 108801, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0005109819306648>
- [22] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York, NY, USA: Springer, 2006.
- [23] S. Wright, *Primal-Dual Interior-Point Methods*. SIAM, 1997.
- [24] E. Wong, "Active-set methods for quadratic programming," Ph.D. dissertation, University of California, San Diego, 2011.
- [25] Y. Ye, M. J. Todd, and S. Mizuno, "An $\mathcal{O}(\sqrt{n/l})$ -iteration homogeneous and self-dual linear programming algorithm," *Mathematics of Operations Research*, vol. 19, pp. 53–67, 1994.
- [26] X. Xu, P. Hung, and Y. Ye, "A simplified homogeneous and self-dual linear programming algorithm and its implementation," *Annals of Operations Research*, vol. 62, pp. 151–172, 1996.
- [27] E. Andersen and K. Andersen, "The mosek interior point optimizer for linear programming: an implementation of the homogeneous algorithm," *High Performance Optimization*, vol. 33, p. 197–232, 2000.
- [28] C. J. Nohra, A. U. Raghunathan, and N. Sahinidis, "Spectral relaxations and branching strategies for global optimization of mixed-integer quadratic programs," *SIAM Journal on Optimization*, vol. 31, no. 1, pp. 142–171, 2021.
- [29] C. J. Nohra, A. U. Raghunathan, and N. V. Sahinidis, "Sdp-quality bounds via convex quadratic relaxations for global optimization of mixed-integer quadratic programs," *Math. Program.*, 2021.
- [30] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Johns Hopkins University Press, 1996.
- [31] O. Mangasarian, *Nonlinear Programming*, ser. Classics in Applied Mathematics. SIAM, 1994.
- [32] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical programming*, vol. 106, no. 1, pp. 25–57, 2006.