Constrained LQR Design Using Interior-Point Arc-Search Method for Convex Quadratic Programming with Box Constraints

Yaguang Yang[*]

June 20, 2018

Abstract

Although the classical LQR design method has been very successful in real world engineering designs, in some cases, the classical design method needs modifications because of the saturation in actuators. This modified problem is sometimes called the constrained LQR design. For discrete systems, the constrained LQR design problem is equivalent to a convex quadratic programming problem *with box constraints*. We will show that the interior-point method is very efficient for this problem because an initial interior point is available, a condition which is not true for general convex quadratic programming problem. We will devise an effective and efficient algorithm for the constrained LQR design problem using the special structure of the box constraints and a recently introduced arc-search technique for the interior-point algorithm. We will prove that the algorithm is polynomial and has the best-known complexity bound $O(\sqrt{n}\log(1/\epsilon))$ for the convex quadratic programming. The proposed algorithm is implemented in MATLAB. An example for the constrained LQR design is provided to show the effectiveness and efficiency of the design method. The proposed algorithm can easily be used for model predictive control.

Keywords: Constrained LQR, arc-search, convex quadratic programming, polynomial algorithm.

1   Introduction

The LQR design of linear control system theory has been proven to be one of the most effective ways to design real world control systems because, in many cases, the linear system model has adequate accuracy and fidelity to describe the real world systems, the design method has manageable complexity, and is

---

[*]NRC, Office of Research, 21 Church Street, Rockville, 20850. Email: yaguang.yang@verizon.net

better understood and more mature than the design methods in nonlinear system theory. For example, a beautiful application that demonstrated all these nice features is described in [18]. One of the major obstacles in some applications for the LQR design or, in general, the classical linear control system theory is that it does not consider the reality that the actuators have some saturation limits. This drawback in the linear control system theory has been addressed by many people and some of the latest results are presented in some recently published books such as [11] [25]. Two of the most attractive and related methods are the constrained LQR design [4, 21] and the Model Predictive Control (MPC) design [17, 2], both involve solving some convex quadratic programming.

In [21], a general framework of the constrained LQR problem is formulated. The problem can be represented as a standard convex quadratic programming with $Nr + Nm$ variables, where N is the number of horizons, m is the number of control inputs, and r is the number of states. This problem is then solved by using two interior-point methods, The first method is similar to the one proposed in [9], the second method is similar to the one proposed in [13]. The analysis and computational experience demonstrate the same frustration that the interior-point algorithms with desirable theoretical properties (polynomial complexity) tend to be slow in computation, while little can be proven about the algorithms (such as Mehrotra Predictor Corrector algorithm [12, 20]) that perform much better in practice.

MPC design is a very active research area which has many industry applications [16] and numerous theoretical investigations [1]. A very attractive and important method is proposed in [2] where a general frame work of model predictive control problem is formulated. To reduce the on-line computational burden, the off-line design strategy is emphasized. The problem is solved by a multi-parametric program in which some convex quadratic programming will be solved repeatedly. Further considerations on efficiency is discussed recently by many people, for example [19]. These papers focus on the off-line strategy but pays no attention to the development of the best convex quadratic programming method for the MPC design.

In this paper, we will consider constrained LQR design problem in which the actuators have lower and upper bounds. This problem is slightly simpler than the problems considered in [21, 2, 19] but is still general enough for most real world problems. We will show, by using the state space equations, that the number of the variables of the constrained LQR problem described in [21] can easily be reduced from $Nr + Nm$ to Nm and all equality constraints can be removed. This means that the reduced problem is not only much smaller but also has a special structure which is called convex quadratic programming subject to box constraints. We will solve the reduced problem by an interior-point algorithm that searches the optimizer along an arc that approximates the central path. This algorithm is similar to a recently

developed algorithm [24] but it is especially designed for convex quadratic programming subject only to the box constraints and it is more efficient than the algorithm in [24] because of two improvements due to the special structure: (1) the enlarged search neighborhood, and (2) an explicit initial interior point (finding an initial interior point may be a major obstacle for feasible interior-point methods). We will show that this algorithm has both desirable theoretical properties (polynomial complexity) and superior performance in computation because of the above mentioned improvements. Although the idea of the proof of polynomiality is similar to that used in [24], we provided the proof in the appendix for several reasons: (1) it also shows the enlarged search neighborhood comparing to the linearly constrained convex quadratic programming, (2) some special cares are needed for box constraints, and (3) the completeness. We have implemented the algorithm in MATLAB. We will demonstrate by some numerical LQR design example that the proposed constrained LQR design is very effective and efficient.

The remainder of the paper is organized as follows. Section 2 introduces notations and some technical lemmas that will be used in the rest of the paper. Section 3 discusses the constrained LQR design and convex quadratic programming with box constraints. Section 4 describes the central path of quadratic programming with box constraints. Section 5 proposes an arc-search algorithm for the convex quadratic programming with box constraints. Section 6 gives convergence analysis. Section 7 addresses implementation issues. Section 8 presents some LQR design example. Section 9 summarizes the conclusions. Some technical proofs are in the appendix to enhance the readability of the paper.

2   Some Notations and Technical Lemmas

Throughout the paper, we will use notations adopted in [24]. We denote n-dimensional vector space by $\mathbf{R}^n$, $n \times m$-dimensional matrix space by $\mathbf{R}^{n \times m}$, Hadamard (element-wise) product of two vectors $y \in \mathbf{R}^n$ and $\lambda \in \mathbf{R}^n$ by $y \circ \lambda$, the ith component of y by y_i, element-wise division of the two vectors by $\frac{y}{\lambda}$ if $\min |\lambda_i| > 0$, the Euclidean norm of y by $\|y\|$, the identity matrix of any dimension by I, the vector of all ones with appropriate dimension by e, element-wise absolute value vector by $|y| = [|y_1|, \ldots, |y_n|]^{\mathrm{T}}$. To simplify the notation for block column vectors, we will denote, for example, $[y^{\mathrm{T}}, \lambda^{\mathrm{T}}]^{\mathrm{T}}$ by (y, λ). For vectors $x \in \mathbf{R}^n$, $y \in \mathbf{R}^n$, $z \in \mathbf{R}^n$, $\lambda \in \mathbf{R}^n$, and $\gamma \in \mathbf{R}^n$, we will use capital letters X, Y, Z, Λ, and Γ for some related diagonal matrices whose diagonal elements are the components of the corresponding vectors. For example, we will use $\Lambda = \mathrm{diag}(\lambda)$ and $\Gamma = \mathrm{diag}(\gamma)$ for the diagonal matrices. For a matrix $H \in \mathbf{R}^{n \times n}$, we use $H \geq 0$ if H is positive semidefinite, and $H > 0$ if H is positive definite. Finally, we define an initial point of any algorithm by x^0, the point after the kth iteration by x^k.

We will also use some technical lemmas which are independent of the problem. The first two simple lemmas are given in [22].

Lemma 2.1 *Let $p > 0$, $q > 0$, and $r > 0$ be some constants. If $p + q \leq r$, then $pq \leq \frac{r^2}{4}$.*

Lemma 2.2 *For $\alpha \in [0, \frac{\pi}{2}]$,*

$$\sin(\alpha) \geq \sin^2(\alpha) = 1 - \cos^2(\alpha) \geq 1 - \cos(\alpha).$$

The following Lemma is proved in [14].

Lemma 2.3 *Let u, v, and w be real vectors of same size satisfying $u + v = w$ and $u^{\mathrm{T}} v \geq 0$. Then,*

$$2\|u\| \cdot \|v\| \leq \|u\|^2 + \|v\|^2 \leq \|u\|^2 + \|v\|^2 + 2u^{\mathrm{T}} v = \|u + v\|^2 = \|w\|^2. \tag{1}$$

The next technical lemma is from [20, page 88].

Lemma 2.4 *Let u and v be the vectors of the same dimension, and $u^{\mathrm{T}} v \geq 0$. Then*

$$\|u \circ v\| \leq 2^{-\frac{3}{2}} \|u + v\|^2.$$

We will use the famous Cardano's formula which can be found in [15].

Lemma 2.5 *Let p and q be the real numbers that are related to the following cubic algebra equation*

$$x^3 + px + q = 0.$$

If

$$\Delta = \left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3 > 0,$$

then the cubic equation has one real root that is given by

$$x = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}.$$

For quartic polynomials, the roots can also be represented by formulae, we do not list all the possible cases and solutions but refer to [10] for the detailed discussion. The last technical lemma in this section is as follows.

Lemma 2.6 *Let u and v be the n-dimensional vectors. Then*

$$\left\| u \circ v - \frac{1}{n} \left(u^{\mathrm{T}} v \right) e \right\| \leq \left\| u \circ v \right\|.$$

Proof: Omitted.

3   Constrained LQR and Convex QP with Box Constraints

We will consider the following constrained LQR (or MPC) design problem. Let $\mathbf{x} \in \mathbf{R}^r$ be the system state, and $\mathbf{u} \in \mathbf{R}^m$ be the control vector. The discrete linear time-invariant system is given by

$$\mathbf{x}_{s+1} = A\mathbf{x}_s + B\mathbf{u}_s, \tag{2}$$

while fulfilling the constraints

$$- e \leq \mathbf{u}_s \leq e, \tag{3}$$

where $s = t, \ldots, t + N - 1$. Let P, Q, and R be positive definite matrices. The design is to optimize the following cost function

$$J = \min_{\mathbf{u}_t, \mathbf{u}_{t+1}, \cdots, \mathbf{u}_{t+N-1}} \frac{1}{2}\mathbf{x}_{t+N}^{\mathrm{T}} P\mathbf{x}_{t+N} + \frac{1}{2} \sum_{k=0}^{N-1} \left[\mathbf{x}_{t+k}^{\mathrm{T}} Q\mathbf{x}_{t+k} + \mathbf{u}_{t+k}^{\mathrm{T}} R\mathbf{u}_{t+k} \right] \tag{4}$$

under the system dynamics equality constraints (2) and control saturation inequality constraints (3). Given current state $\mathbf{x}_t$, this LQR (or MPC) design problem is a typical convex quadratic programming problems with $Nr + Nm$ variables $\mathbf{x}_{t+1}, \cdots, \mathbf{x}_{t+N}, \mathbf{u}_t, \cdots, \mathbf{u}_{t+N-1}$. Though this problem can be directly solved as suggested by [4, 21], it can be significantly reduced to an equivalent but much smaller convex quadratic programming problem subject only to box constraints. Denote

$$A^k = \underbrace{A \cdots A}_{\text{product of k A}} := A_k$$

with $A_0 = I$. Since

$$\mathbf{x}_{t+k} = A\mathbf{x}_{t+k-1} + B\mathbf{u}_{t+k-1} = A^k\mathbf{x}_t + \sum_{j=0}^{k-1} A^j B\mathbf{u}_{t+k-j-1} = A_k\mathbf{x}_t + \sum_{j=0}^{k-1} A_j B\mathbf{u}_{t+k-j-1}, \tag{5}$$

5

(4) can be rewritten as

$$
\begin{aligned}
J \;=\; & \min_{\mathbf{u}_t, \mathbf{u}_{t+1}, \cdots, \mathbf{u}_{t+N-1}} \frac{1}{2} \left(A_N \mathbf{x}_t + \sum_{j=0}^{N-1} A_j B \mathbf{u}_{t+N-j-1} \right)^{\mathrm{T}} P \left(A_N \mathbf{x}_t + \sum_{j=0}^{N-1} A_j B \mathbf{u}_{t+N-j-1} \right) \\
& + \frac{1}{2} \sum_{k=1}^{N-1} \left(A_k \mathbf{x}_t + \sum_{j=0}^{k-1} A_j B \mathbf{u}_{t+k-j-1} \right)^{\mathrm{T}} Q \left(A_k \mathbf{x}_t + \sum_{j=0}^{k-1} A_j B \mathbf{u}_{t+k-j-1} \right) + \frac{1}{2} \sum_{k=0}^{N-1} \left(\mathbf{u}_{t+k}^{\mathrm{T}} R \mathbf{u}_{t+k} \right)
\end{aligned}
\tag{6}
$$

Notice that $\mathbf{x}_t$ is a constant vector, A_j, P, Q, and R are constant matrices, the cost function (6) can be reduced to

$$
\begin{aligned}
J_0 \;=\; & \min_{\mathbf{u}_t, \mathbf{u}_{t+1}, \cdots, \mathbf{u}_{t+N-1}} \frac{1}{2} \left(\sum_{j=0}^{N-1} A_j B \mathbf{u}_{t+N-j-1} \right)^{\mathrm{T}} P \left(\sum_{j=0}^{N-1} A_j B \mathbf{u}_{t+N-j-1} \right) \\
& + (A_N \mathbf{x}_t)^{\mathrm{T}} P \left(\sum_{j=0}^{N-1} A_j B \mathbf{u}_{t+N-j-1} \right) \\
& + \frac{1}{2} \sum_{k=1}^{N-1} \left(\sum_{j=0}^{k-1} A_j B \mathbf{u}_{t+k-j-1} \right)^{\mathrm{T}} Q \left(\sum_{j=0}^{k-1} A_j B \mathbf{u}_{t+k-j-1} \right) \\
& + \sum_{k=1}^{N-1} \left((A_k \mathbf{x}_t)^{\mathrm{T}} Q \left(\sum_{j=0}^{k-1} A_j B \mathbf{u}_{t+k-j-1} \right) \right) \\
& + \frac{1}{2} \sum_{k=0}^{N-1} \left(\mathbf{u}_{t+k}^{\mathrm{T}} R \mathbf{u}_{t+k} \right).
\end{aligned}
\tag{7}
$$

Denote

$$
\sum_{j=0}^{k-1} A_j B \mathbf{u}_{t+k-j-1} = \underbrace{[A_{k-1} B, A_{k-2} B, \cdots, B]}_{\phi_k} \underbrace{\begin{bmatrix} \mathbf{u}_t \\ \vdots \\ \mathbf{u}_{t+k-1} \end{bmatrix}}_{v_k} = \phi_k v_k,
\tag{8}
$$

$$
Q_k = \begin{bmatrix} \phi_k^{\mathrm{T}} Q \phi_k & 0 \\ 0 & 0 \end{bmatrix},
\tag{9}
$$

$$
R_N = \underbrace{\begin{bmatrix} R & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & R \end{bmatrix}}_{N \text{ diagonal matrices}},
\tag{10}
$$

6

and

$$S_k = \begin{bmatrix} A_k^{\mathrm{T}} Q \phi_k & 0 \end{bmatrix}, \tag{11}$$

where 0 are zero matrices with appropriate dimensions. The constrained LQR (or MPC) design is reduced further to

$$J_0 = \min_{\mathbf{u}_t, \mathbf{u}_{t+1}, \cdots, \mathbf{u}_{t+N-1}} \frac{1}{2} v_N^{\mathrm{T}} \left(\phi_N^{\mathrm{T}} P \phi_N + \sum_{k=1}^{N-1} Q_k + R_N \right) v_N + \mathbf{x}_t^{\mathrm{T}} \left(A_N^{\mathrm{T}} P \phi_N + \sum_{k=1}^{N-1} S_k \right) v_N$$
$$s.t. \quad -e \le v_N \le e. \tag{12}$$

This is a convex quadratic programming problem with Nm variables and $2Nm$ box constraints, a much smaller and simpler problem than the original one. Let $n = Nm$,

$$x = v_N, \tag{13}$$

$$H = \left(\phi_N^{\mathrm{T}} P \phi_N + \sum_{k=1}^{N-1} Q_k + R_N \right), \tag{14}$$

$$c^{\mathrm{T}} = \mathbf{x}_t^{\mathrm{T}} \left(A_N^{\mathrm{T}} P \phi_N + \sum_{k=1}^{N-1} S_k \right). \tag{15}$$

The LQR (or MPC) design problem can be written in a standard form of convex quadratic problem with box constraints:

$$(QP) \quad \min \tfrac{1}{2} x^{\mathrm{T}} H x + c^{\mathrm{T}} x, \quad \text{subject to} \quad -e \le x \le e, \tag{16}$$

where $0 < H \in \mathbf{R}^{n \times n}$ is a positive definite matrix, $c \in \mathbf{R}^n$ is given, and $x \in \mathbf{R}^n$ is the control vector to be optimized. The remaining discussion of this paper is focused on the solution to the convex quadratic programming problem with box constraints described by (16).

4 Central Path of Convex QP with Box Constraints

It is well known that x is an optimal solution of (16) if and only if x, λ, and γ meet the following KKT conditions

$$-\lambda + \gamma - Hx = c, \tag{17a}$$

$$-e \le x \le e, \tag{17b}$$

$$(\lambda, \gamma) \ge 0, \tag{17c}$$

$$\lambda_i(e_i - x_i) = 0, \quad \gamma_i(e_i + x_i) = 0, \quad i = 1, \ldots, n. \tag{17d}$$

Denote $y = e - x \geq 0$, $z = e + x \geq 0$. The KKT condition can be rewritten as

$$Hx + c + \lambda - \gamma = 0, \tag{18a}$$

$$x + y = e, \quad x - z = -e, \tag{18b}$$

$$(y, z, \lambda, \gamma) \geq 0, \tag{18c}$$

$$\lambda_i y_i = 0, \quad \gamma_i z_i = 0, \quad i = 1, \ldots, n. \tag{18d}$$

For the convex (QP) problem, the KKT condition is also sufficient for x to be a global optimal solution. Denote the feasible set $\mathcal{F}$ as a collection of all points that meet the constraints (18a), (18b), (18c)

$$\mathcal{F} = \{(x, y, z, \lambda, \gamma) : Hx + c + \lambda - \gamma = 0, (y, z, \lambda, \gamma) \geq 0, x + y = e, x - z = -e\}, \tag{19}$$

and the strictly feasible set $\mathcal{F}^o$ as a collection of all points that meet the constraints (18a), (18b), and are strictly positive in (18c)

$$\mathcal{F}^o = \{(x, y, z, \lambda, \gamma) : Hx + c + \lambda - \gamma = 0, (y, z, \lambda, \gamma) > 0, x + y = e, x - z = -e\}. \tag{20}$$

Similar to the linear programming, we define the central path $\mathcal{C} \in \mathcal{F}^o \subset \mathcal{F}$, as a curve in finite dimensional space parameterized by a scalar $\tau > 0$ as follows. For each interior point $(x, y, z, \lambda, \gamma) \in \mathcal{F}^o$ on the central path, there is a $\tau > 0$ such that

$$Hx + c + \lambda - \gamma = 0, \tag{21a}$$

$$x + y = e, \quad x - z = -e, \tag{21b}$$

$$(y, z, \lambda, \gamma) > 0, \tag{21c}$$

$$\lambda_i y_i = \tau, \quad \gamma_i z_i = \tau, \quad i = 1, \ldots, n. \tag{21d}$$

Therefore, the central path is an arc that is parameterized as a function of τ and is denoted as

$$\mathcal{C} = \{(x(\tau), y(\tau), z(\tau), \lambda(\tau), \gamma(\tau)) : \tau > 0\}. \tag{22}$$

As $\tau \to 0$, the moving point $(x(\tau), y(\tau), z(\tau), \lambda(\tau), \gamma(\tau))$ on the central path represented by (21) approaches the solution of (QP) represented by (16). Throughout the paper, we make the following assumption.

Assumption:

1. $\mathcal{F}^o$ is not empty.

Assumption 1 implies the existence of a central path. This assumption is always true for the LQR problem, and we will provide an explicit initial interior point in Section 7.

Let $1 > \theta > 0$, denote $p = (y, z)$, $\omega = (\lambda, \gamma)$, and the duality gap

$$\mu = \frac{\lambda^{\mathrm{T}} y + \gamma^{\mathrm{T}} z}{2n} = \frac{p^{\mathrm{T}} \omega}{2n}. \tag{23}$$

We define a set of neighborhood of the central path as

$$\mathcal{N}_2(\theta) = \{(x, y, z, \lambda, \gamma) \in \mathcal{F}^o : \|p \circ \omega - \mu e\| \le \theta\mu, \} \subset \mathcal{F}^o. \tag{24}$$

As we reduce the duality gap to zero, the neighborhood of $\mathcal{N}_2(\theta)$ will be a neighborhood of the central path that approaches the optimizer(s) of the QP problem, therefore, all points inside $\mathcal{N}_2(\theta)$ will approach the optimizer(s) of the QP problem. For $(x, y, z, \lambda, \gamma) \in \mathcal{N}_2(\theta)$, since $(1 - \theta)\mu \le \omega_i p_i \le (1 + \theta)\mu$, where ω_i are either λ_i or γ_i, and p_i are either y_i or z_i, we have

$$\frac{\omega_i p_i}{1 + \theta} \le \frac{\max_i \omega_i p_i}{1 + \theta} \le \mu \le \frac{\min_i \omega_i p_i}{1 - \theta} \le \frac{\omega_i p_i}{1 - \theta}. \tag{25}$$

5 An Arc-search Algorithm for Convex QP with Box Constraints

The idea of arc-search proposed in this paper is very simple. The algorithm starts from a feasible point in $\mathcal{N}_2(\theta)$ close to the central path, constructs an arc that passes through the point and approximates the central path, searches along the arc to a new point in a larger area $\mathcal{N}_2(2\theta)$ that reduces the duality gap $p^{\mathrm{T}} \omega$ and meets (21a), (21b), and (21c). The process is repeated by finding a better point close to the central path or on the central path in $\mathcal{N}_2(\theta)$ that simultaneously meets (21a), (21b), and (21c).

Following the idea used in [24], we will use an ellipse $\mathcal{E}$ [6] in an appropriate dimensional space to approximate the central path $\mathcal{C}$ described by (21), where

$$\mathcal{E} = \{(x(\alpha), y(\alpha), z(\alpha), \lambda(\alpha), \gamma(\alpha)) : (x(\alpha), y(\alpha), z(\alpha), \lambda(\alpha), \gamma(\alpha)) = \vec{a} \cos(\alpha) + \vec{b} \sin(\alpha) + \vec{c}\}, \tag{26}$$

$\vec{a} \in \mathbf{R}^{5n}$ and $\vec{b} \in \mathbf{R}^{5n}$ are the axes of the ellipse, $\vec{c} \in \mathbf{R}^{5n}$ is the center of the ellipse. Given a point $(x, y, z, \lambda, \gamma) = (x(\alpha_0), y(\alpha_0), z(\alpha_0), \lambda(\alpha_0), \gamma(\alpha_0)) \in \mathcal{E}$ which is close to or on the central path, $\vec{a}$, $\vec{b}$, $\vec{c}$ are functions of α, $(x, \lambda, \gamma, y, z)$, $(\dot{x}, \dot{y}, \dot{z}, \dot{\lambda}, \dot{\gamma})$, and $(\ddot{x}, \ddot{y}, \ddot{z}, \ddot{\lambda}, \ddot{\gamma})$, where $(\dot{x}, \dot{y}, \dot{z}, \dot{\lambda}, \dot{\gamma})$ and $(\ddot{x}, \ddot{y}, \ddot{z}, \ddot{\lambda}, \ddot{\gamma})$ are defined as

9

$$
\begin{bmatrix} H & 0 & 0 & I & -I \\ I & I & 0 & 0 & 0 \\ I & 0 & -I & 0 & 0 \\ 0 & \Lambda & 0 & Y & 0 \\ 0 & 0 & \Gamma & 0 & Z \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ \dot{\lambda} \\ \dot{\gamma} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \lambda \circ y \\ \gamma \circ z \end{bmatrix}, \tag{27}
$$

$$
\begin{bmatrix} H & 0 & 0 & I & -I \\ I & I & 0 & 0 & 0 \\ I & 0 & -I & 0 & 0 \\ 0 & \Lambda & 0 & Y & 0 \\ 0 & 0 & \Gamma & 0 & Z \end{bmatrix} \begin{bmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \\ \ddot{\lambda} \\ \ddot{\gamma} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -2\dot{\lambda} \circ \dot{y} \\ -2\dot{\gamma} \circ \dot{z} \end{bmatrix}. \tag{28}
$$

The first rows of (27) and (28) are equivalent to

$$
H\dot{x} = \dot{\gamma} - \dot{\lambda}, \qquad H\ddot{x} = \ddot{\gamma} - \ddot{\lambda}. \tag{29}
$$

The next 2 rows of (27) and (28) are equivalent to

$$
\dot{x} = -\dot{y}, \qquad \dot{x} = \dot{z}, \qquad \ddot{x} = -\ddot{y}, \qquad \ddot{x} = \ddot{z}. \tag{30}
$$

The last 2 rows of (27) and (28) are equivalent to

$$
p \circ \dot{\omega} + \dot{p} \circ \omega = p \circ \omega, \tag{31}
$$

$$
p \circ \ddot{\omega} + \ddot{p} \circ \omega = -2\dot{p} \circ \dot{\omega}. \tag{32}
$$

It has been shown in [22] that one can avoid the calculation of $\vec{a}$, $\vec{b}$, and $\vec{c}$ in the expression of the ellipse. The following formulas are used instead.

Theorem 5.1 *Let $(x(\alpha), y(\alpha), z(\alpha), \lambda(\alpha), \gamma(\alpha))$ be an arc defined by (26) passing through a point $(x, y, z, \lambda, \gamma) \in \mathcal{E}$, and its first and second derivatives at $(x, y, z, \lambda, \gamma)$ be $(\dot{x}, \dot{y}, \dot{z}, \dot{\lambda}, \dot{\gamma})$ and $(\ddot{x}, \ddot{y}, \ddot{z}, \ddot{\lambda}, \ddot{\gamma})$ which are defined by (27) and (28). Then an ellipse approximation of the central path is given by*

$$
x(\alpha) = x - \dot{x}\sin(\alpha) + \ddot{x}(1 - \cos(\alpha)), \tag{33}
$$

$$
y(\alpha) = y - \dot{y}\sin(\alpha) + \ddot{y}(1 - \cos(\alpha)), \tag{34}
$$

$$
z(\alpha) = z - \dot{z}\sin(\alpha) + \ddot{z}(1 - \cos(\alpha)), \tag{35}
$$

$$
\lambda(\alpha) = \lambda - \dot{\lambda}\sin(\alpha) + \ddot{\lambda}(1 - \cos(\alpha)), \tag{36}
$$

10

$$\gamma(\alpha) = \gamma - \dot{\gamma}\sin(\alpha) + \ddot{\gamma}(1 - \cos(\alpha)). \tag{37}$$

■

We will also use a compact format for $p(\alpha) = (y(\alpha), z(\alpha))$ and $\omega(\alpha) = (\lambda(\alpha), \gamma(\alpha))$, which are given by

$$p(\alpha) = p - \dot{p}\sin(\alpha) + \ddot{p}(1 - \cos(\alpha)), \tag{38}$$

$$\omega(\alpha) = \omega - \dot{\omega}\sin(\alpha) + \ddot{\omega}(1 - \cos(\alpha)). \tag{39}$$

We denote the duality gap at point $(x(\alpha), p(\alpha), \omega(\alpha))$ as

$$\mu(\alpha) = \frac{\lambda(\alpha)^{\mathrm{T}}y(\alpha) + \gamma(\alpha)^{\mathrm{T}}z(\alpha)}{2n} = \frac{p(\alpha)^{\mathrm{T}}\omega(\alpha)}{2n}. \tag{40}$$

Assuming $(y, z, \lambda, \gamma) > 0$, one can easily see that if $\frac{\dot{y}}{y}$, $\frac{\dot{z}}{z}$, $\frac{\dot{\lambda}}{\lambda}$, $\frac{\dot{\gamma}}{\gamma}$, $\frac{\ddot{y}}{y}$, $\frac{\ddot{z}}{z}$, $\frac{\ddot{\lambda}}{\lambda}$, $\frac{\ddot{\gamma}}{\gamma}$ are bounded (we will show that this is true), and if α is small enough, then $y(\alpha) > 0$, $z(\alpha) > 0$, $\lambda(\alpha) > 0$, and $\gamma(\alpha) > 0$. We will also show that searching along this ellipse will reduce the duality gap, i.e., $\mu(\alpha) < \mu$.

Lemma 5.1 *Let $(x, y, z, \lambda, \gamma)$ be a strictly feasible point of (QP), $(\dot{x}, \dot{y}, \dot{z}, \dot{\lambda}, \dot{\gamma})$ and $(\ddot{x}, \ddot{y}, \ddot{z}, \ddot{\lambda}, \ddot{\gamma})$ meet (27) and (28), $(x(\alpha), y(\alpha), z(\alpha), \lambda(\alpha), \gamma(\alpha))$ be calculated using (33), (34), (35), (36), and (37), then the following conditions hold.*

$$x(\alpha) + y(\alpha) = e, \quad x(\alpha) - z(\alpha) = -e, \quad Hx(\alpha) + c + \lambda(\alpha) + \gamma(\alpha) = 0.$$

Proof: Since $(x, y, z, \lambda, \gamma)$ is a strictly feasible point, the result follows from direct calculation by using (20), (27), (28), and Theorem 5.1. ■

Lemma 5.2 *Let $(\dot{x}, \dot{p}, \dot{\omega})$ be defined by (27), $(\ddot{x}, \ddot{p}, \ddot{\omega})$ be defined by (28), and H be positive definite matrix. Then the following relations hold.*

$$\dot{p}^{\mathrm{T}}\dot{\omega} = \dot{x}^{\mathrm{T}}(\dot{\gamma} - \dot{\lambda}) = \dot{x}^{\mathrm{T}}H\dot{x} \geq 0, \tag{41}$$

The equality holds if and only if $\|\dot{x}\| = 0$.

$$\ddot{p}^{\mathrm{T}}\ddot{\omega} = \ddot{x}^{\mathrm{T}}(\ddot{\gamma} - \ddot{\lambda}) = \ddot{x}^{\mathrm{T}}H\ddot{x} \geq 0, \tag{42}$$

The equality holds if and only if $\|\ddot{x}\| = 0$.

$$\ddot{p}^{\mathrm{T}}\dot{\omega} = \ddot{x}^{\mathrm{T}}(\dot{\gamma} - \dot{\lambda}) = \dot{x}^{\mathrm{T}}(\ddot{\gamma} - \ddot{\lambda}) = \dot{p}^{\mathrm{T}}\ddot{\omega} = \dot{x}^{\mathrm{T}}H\ddot{x}. \tag{43}$$

$$-(\dot{x}^{\mathrm{T}}H\dot{x})(1-\cos(\alpha))^2 - (\ddot{x}^{\mathrm{T}}H\ddot{x})\sin^2(\alpha)$$

$$\leq \quad (\ddot{x}^{\mathrm{T}}(\dot{\gamma}-\dot{\lambda}) + \dot{x}^{\mathrm{T}}(\ddot{\gamma}-\ddot{\lambda}))\sin(\alpha)(1-\cos(\alpha))$$

$$\leq \quad (\dot{x}^{\mathrm{T}}H\dot{x})(1-\cos(\alpha))^2 + (\ddot{x}^{\mathrm{T}}H\ddot{x})\sin^2(\alpha). \tag{44}$$

$$-(\dot{x}^{\mathrm{T}}H\dot{x})\sin^2(\alpha) - (\ddot{x}^{\mathrm{T}}H\ddot{x})(1-\cos(\alpha))^2$$

$$\leq \quad (\ddot{x}^{\mathrm{T}}(\dot{\gamma}-\dot{\lambda}) + \dot{x}^{\mathrm{T}}(\ddot{\gamma}-\ddot{\lambda}))\sin(\alpha)(1-\cos(\alpha))$$

$$\leq \quad (\dot{x}^{\mathrm{T}}H\dot{x})\sin^2(\alpha) + (\ddot{x}^{\mathrm{T}}H\ddot{x})(1-\cos(\alpha))^2. \tag{45}$$

For $\alpha = \frac{\pi}{2}$, (44) and (45) reduce to

$$-\left(\dot{x}^{\mathrm{T}}H\dot{x} + \ddot{x}^{\mathrm{T}}H\ddot{x}\right) \leq (\ddot{x}^{\mathrm{T}}H\dot{x} + \dot{x}^{\mathrm{T}}H\ddot{x}) \leq \dot{x}^{\mathrm{T}}H\dot{x} + \ddot{x}^{\mathrm{T}}H\ddot{x}. \tag{46}$$

Proof: See Appendix A.

Using Lemmas 5.2, 2.1, and 2.3, we can show that $\frac{\dot{p}}{p} := \left(\frac{\dot{y}}{y}, \frac{\dot{z}}{z}\right)$, $\frac{\dot{\omega}}{\omega} := \left(\frac{\dot{\lambda}}{\lambda}, \frac{\dot{\gamma}}{\gamma}\right)$, $\frac{\ddot{p}}{p} := \left(\frac{\ddot{y}}{y}, \frac{\ddot{z}}{z}\right)$ and $\frac{\ddot{\omega}}{\omega} := \left(\frac{\ddot{\lambda}}{\lambda}, \frac{\ddot{\gamma}}{\gamma}\right)$ are all bounded as claimed in the following two Lemmas.

Lemma 5.3 *Let $(x, p, \omega) = (x, y, z, \lambda, \gamma) \in \mathcal{N}_2(\theta)$ and $(\dot{x}, \dot{p}, \dot{\omega}) = (\dot{x}, \dot{y}, \dot{z}, \dot{\lambda}, \dot{\gamma})$ meet (27). Then,*

$$\left\|\frac{\dot{p}}{p}\right\|^2 + \left\|\frac{\dot{\omega}}{\omega}\right\|^2 \leq \frac{2n}{1-\theta}, \tag{47}$$

$$\left\|\frac{\dot{p}}{p}\right\|^2 \left\|\frac{\dot{\omega}}{\omega}\right\|^2 \leq \left(\frac{n}{1-\theta}\right)^2, \tag{48}$$

$$0 \leq \frac{\dot{p}^{\mathrm{T}}\dot{\omega}}{\mu} \leq \frac{1+\theta}{1-\theta}n := \delta_1 n. \tag{49}$$

Proof: See Appendix A.

Lemma 5.4 *Let $(x, p, \omega) = (x, y, z, \lambda, \gamma) \in \mathcal{N}_2(\theta)$, $(\dot{x}, \dot{y}, \dot{z}, \dot{\lambda}, \dot{\gamma})$ and $(\ddot{x}, \ddot{y}, \ddot{z}, \ddot{\lambda}, \ddot{\gamma})$ meet (27) and (28). Then*

$$\left\|\frac{\ddot{p}}{p}\right\|^2 + \left\|\frac{\ddot{\omega}}{\omega}\right\|^2 \leq \frac{4(1+\theta)n^2}{(1-\theta)^3}, \tag{50}$$

$$\left\|\frac{\ddot{p}}{p}\right\|^2 \left\|\frac{\ddot{\omega}}{\omega}\right\|^2 \leq \left(\frac{2(1+\theta)n^2}{(1-\theta)^3}\right)^2, \tag{51}$$

$$0 \leq \frac{\ddot{p}^{\mathrm{T}}\ddot{\omega}}{\mu} \leq \frac{2(1+\theta)^2}{(1-\theta)^3}n^2 := \delta_2 n^2, \tag{52}$$

$$\left|\frac{\dot{p}^{\mathrm{T}}\ddot{\omega}}{\mu}\right| \leq \frac{(2n(1+\theta))^{\frac{3}{2}}}{(1-\theta)^2} := \delta_3 n^{\frac{3}{2}}, \quad \left|\frac{\ddot{p}^{\mathrm{T}}\dot{\omega}}{\mu}\right| \leq \frac{(2n(1+\theta))^{\frac{3}{2}}}{(1-\theta)^2} := \delta_3 n^{\frac{3}{2}}. \tag{53}$$

Proof: See Appendix A.

Using the bounds established in Lemmas 5.2, 5.3, 5.4, and 2.2, we can obtain the lower bound and upper bound for $\mu(\alpha)$.

Lemma 5.5 Let $(x, p, \omega) = (x, y, z, \lambda, \gamma) \in \mathcal{N}_2(\theta)$, $(\dot{x}, \dot{y}, \dot{z}, \dot{\lambda}, \dot{\gamma})$ and $(\ddot{x}, \ddot{y}, \ddot{z}, \ddot{\lambda}, \ddot{\gamma})$ meet (27) and (28). Let $x(\alpha)$, $y(\alpha)$, $z(\alpha)$, $\lambda(\alpha)$, and $\gamma(\alpha)$ be defined by (33), (34), (35), (36), and (37). Then,

$$\mu(1 - \sin(\alpha)) - \frac{1}{2n}\dot{x}^{\mathrm{T}}H\dot{x}\left((1 - \cos(\alpha))^2 + \sin^2(\alpha)\right)$$

$$\leq \mu(\alpha) = \mu(1 - \sin(\alpha)) + \frac{1}{2n}\left(\ddot{x}^{\mathrm{T}}(\dot{\gamma} - \ddot{\lambda}) - \dot{x}^{\mathrm{T}}(\dot{\gamma} - \dot{\lambda})\right)(1 - \cos(\alpha))^2$$

$$- \frac{1}{2n}\left(\dot{x}^{\mathrm{T}}(\ddot{\gamma} - \ddot{\lambda}) + \ddot{x}^{\mathrm{T}}(\dot{\gamma} - \dot{\lambda})\right)\sin(\alpha)(1 - \cos(\alpha))$$

$$\leq \mu(1 - \sin(\alpha)) + \frac{1}{2n}\ddot{x}^{\mathrm{T}}H\ddot{x}\left((1 - \cos(\alpha))^2 + \sin^2(\alpha)\right). \tag{54}$$

Proof: See Appendix A.

To keep all the iterates of the algorithm inside the strictly feasible set, we need $(p(\alpha), \omega(\alpha)) > 0$ for all iterations. We will prove that this is guaranteed if $\mu(\alpha) > 0$ holds. The following corollary states the condition for $\mu(\alpha) > 0$ to hold.

Corollary 5.1 If $\mu > 0$, then for any fixed $\theta \in (0, 1)$, there is an $\bar{\alpha} > 0$ depending on θ, such that for any $\sin(\alpha) \leq \sin(\bar{\alpha})$, $\mu(\alpha) > 0$. In particular, if $\theta = 0.19$, $\sin(\bar{\alpha}) \geq 0.6158$.

Proof: From Lemmas 5.2 and 2.2, we have $\dot{x}^{\mathrm{T}}H\dot{x}^{\mathrm{T}} = \dot{x}^{\mathrm{T}}(\dot{\gamma} - \dot{\lambda}) = \dot{p}^{\mathrm{T}}\dot{\omega}$ and $((1 - \cos(\alpha))^2 \leq \sin^4(\alpha)$. Therefore, from Lemmas 5.5 and 5.3, we have

$$\mu(\alpha) \geq \mu\left(1 - \sin(\alpha) - \frac{1}{2n\mu}\dot{p}^{\mathrm{T}}\dot{\omega}\left(\sin^4(\alpha) + \sin^2(\alpha)\right)\right)$$

$$\geq \mu\left(1 - \sin(\alpha) - \frac{(1 + \theta)}{2(1 - \theta)}\left(\sin^4(\alpha) + \sin^2(\alpha)\right)\right) := \mu r(\alpha).$$

Since $\mu > 0$, and $r(\alpha)$ is a monotonic decreasing function in $[0, \frac{\pi}{2}]$ with $r(0) > 0$, $r(\frac{\pi}{2}) < 0$, there is a unique real solution $\sin(\bar{\alpha}) \in (0, 1)$ of $r(\alpha) = 0$ such that for all $\sin(\alpha) < \sin(\bar{\alpha})$, $r(\alpha) > 0$, or $\mu(\alpha) > 0$. It is easy to check that if $\theta = 0.19$, $\sin(\bar{\alpha}) = 0.6158$ is the solution of $r(\alpha) = 0$. ∎

Remark 5.1 Corollary 5.1 indicates that for any $\theta \in (0, 1)$, there is a positive $\bar{\alpha}$ such that for $\alpha \leq \bar{\alpha}$, $\mu(\alpha) > 0$. Intuitively, to search in a wider region will generate a longer step. Therefore, the larger the θ is, the better. But to derive the convergence result, $\theta \leq 0.22$ is imposed in Lemma 5.9 and $\theta \leq 0.19$ is imposed in Lemma 5.13.

To reduce the duality gap in an iteration, we need to have $\mu(\alpha) \le \mu$. For linear programming, it is known [22] that $\mu(\alpha) \le \mu$ for $\alpha \in [0, \hat{\alpha}]$ with $\hat{\alpha} = \frac{\pi}{2}$, and the larger the α in the interval is, the smaller the $\mu(\alpha)$ will be. This claim is not true for the convex quadratic programming with box constraints and it needs to be modified as follows.

Lemma 5.6 Let $(x, p, \omega) = (x, y, z, \lambda, \gamma) \in \mathcal{N}_2(\theta)$, $(\dot{x}, \dot{y}, \dot{z}, \dot{\lambda}, \dot{\gamma})$ and $(\ddot{x}, \ddot{y}, \ddot{z}, \ddot{\lambda}, \ddot{\gamma})$ meet (27) and (28). Let $x(\alpha)$, $y(\alpha)$, $z(\alpha)$, $\lambda(\alpha)$, and $\gamma(\alpha)$ be defined by (33), (34), (35), (36), and (37). Then, there exists

$$\hat{\alpha} = \begin{cases} \frac{\pi}{2}, & \text{if } \frac{\ddot{x}^{\mathrm{T}} H \ddot{x}}{n\mu} \le 1 \\ \\ \sin^{-1}(g), & \text{if } \frac{\ddot{x}^{\mathrm{T}} H \ddot{x}}{n\mu} > 1 \end{cases} \tag{55}$$

where

$$g = \sqrt[3]{\frac{n\mu}{\ddot{x}^{\mathrm{T}} H \ddot{x}} + \sqrt{\left(\frac{n\mu}{\ddot{x}^{\mathrm{T}} H \ddot{x}}\right)^2 + \left(\frac{1}{3}\right)^3}} + \sqrt[3]{\frac{n\mu}{\ddot{x}^{\mathrm{T}} H \ddot{x}} - \sqrt{\left(\frac{n\mu}{\ddot{x}^{\mathrm{T}} H \ddot{x}}\right)^2 + \left(\frac{1}{3}\right)^3}},$$

such that for every $\alpha \in [0, \hat{\alpha}]$, $\mu(\alpha) \le \mu$.

Proof: See Appendix A.

According to Theorem 5.1, Lemmas 5.1, 5.3, 5.4, and 5.6, if α is small enough, then $(p(\alpha), \omega(\alpha)) > 0$, and $\mu(\alpha) < \mu$, i.e., the search along the ellipse defined by Theorem 5.1 will generate a strictly feasible point with a smaller duality gap. Since $(p, \omega) > 0$ holds in all iterations, reducing the duality gap to zero means approaching the solution of the convex quadratic programming. We will apply a similar idea used in [13, 23], i.e., starting with an iterate in $\mathcal{N}_2(\theta)$, searching along the approximated central path to reduce the duality gap and to keep the iterate in $\mathcal{N}_2(2\theta)$, and then making a correction to move the iterate back to $\mathcal{N}_2(\theta)$. First, we will introduce the following notations.

$$a_0 = -\theta\mu < 0,$$

$$a_1 = \theta\mu > 0,$$

$$a_2 = 2\theta\frac{\dot{p}^{\mathrm{T}}\dot{\omega}}{2n} = 2\theta\frac{\dot{x}^{\mathrm{T}}(\dot{\gamma} - \dot{\lambda})}{2n} = 2\theta\frac{\dot{x}^{\mathrm{T}} H \dot{x}}{2n} \ge 0,$$

$$a_3 = \left\| \dot{p} \circ \ddot{\omega} + \dot{\omega} \circ \ddot{p} - \frac{1}{2n}(\dot{p}^{\mathrm{T}}\ddot{\omega} + \dot{\omega}^{\mathrm{T}}\ddot{p})e \right\| \ge 0,$$

$$a_4 = \left\| \ddot{p} \circ \ddot{\omega} - \dot{\omega} \circ \dot{p} - \frac{1}{2n}(\ddot{p}^{\mathrm{T}}\ddot{\omega} - \dot{\omega}^{\mathrm{T}}\dot{p})e \right\| + 2\theta\frac{\dot{p}^{\mathrm{T}}\dot{\omega}}{2n}$$
$$= \left\| \ddot{p} \circ \ddot{\omega} - \dot{\omega} \circ \dot{p} - \frac{1}{2n}(\ddot{p}^{\mathrm{T}}\ddot{\omega} - \dot{\omega}^{\mathrm{T}}\dot{p})e \right\| + 2\theta\frac{\dot{x}^{\mathrm{T}} H \dot{x}}{2n} \ge 0.$$

14

We also define a quartic polynomial in terms of $\sin(\alpha)$ as follows

$$q(\alpha) = a_4 \sin^4(\alpha) + a_3 \sin^3(\alpha) + a_2 \sin^2(\alpha) + a_1 \sin(\alpha) + a_0 = 0. \tag{56}$$

Since $q(\alpha)$ is a monotonic increasing function of $\alpha \in [0, \frac{\pi}{2}]$, $q(0) = -\theta\mu < 0$ and $q(\frac{\pi}{2}) = a_2 + a_3 + a_4 > 0$ if $\dot{x} \neq 0$, the polynomial has exactly one positive root in $[0, \frac{\pi}{2}]$. Moreover, since (56) is a quartic equation, all the solutions are analytical and the computational cost is independent of the size of H and negligible [10].

Lemma 5.7 *Let* $(x, p, \omega) = (x, y, z, \lambda, \omega) \in \mathcal{N}_2(\theta)$, $(\dot{x}, \dot{y}, \dot{z}, \dot{\lambda}, \dot{\omega})$ *and* $(\ddot{x}, \ddot{y}, \ddot{z}, \ddot{\lambda}, \ddot{\omega})$ *be calculated from (27) and (28). Denote* $\sin(\tilde{\alpha})$ *be the only positive real solution of (56) in* $[0, 1]$. *Assume* $\sin(\alpha) \leq \min\{\sin(\tilde{\alpha}), \sin(\bar{\alpha})\}$, *let* $(x(\alpha), y(\alpha), z(\alpha), \lambda(\alpha), \gamma(\alpha))$ *and* $\mu(\alpha)$ *be updated as follows*

$$(x(\alpha), y(\alpha), z(\alpha), \lambda(\alpha), \gamma(\alpha)) = (x, y, z, \lambda, \gamma) - (\dot{x}, \dot{y}, \dot{z}, \dot{\lambda}, \dot{\gamma})\sin(\alpha) + (\ddot{x}, \ddot{y}, \ddot{z}, \ddot{\lambda}, \ddot{\gamma})(1 - \cos(\alpha)), \tag{57}$$

$$\mu(\alpha) = \mu(1 - \sin(\alpha)) + \frac{1}{2n}\left((\ddot{p}^{\mathrm{T}}\ddot{\omega} - \dot{p}^{\mathrm{T}}\dot{\omega})(1 - \cos(\alpha))^2 - (\dot{p}^{\mathrm{T}}\ddot{\omega} + \ddot{p}^{\mathrm{T}}\dot{\omega})\sin(\alpha)(1 - \cos(\alpha))\right). \tag{58}$$

Then $(x(\alpha), y(\alpha), z(\alpha), \lambda(\alpha), \gamma(\alpha)) \in \mathcal{N}_2(2\theta)$.

Proof: See Appendix A.

Remark 5.2 *It is worthwhile to note, by examining the proof of Lemma 5.7, that* $\sin(\tilde{\alpha})$ *is selected for the proximity condition (103) to hold, and* $\sin(\bar{\alpha})$ *is selected for* $\mu(\alpha) > 0$, *thereby assuring the positivity condition (104) to hold.*

The lower bound of $\sin(\bar{\alpha})$ is estimated in Corollary 5.1. To estimate the lower bound of $\sin(\tilde{\alpha})$, we need the following lemma.

Lemma 5.8 *Let* $(x, p, \omega) \in \mathcal{N}_2(\theta)$, $(\dot{x}, \dot{p}, \dot{\omega})$ *and* $(\ddot{x}, \ddot{p}, \ddot{\omega})$ *meet (27) and (28). Then*

$$\left\|\dot{p} \circ \dot{\omega}\right\| \leq \frac{(1 + \theta)}{(1 - \theta)}n\mu, \tag{59}$$

$$\left\|\ddot{p} \circ \ddot{\omega}\right\| \leq \frac{2(1 + \theta)^2}{(1 - \theta)^3}n^2\mu, \tag{60}$$

$$\left\|\ddot{p} \circ \dot{\omega}\right\| \leq \frac{2\sqrt{2}(1 + \theta)^{\frac{3}{2}}}{(1 - \theta)^2}n^{\frac{3}{2}}\mu, \tag{61}$$

$$\left\|\dot{p} \circ \ddot{\omega}\right\| \leq \frac{2\sqrt{2}(1 + \theta)^{\frac{3}{2}}}{(1 - \theta)^2}n^{\frac{3}{2}}\mu. \tag{62}$$

15

Proof: See Appendix A.

Lemma 5.9 *Let $\theta \leq 0.22$. Then $\sin(\tilde{\alpha}) \geq \frac{\theta}{\sqrt{n}}$.*

Proof: See Appendix A.

Corollary 5.1, Lemmas 5.7, and 5.9 prove the feasibility of searching optimizer along the ellipse. To move the iterate back to $\mathcal{N}_2(\theta)$, we use the direction $(\Delta x, \Delta y, \Delta z, \Delta \lambda, \Delta \gamma)$ defined by

$$
\begin{bmatrix}
H & 0 & 0 & I & -I \\
I & I & 0 & 0 & 0 \\
I & 0 & -I & 0 & 0 \\
0 & \Lambda(\alpha) & 0 & Y(\alpha) & 0 \\
0 & 0 & \Gamma(\alpha) & 0 & Z(\alpha)
\end{bmatrix}
\begin{bmatrix}
\Delta x \\
\Delta y \\
\Delta z \\
\Delta \lambda \\
\Delta \gamma
\end{bmatrix}
=
\begin{bmatrix}
0 \\
0 \\
0 \\
\mu(\alpha)e - \lambda(\alpha) \circ y(\alpha) \\
\mu(\alpha)e - \gamma(\alpha) \circ z(\alpha)
\end{bmatrix}.
\tag{63}
$$

and we update $(x^{k+1}, p^{k+1}, \omega^{k+1})$ and μ^{k+1} by

$$
(x^{k+1}, p^{k+1}, \omega^{k+1}) = (x(\alpha), p(\alpha), \omega(\alpha)) + (\Delta x, \Delta p, \Delta \omega),
\tag{64}
$$

$$
\mu^{k+1} = \frac{p^{k+1^\mathrm{T}} \omega^{k+1}}{2n},
\tag{65}
$$

where $\Delta p = (\Delta y, \Delta z)$ and $\Delta \omega = (\Delta \lambda, \Delta \gamma)$. Denote $P(\alpha) = \begin{bmatrix} Y(\alpha) & 0 \\ 0 & Z(\alpha) \end{bmatrix}$, $\Omega(\alpha) = \begin{bmatrix} \Lambda(\alpha) & 0 \\ 0 & \Gamma(\alpha) \end{bmatrix}$,

and $D = P^{\frac{1}{2}}(\alpha)\Omega^{-\frac{1}{2}}(\alpha)$. Then, the last 2 rows of (63) can be rewritten as

$$
P\Delta\omega + \Omega\Delta p = u(\alpha)e - P(\alpha)\Omega(\alpha)e.
\tag{66}
$$

Now, we show that the correction step brings the iterate from $\mathcal{N}_2(2\theta)$ back to $\mathcal{N}_2(\theta)$.

Lemma 5.10 *Let $(x(\alpha), p(\alpha), \omega(\alpha)) \in \mathcal{N}_2(2\theta)$ and $(\Delta x, \Delta p, \Delta \omega)$ be defined as in (63). Let $(x^{k+1}, p^{k+1}, \omega^{k+1})$ be updated by using (64). Then, for $\theta \leq 0.29$ and $\sin(\alpha) \leq \sin(\bar{\alpha})$, $(x^{k+1}, p^{k+1}, \omega^{k+1}) \in \mathcal{N}_2(\theta)$.*

Proof: See Appendix A.

Next, we show that the combined step (searching along the arc in $\mathcal{N}_2(2\theta)$ and moving back to $\mathcal{N}_2(\theta)$) will reduce the duality gap of the iterate, i.e., $\mu^{k+1} < \mu^k$, if we select some appropriate θ and α. We introduce the following two Lemmas before we prove this result.

Lemma 5.11 *Let $(x(\alpha), p(\alpha), \omega(\alpha)) \in \mathcal{N}_2(2\theta)$ and $(\Delta x, \Delta p, \Delta \omega)$ be defined as in (63). Then*

$$
0 \leq \frac{\Delta p^\mathrm{T} \Delta \omega}{2n} \leq \frac{\theta^2(1 + 2\theta)}{n(1 - 2\theta)^2}\mu(\alpha) := \frac{\delta_0}{n}\mu(\alpha).
\tag{67}
$$

Proof: See Appendix A.

Lemma 5.12 *Let $(x(\alpha), p(\alpha), \omega(\alpha)) \in \mathcal{N}_2(2\theta)$ and $(\Delta x, \Delta p, \Delta \omega)$ be defined as in (63). Let $(x^{k+1}, p^{k+1}, \omega^{k+1})$ be defined as in (64). Then*

$$\mu(\alpha) \leq \mu^{k+1} := \frac{{p^{k+1}}^{\mathrm{T}} \omega^{k+1}}{2n} \leq \mu(\alpha) \left(1 + \frac{\theta^2(1 + 2\theta)}{n(1 - 2\theta)^2} \right) = \mu(\alpha) \left(1 + \frac{\delta_0}{n} \right)$$

Proof: Using the fact that $p(\alpha)^{\mathrm{T}} \Delta \omega + \omega(\alpha)^{\mathrm{T}} \Delta p = 0$ established in (110) in the proof of Lemma 5.10, and Lemma 5.11, it is therefore straightforward to obtain

$$\mu(\alpha) \leq \frac{p(\alpha)^{\mathrm{T}} \omega(\alpha)}{2n} + \frac{1}{2n} \Delta p^{\mathrm{T}} \Delta \omega = \frac{(p(\alpha) + \Delta p)^{\mathrm{T}} (\omega(\alpha) + \Delta \omega)}{2n} = \mu^{k+1} \leq \mu(\alpha) + \frac{\theta^2(1 + 2\theta)}{n(1 - 2\theta)^2} \mu(\alpha).$$

This proves the lemma. ∎

For linear programming, it is known [13, 22] that $\mu^{k+1} = \mu(\alpha)$. This claim is not always true for the convex quadratic programming as is pointed out in Lemma 5.12. Therefore, some extra work is needed to make sure that the duality gap will be reduced in every iteration.

Lemma 5.13 *For $\theta \leq 0.19$, if*

$$\sin(\alpha) = \frac{\theta}{\sqrt{n}}, \tag{68}$$

then $\mu^{k+1} < \mu^k$. Moreover, for $\sin(\alpha) = \frac{\theta}{\sqrt{n}} = \frac{0.19}{\sqrt{n}}$,

$$\mu^{k+1} \leq \mu^k \left(1 - \frac{0.0185}{\sqrt{n}} \right). \tag{69}$$

Proof: See Appendix A.

Remark 5.3 *As we have seen in this section that starting with (x^0, p^0, ω^0), the interior-point algorithm proceeds with finding $(x(\alpha), p(\alpha), \omega(\alpha)) \in \mathcal{N}_2(2\theta)$ and $(x^{k+1}, p^{k+1}, \omega^{k+1}) \in \mathcal{N}_2(\theta)$ such that $\mu^{k+1} < \mu^k$. In view of the proofs of Lemmas 5.7, 5.10, and 5.13, the positivity of $(x(\alpha), p(\alpha), \omega(\alpha)) > 0$ and $(x^{k+1}, p^{k+1}, \omega^{k+1}) > 0$ relies on $\mu(\alpha) > 0$ which, according to Corollary 5.1, is achievable for any θ and is given by a bound in terms of $\bar{\alpha}$. The proximity condition for $(x(\alpha), p(\alpha), \omega(\alpha))$ relies on the real positive root of $q(\sin(\alpha))$, denoted by $\sin(\tilde{\alpha})$, which is conservatively estimated in Lemma 5.9 under the condition that $\theta \leq 0.22$; the proximity condition for $(x^{k+1}, p^{k+1}, \omega^{k+1})$ is established in Lemma 5.10 under the condition that $\theta \leq 0.29$. Finally, duality gap reduction $\mu^{k+1} < \mu^k$ is established in Lemma 5.13 under the condition that $\theta \leq 0.19$. For all these results to hold, we just need to take the smallest bound $\theta = 0.19$.*

We summarize all the results in this section as the following theorem.

Theorem 5.2 *Let $\theta = 0.19$ and $(x^k, p^k, \omega^k) \in \mathcal{N}_2(\theta)$. Then, $(x(\alpha), p(\alpha), \omega(\alpha)) \in \mathcal{N}_2(2\theta)$; $(x^{k+1}, p^{k+1}, \omega^{k+1}) \in \mathcal{N}_2(\theta)$; and $\mu^{k+1} \le \mu^k \left(1 - \frac{0.0185}{\sqrt{n}}\right)$.*

Proof: From Corollary 5.1 and Lemma 5.9, we can select $\sin(\alpha) \le \min\{\sin(\tilde{\alpha}), \sin(\bar{\alpha})\}$. Therefore, Lemma 5.7 holds, i.e., $(x(\alpha), p(\alpha), \omega(\alpha)) \in \mathcal{N}_2(2\theta)$. Since $\sin(\alpha) \le \sin(\bar{\alpha})$ and $(x(\alpha), p(\alpha), \omega(\alpha)) \in \mathcal{N}_2(2\theta)$, Lemma 5.10 states $(x^{k+1}, p^{k+1}, \omega^{k+1}) \in \mathcal{N}_2(\theta)$. For $\theta = 0.19$ and $\sin(\alpha) = \frac{\theta}{\sqrt{n}}$, Lemma 5.13 states $\mu^{k+1} \le \mu^k \left(1 - \frac{0.0185}{\sqrt{n}}\right)$. This finishes the proof. ∎

Remark 5.4 *It is worthwhile to point out that $\theta = 0.19$ for the box constrained quadratic optimization problem is larger than the $\theta = 0.148$ for linearly constrained quadratic optimization problem. This makes the searching neighborhood larger and the following algorithm more efficient in computation than the algorithm in [24].*

We present the proposed method as the following

Algorithm 5.1 (Arc-search path-following)

Data: $H \ge 0$, c, n, $\theta = 0.19$, $\epsilon > 0$, initial point $(x^0, p^0, \omega^0) \in \mathcal{N}_2(\theta)$, and $\mu^0 = \frac{{p^0}^{\mathrm{T}} \omega^0}{2n}$.

for *iteration $k = 1, 2, \ldots$*

    Step 1: Solve the linear systems of equations (27) and (28) to get $(\dot{x}, \dot{p}, \dot{\omega})$ and $(\ddot{x}, \ddot{p}, \ddot{\omega})$.

    Step 2: Let $\sin(\alpha) = \frac{\theta}{\sqrt{n}}$. Update $(x(\alpha), p(\alpha), \omega(\alpha))$ and $\mu(\alpha)$ by (57) and (58).

    Step 3: Solve (63) to get $(\Delta x, \Delta p, \Delta \omega)$, update $(x^{k+1}, p^{k+1}, \omega^{k+1})$ and μ^{k+1} by using (64) and (65).

    Step 4: Set $k + 1 \to k$. Go back to Step 1.

end (for)

6 Convergence Analysis

The first result in this section extends a result of linear programming (c.f. [20]) to convex quadratic programming subject to box constraints.

Lemma 6.1 *Suppose $\mathcal{F}^o \ne \emptyset$. Then for each $K \ge 0$, the set*

$$\{(x, p, \omega) \mid (x, p, \omega) \in \mathcal{F}, \quad p^{\mathrm{T}} \omega \le K\}$$

is bounded.

Proof: First, x is bounded because $-e \leq x \leq e$. Since $x + y = e$ and $-e \leq x \leq e$, we have $0 \leq y = e - x \leq 2e$. Since $x - z = -e$, we have $0 \leq z = x + e \leq 2e$. Therefore, y and z are also bounded. Let $(\bar{x}, \bar{y}, \bar{z}, \bar{\lambda}, \bar{\gamma})$ be any fixed point in $\mathcal{F}^o$, and $(x, y, z, \lambda, \gamma)$ be any point in $\mathcal{F}$ with $y^{\mathrm{T}}\lambda + z^{\mathrm{T}}\gamma \leq K$. Then

$$H(\bar{x} - x) + (\bar{\lambda} - \lambda) - (\bar{\gamma} - \gamma) = 0.$$

Therefore

$$(\bar{x} - x)^{\mathrm{T}} H(\bar{x} - x) + (\bar{x} - x)^{\mathrm{T}}(\bar{\lambda} - \lambda) - (\bar{x} - x)^{\mathrm{T}}(\bar{\gamma} - \gamma) = 0,$$

or equivalently

$$(\bar{x} - x)^{\mathrm{T}}(\bar{\gamma} - \gamma) - (\bar{x} - x)^{\mathrm{T}}(\bar{\lambda} - \lambda) = (\bar{x} - x)^{\mathrm{T}} H(\bar{x} - x) \geq 0.$$

Using the relations $x - e = -y$ and $x + e = z$, we have

$$((\bar{x} + e) - (x + e))^{\mathrm{T}}(\bar{\gamma} - \gamma) - ((\bar{x} - e) - (x - e))^{\mathrm{T}}(\bar{\lambda} - \lambda) \geq 0,$$

or equivalently

$$(\bar{z} - z)^{\mathrm{T}}(\bar{\gamma} - \gamma) + (\bar{y} - y)^{\mathrm{T}}(\bar{\lambda} - \lambda) \geq 0.$$

This leads to

$$\bar{z}^{\mathrm{T}}\bar{\gamma} + z^{\mathrm{T}}\gamma - z^{\mathrm{T}}\bar{\gamma} - \bar{z}^{\mathrm{T}}\gamma + \bar{y}^{\mathrm{T}}\bar{\lambda} + y^{\mathrm{T}}\lambda - y^{\mathrm{T}}\bar{\lambda} - \bar{y}^{\mathrm{T}}\lambda \geq 0,$$

or in a compact form

$$\bar{p}^{\mathrm{T}}\bar{\omega} + p^{\mathrm{T}}\omega - p^{\mathrm{T}}\bar{\omega} - \bar{p}^{\mathrm{T}}\omega \geq 0.$$

Sine $(\bar{p}, \bar{\omega}) > 0$ is fixed, let

$$\xi = \min_{i=1,\cdots,n} \quad \min\{\bar{p}_i, \bar{\omega}_i\}.$$

Then, using $p^{\mathrm{T}}\omega \leq K$,

$$\bar{p}^{\mathrm{T}}\bar{\omega} + K \geq \xi e^{\mathrm{T}}(p + \omega) \geq \max_{i=1,\cdots,n} \max\{\xi p_i, \xi \omega_i\},$$

i.e., for $i \in \{1, \cdots, n\}$,

$$0 \leq p_i \leq \frac{1}{\xi}(K + \bar{p}^{\mathrm{T}}\bar{\omega}), \quad 0 \leq \omega_i \leq \frac{1}{\xi}(K + \bar{p}^{\mathrm{T}}\bar{\omega}).$$

This proves the lemma. ■

The following theorem is a direct result of Lemmas 6.1, 5.1, Theorem 5.2, KKT conditions, Theorem A.2 in [20].

Theorem 6.1 *Suppose that Assumption 1 holds, then the sequence generated by Algorithm 5.1 converges to a set of accumulation points, and all these accumulation points are global optimal solutions of the convex quadratic programming subject to box constraints.*

19

Let (x^*, p^*, ω^*) be any solution of (17), following the notation of [3], we denote index sets $\mathcal{B}$, $\mathcal{S}$, and $\mathcal{T}$ as

$$\mathcal{B} = \{j \in \{1, \ldots, 2n\} \mid p_j^* \neq 0\}. \tag{70}$$

$$\mathcal{S} = \{j \in \{1, \ldots, 2n\} \mid \omega_j^* \neq 0\}. \tag{71}$$

$$\mathcal{T} = \{j \in \{1, \ldots, 2n\} \mid p_j^* = \omega_j^* = 0\}. \tag{72}$$

According to Goldman-Tucker theorem [7], for the linear programming, $\mathcal{B} \cap \mathcal{S} = \emptyset = \mathcal{T}$ and $\mathcal{B} \cup \mathcal{S} = \{1, \ldots, 2n\}$. A solution with this property is called strictly complementary. This property has been used in many papers to prove the locally super-linear convergence of interior-point algorithms in linear programming. However, it is pointed out in [8] that this partition does not hold for general quadratic programming problems. We will show that as long as a convex quadratic programming subject to box constraints has strictly complementary solution(s), an interior-point algorithm will generate a sequence to approach strict complementary solution(s). As a matter of fact, from Lemma 6.1, we can extend the result of [20, Lemma 5.13] to the case of convex quadratic programming subject to box constraints, and obtain the following lemma which is independent of any algorithm.

Lemma 6.2 Let $\mu^0 > 0$, and $\rho \in (0, 1)$. Assume that the convex QP (16) has strictly complementary solution(s). Then for all points (x, p, ω) with $(x, p, \omega) \in \mathcal{F}^o$, $p_i \omega_i > \rho\mu$, and $\mu < \mu^0$, there are constants M, C_1, and C_2 such that

$$\|(p, \omega)\| \leq M, \tag{73}$$

$$0 < p_i \leq \mu/C_1 \ (i \in \mathcal{S}), \quad 0 < \omega_i \leq \mu/C_1 \ (i \in \mathcal{B}). \tag{74}$$

$$\omega_i \geq C_2 \rho \ (i \in \mathcal{S}), \quad p_i \geq C_2 \rho \ (i \in \mathcal{B}). \tag{75}$$

Proof: The first result (73) follows immediately from Lemma 6.1 by setting $K = 2n\mu^0$. Let (x^*, p^*, ω^*) be any strictly complementary solution. Since (x^*, p^*, ω^*) and (x, p, ω) are both feasible, we have

$$(y - y^*) = -(x - x^*) = -(z - z^*), \quad H(x - x^*) + (\lambda - \lambda^*) - (\gamma - \gamma^*) = 0.$$

Therefore,

$$(y - y^*)^{\mathrm{T}}(\lambda - \lambda^*) + (z - z^*)^{\mathrm{T}}(\gamma - \gamma^*) = (x - x^*)^{\mathrm{T}} H(x - x^*) \geq 0. \tag{76}$$

Since $(x^*, y^*, z^*, \lambda^*, \gamma^*) = (x^*, p^*, \omega^*)$ is strictly complementary solution, $\mathcal{T} = \emptyset$, $p_i^* = 0$ for $i \in \mathcal{S}$, and $\omega_i^* = 0$ for $i \in \mathcal{B}$. Since $p^{\mathrm{T}}\omega = 2n\mu$, $(p^*)^{\mathrm{T}}\omega^* = 0$, from (76), we have

$$p^{\mathrm{T}}\omega = y^{\mathrm{T}}\lambda + z^{\mathrm{T}}\gamma + \left((y^*)^{\mathrm{T}}\lambda^* + (z^*)^{\mathrm{T}}\gamma^*\right) \geq y^{\mathrm{T}}\lambda^* + z^{\mathrm{T}}\gamma^* + \left((y^*)^{\mathrm{T}}\lambda + (z^*)^{\mathrm{T}}\gamma\right) = p^{\mathrm{T}}\omega^* + \omega^{\mathrm{T}}p^*$$

$$\Longleftrightarrow \qquad 2n\mu \geq p^{\mathrm{T}}\omega^* + \omega^{\mathrm{T}}p^* = \sum_{i \in \mathcal{S}} p_i \omega_i^* + \sum_{i \in \mathcal{B}} p_i^* \omega_i. \tag{77}$$

Since each term in the summations is positive and bounded above by $2n\mu$, we have for any $i \in \mathcal{S}$, $\omega_i^* > 0$, therefore,

$$0 < p_i \leq \frac{2n\mu}{\omega_i^*}.$$

Denote $\Omega_D = \{(p^*, \omega^*) | \omega_i^* > 0\}$ and $\Omega_P = \{(p^*, \omega^*) | p_i^* > 0\}$, we have

$$0 < p_i \leq \frac{2n\mu}{\sup_{(p^*, \omega^*) \in \Omega_D} \omega_i^*}.$$

This leads to

$$\max_{i \in \mathcal{S}} p_i \leq \frac{2n\mu}{\min_{i \in \mathcal{S}} \sup_{(p^*, \omega^*) \in \Omega_D} \omega_i^*}.$$

Similarly,

$$\max_{i \in \mathcal{B}} \omega_i \leq \frac{2n\mu}{\min_{i \in \mathcal{B}} \sup_{(p^*, \omega^*) \in \Omega_P} p_i^*}.$$

Combining these 2 inequalities gives

$$\max\{\max_{i \in \mathcal{S}} p_i, \max_{i \in \mathcal{B}} \omega_i\} \leq \frac{2n\mu}{\min\{\min_{i \in \mathcal{S}} \sup_{(p^*, \omega^*) \in \Omega_D} \omega_i^*, \min_{i \in \mathcal{B}} \sup_{(p^*, \omega^*) \in \Omega_P} p_i^*\}} = \frac{\mu}{C_1}.$$

This proves (74). Finally, $p_i \omega_i \geq \rho\mu$, hence for any $i \in \mathcal{S}$,

$$\omega_i \geq \frac{\rho\mu}{p_i} \geq \frac{\rho\mu}{\mu/C_1} = C_2\rho.$$

Similarly, for any $i \in \mathcal{B}$,

$$p_i \geq \frac{\rho\mu}{\omega_i} \geq \frac{\rho\mu}{\mu/C_1} = C_2\rho.$$

∎

Lemma 6.2 leads to the following

Theorem 6.2 *Let $(x^k, p^k, \omega^k) \in \mathcal{N}_2(\theta)$ be generated by Algorithms 5.1. Assume that the convex QP with box constraints has strictly complementary solution(s). Then every limit point of the sequence is a strictly complementary solution of the convex quadratic programming with box constraints, i.e.,*

$$\omega_i^* \geq C_2\rho \quad (i \in \mathcal{S}), \qquad p_i^* \geq C_2\rho \quad (i \in \mathcal{B}). \tag{78}$$

Proof: From Lemma 6.2, (p^k, ω^k) is bounded, therefore there is at least one limit point (p^*, ω^*). Since (p_i^k, ω_i^k) is in the neighborhood of the central path, i.e., $p_i^k \omega_i^k > \rho \mu^k := (1 - 3\theta)\mu^k$,

$$\omega_i^k \geq C_2 \rho \ \ (i \in \mathcal{S}), \quad p_i^k \geq C_2 \rho \ \ (i \in \mathcal{B}),$$

every limit point will meet (78) due to the fact that $C_2 \rho$ is a constant. ∎

We now show that the complexity bound of Algorithm 5.1 is $O(\sqrt{n} \log(1/\epsilon))$. We need the following theorem from [20] for this purpose.

Theorem 6.3 *Let $\epsilon \in (0, 1)$ be given. Suppose that an algorithm for solving (17) generates a sequence of iterations that satisfies*

$$\mu^{k+1} \leq \left(1 - \frac{\delta}{n^\chi}\right) \mu^k, \quad k = 0, 1, 2, \ldots, \tag{79}$$

for some positive constants δ and χ. Suppose that the starting point (x^0, p^0, ω^0) satisfies $\mu^0 \leq 1/\epsilon$. Then there exists an index K with

$$K = O(n^\chi \log(1/\epsilon))$$

such that

$$\mu^k \leq \epsilon \ \ \text{for} \ \ \forall k \geq K.$$

Combining Lemma 5.13 and Theorems 6.3 gives

Theorem 6.4 *The complexity of Algorithm 5.1 is bounded by $O(\sqrt{n} \log(1/\epsilon))$.*

7 Implementation Issues

Algorithm 5.1 is presented in a form that is convenient for the convergence analysis. Some implementation details that make the algorithm effective and efficient are discussed in this section.

7.1 Termination criterion

Algorithm 5.1 needs a termination criterion in real implementation. One can use

$$\mu^k \leq \epsilon, \tag{80a}$$

$$\|r_X\| = \|Hx^k + \lambda^k - \gamma^k + c\| \leq \epsilon, \tag{80b}$$

$$\|r_Y\| = \|x^k + y^k - e\| \leq \epsilon, \tag{80c}$$

$$\|r_Z\| = \|x^k - z^k + e\| \leq \epsilon, \tag{80d}$$

$$\|r_t\| = \|P^k \Omega^k e - \mu e\| \leq \epsilon, \tag{80e}$$

$$(p^k, \omega^k) > 0. \tag{80f}$$

An alternate criterion is similar to the one used in `linprog` [26]

$$\kappa := \frac{\|r_Y\| + \|r_Z\|}{2n} + \frac{\|r_X\|}{\max\{1, \|c\|\}} + \frac{\mu^k}{\max\{1, \|x^{k\mathrm{T}} H x^k + c^\mathrm{T} x^k\|\}} \leq \epsilon. \tag{81}$$

7.2 Initial $(x^0, \lambda^0, s^0) \in \mathcal{N}_2(\theta)$

For feasible interior-point algorithms, an important prerequisite is to start with a feasible interior point. While finding an initial feasible point may not be a simple and trivial task for even linear programming with equality constraints [5], for quadratic programming subject to box constraints, finding the initial point is not an issue. We show that the following initial point $(x^0, y^0, z^0, \lambda^0, \gamma^0)$ is an interior point, moreover $(x^0, y^0, z^0, \lambda^0, \gamma^0) \in \mathcal{N}_2(\theta)$.

$$x^0 = 0, \quad y^0 = z^0 = e > 0, \tag{82a}$$

$$\lambda_i^0 = 4(1 + \|c\|^2) - \frac{c_i}{2} > 0, \tag{82b}$$

$$\gamma_i^0 = 4(1 + \|c\|^2) + \frac{c_i}{2} > 0. \tag{82c}$$

It is easy to see that this selected point meets (20). Therefore, we will show that it meets (24). Since

$$\mu^0 = \frac{\sum_{i=1}^n \left(\lambda_i^0 + \gamma_i^0\right)}{2n} = \frac{\sum_{i=1}^n \left(8(1 + \|c\|^2)\right)}{2n} = 4(1 + \|c\|^2), \tag{83}$$

we have, for $\theta = 0.19$,

$$\left\|p^0 \circ \omega^0 - \mu^0 e\right\|^2 = \sum_{i=1}^n (\lambda_i^0 - \mu^0)^2 + \sum_{i=1}^n (\gamma_i^0 - \mu^0)^2 = \frac{\|c\|^2}{2} \leq 16\theta^2 (1 + \|c\|^2)^2 = \theta^2 (\mu^0)^2.$$

7.3 Step size

Directly using $\sin(\alpha) = \frac{\theta}{\sqrt{n}}$ in Algorithm 5.1 provides an effective formula to prove the polynomiality. However, this choice of $\sin(\alpha)$ is too conservative in practice because this search step in $\mathcal{N}_2(2\theta)$ is too small and the speed of duality gap reduction is slow. A better choice of $\sin(\alpha)$ should have a larger step in every iteration so that the polynomiality is reserved and fast convergence is achieved. In view of Remark 5.3, conditions that restrict step size are positivity conditions, proximity conditions, and duality reduction condition. We examine how to enlarge the step size under these restrictions.

First, from (104) and (113), $\mu(\alpha) > 0$ is required for positivity conditions $(p(\alpha), \omega(\alpha)) > 0$ and $(p^{k+1}, \omega^{k+1}) > 0$ to hold. Since $\sin(\bar{\alpha})$ estimated in Corollary 5.1 is conservative, we find a better $\bar{\alpha}$ directly from (54).

$$\mu(\alpha) \geq \mu(1 - \sin(\alpha)) - \frac{1}{2n}(\dot{p}^{\mathrm{T}}\dot{\omega})\Big(\sin^4(\alpha) + \sin^2(\alpha)\Big) := f(\sin(\alpha)) = \sigma, \tag{84}$$

where $\sigma > 0$ is a small number, and $f(\sin(\alpha))$ is a monotonic decreasing function of $\sin(\alpha)$ with $f(\sin(0)) = 1$ and $f(\sin(\frac{\pi}{2})) < 0$. Therefore, (84) has a unique positive real solution for $\alpha \in [0, \frac{\pi}{2}]$ Since (84) is a quartic function of $\sin(\alpha)$, the cost of finding the smallest positive solution is negligible [10].

Second, for $\theta \leq 0.19$, from (112), the proximity condition for $(x^{k+1}, y^{k+1}, z^{k+1}, \lambda^{k+1}, \gamma^{k+1})$ holds without further restriction. The proximity condition (103) is met for $\sin(\alpha) \in [0, \sin(\tilde{\alpha})]$, where $\sin(\tilde{\alpha})$ is the smallest positive solution of (56) and it is estimated very conservatively in Lemma 5.9. An efficient implementation should use $\sin(\tilde{\alpha})$, the smallest positive solution of (56). Actually, there exist a $\acute{\alpha}$ which is normally larger than $\tilde{\alpha}$ such that the proximity condition (103) is met for $\sin(\alpha) \in [0, \sin(\acute{\alpha})]$. Let

$$b_0 = -\theta\mu < 0,$$

$$b_1 = \theta\mu > 0,$$

$$b_3 = \left\| \dot{p} \circ \ddot{\omega} + \dot{\omega} \circ \ddot{p} - \frac{1}{2n}(\dot{p}^{\mathrm{T}}\ddot{\omega} + \dot{\omega}^{\mathrm{T}}\ddot{p})e \right\| + \frac{\theta}{n}\left(\dot{p}^{\mathrm{T}}\ddot{\omega} + \ddot{p}^{\mathrm{T}}\dot{\omega}\right),$$

$$b_4 = \left\| \ddot{p} \circ \ddot{\omega} - \dot{\omega} \circ \dot{p} - \frac{1}{2n}(\ddot{p}^{\mathrm{T}}\ddot{\omega} - \dot{\omega}^{\mathrm{T}}\dot{p})e \right\| - \frac{\theta}{n}\left(\ddot{p}^{\mathrm{T}}\ddot{\omega} - \dot{p}^{\mathrm{T}}\dot{\omega}\right),$$

and

$$p(\alpha) := b_4(1 - \cos(\alpha))^2 + b_3\sin(\alpha)(1 - \cos(\alpha)) + b_1\sin(\alpha) + b_0. \tag{85}$$

Applying the second inequality of (45) to $\frac{\theta}{n}\left(\dot{p}^{\mathrm{T}}\ddot{\omega} + \ddot{p}^{\mathrm{T}}\dot{\omega}\right)\sin(\alpha)(1 - \cos(\alpha))$, we can easily show that

$$p(\alpha) \leq q(\alpha),$$

where $q(\alpha)$ is defined in (56). Therefore, the smallest positive solution $\acute{\alpha}$ of $p(\alpha)$ is larger than the smallest positive solution $\tilde{\alpha}$ of $q(\alpha)$. We will show that for $\sin(\alpha) \in [0, \sin(\acute{\alpha})]$, the proximity condition (103) holds. Since for $\sin(\alpha) \in [0, \sin(\acute{\alpha})]$, $p(\alpha) \leq 0$, we have

$$\left\| \ddot{p} \circ \ddot{\omega} - \dot{\omega} \circ \dot{p} - \frac{1}{2n}(\ddot{p}^{\mathrm{T}}\ddot{\omega} - \dot{\omega}^{\mathrm{T}}\dot{p})e \right\| (1 - \cos(\alpha))^2 + \left\| \dot{p} \circ \ddot{\omega} + \dot{\omega} \circ \ddot{p} - \frac{1}{2n}(\dot{p}^{\mathrm{T}}\ddot{\omega} + \dot{\omega}^{\mathrm{T}}\ddot{p})e \right\| \sin(\alpha)(1 - \cos(\alpha))$$

$$\leq (2\theta)\left(\frac{1}{2n}\left(\ddot{p}^{\mathrm{T}}\ddot{\omega} - \dot{p}^{\mathrm{T}}\dot{\omega}\right)(1 - \cos(\alpha))^2 - \frac{1}{2n}\left(\dot{p}^{\mathrm{T}}\ddot{\omega} + \ddot{p}^{\mathrm{T}}\dot{\omega}\right)\sin(\alpha)(1 - \cos(\alpha)) \right) - \theta\mu(1 - \sin(\alpha)). \quad (86)$$

Substituting this inequality into (102) gives

$$\left\| p(\alpha) \circ \omega(\alpha) - \mu(\alpha)e \right\|$$
$$\leq 2\theta\left(\mu(1 - \sin(\alpha)) + \frac{1}{2n}\left(\ddot{x}^{\mathrm{T}}(\ddot{\gamma} - \ddot{\lambda}) - \dot{x}^{\mathrm{T}}(\dot{\gamma} - \dot{\lambda})\right)(1 - \cos(\alpha))^2 \right.$$
$$\left. - \frac{1}{2n}\left(\dot{x}^{\mathrm{T}}(\ddot{\gamma} - \ddot{\lambda}) + \ddot{x}^{\mathrm{T}}(\dot{\gamma} - \dot{\lambda})\right)\sin(\alpha)(1 - \cos(\alpha)) \right) = 2\theta\mu(\alpha). \quad (87)$$

This is the proximity condition for $(x(\alpha), y(\alpha), z(\alpha), \lambda(\alpha), \gamma(\alpha))$. Denote $\hat{b}_0 = b_0$, $\hat{b}_1 = b_1$,

$$\hat{b}_3 = \begin{cases} b_3 & \text{if } b_3 \geq 0, \\ 0 & \text{if } b_3 < 0, \end{cases} \qquad \hat{b}_4 = \begin{cases} b_4 & \text{if } b_4 \geq 0, \\ 0 & \text{if } b_4 < 0, \end{cases}$$

and

$$\hat{p}(\alpha) := \hat{b}_4(1 - \cos(\alpha))^2 + \hat{b}_3\sin(\alpha)(1 - \cos(\alpha)) + \hat{b}_1\sin(\alpha) + \hat{b}_0. \quad (88)$$

Since $\hat{p}(\alpha) \geq p(\alpha)$, the smallest positive solution $\acute{\alpha}$ of $\hat{p}(\alpha)$ is smaller than smallest positive solution $\acute{\alpha}$ of $p(\alpha)$. To estimate the smallest solution of $\acute{\alpha}$, by noticing that $\hat{p}(\alpha)$ is a monotonic increasing function of α and $\hat{p}(0) = -\theta\mu < 0$, we can simply use the bisection method. The computational cost is impendent of the problem size n and is negligible. Since both estimated step sizes $\acute{\alpha}$ and $\tilde{\alpha}$ guarantee the proximity condition for $(x(\alpha), y(\alpha), z(\alpha), \lambda(\alpha), \gamma(\alpha))$ to hold, we select $\check{\alpha} = \max\{\acute{\alpha}, \tilde{\alpha}\} \geq \tilde{\alpha}$ which guarantees the polynomiality claim to hold.

Third, from (121a) and Lemma 5.5, we have

$$\mu^{k+1} \leq \mu^k \left(1 + \frac{\theta^2(1 + 2\theta)}{n(1 - 2\theta)^2} - \left(1 + \frac{\theta^2(1 + 2\theta)}{n(1 - 2\theta)^2}\right)\sin(\alpha) + \left(1 + \frac{\theta^2(1 + 2\theta)}{n(1 - 2\theta)^2}\right)\frac{\ddot{p}^{\mathrm{T}}\ddot{\omega}}{2n\mu}\left(\sin^2(\alpha) + \sin^4(\alpha)\right) \right).$$

For $\mu^{k+1} \leq \mu^k$ to hold, we need

$$\frac{\theta^2(1 + 2\theta)}{n(1 - 2\theta)^2} - \left(1 + \frac{\theta^2(1 + 2\theta)}{n(1 - 2\theta)^2}\right)\sin(\alpha) + \left(1 + \frac{\theta^2(1 + 2\theta)}{n(1 - 2\theta)^2}\right)\frac{\ddot{p}^{\mathrm{T}}\ddot{\omega}}{2n\mu}\left(\sin^2(\alpha) + \sin^4(\alpha)\right) \leq 0.$$

For the sake of convenience in convergence analysis, a conservative estimate is used in Lemma 5.13. For efficient implementation, the following solution should be adopted. Denote $u = \frac{\theta^2(1 + 2\theta)}{n(1 - 2\theta)^2} > 0$, $v = \frac{\ddot{p}^{\mathrm{T}}\ddot{\omega}}{2n\mu} > 0$, $z = \sin(\alpha) \in [0, 1]$, and

$$F(z) = (1 + u)vz^4 + (1 + u)vz^2 - (1 + u)z + u.$$

For $z \in [0, 1]$ and $v \le \frac{1}{6}$, $F'(z) = (1 + u)(4vz^3 + 2vz - 1) \le 0$, therefore, the upper bound of the duality gap is a monotonic decreasing function of $\sin(\alpha)$ for $\alpha \in [0, \frac{\pi}{2}]$. The larger α is, the smaller the upper bound of the duality gap will be. For $v > \frac{1}{6}$, to minimize the upper bound of the duality gap, we can find the solution of $F'(z) = 0$. It is easy to check from discriminator [15] that the cubic polynomial $F'(z)$ has only one real solution which is given by (see Lemma 2.5)

$$\sin(\breve{\alpha}) = \sqrt[3]{\frac{n\mu}{4\ddot{p}^{\mathrm{T}}\ddot{\omega}} + \sqrt{\left(\frac{n\mu}{4\ddot{p}^{\mathrm{T}}\ddot{\omega}}\right)^2 + \left(\frac{1}{6}\right)^3}} + \sqrt[3]{\frac{n\mu}{4\ddot{p}^{\mathrm{T}}\ddot{\omega}} - \sqrt{\left(\frac{n\mu}{4\ddot{p}^{\mathrm{T}}\ddot{\omega}}\right)^2 + \left(\frac{1}{6}\right)^3}}.$$

Since $F''(\sin(\breve{\alpha})) = (1 + u)(12v \sin^2(\breve{\alpha}) + 2v) > 0$, at $\sin(\breve{\alpha}) \in [0, 1)$, the upper bound of the duality gap is minimized. Therefore, we can define

$$\breve{\alpha} = \begin{cases} \frac{\pi}{2}, & \text{if } \frac{\ddot{p}^{\mathrm{T}}\ddot{\omega}}{2n\mu} \le \frac{1}{6} \\[3mm] \sin^{-1}\left(\sqrt[3]{\frac{n\mu}{4\ddot{p}^{\mathrm{T}}\ddot{\omega}} + \sqrt{\left(\frac{n\mu}{4\ddot{p}^{\mathrm{T}}\ddot{\omega}}\right)^2 + \left(\frac{1}{6}\right)^3}} + \sqrt[3]{\frac{n\mu}{4\ddot{p}^{\mathrm{T}}\ddot{\omega}} - \sqrt{\left(\frac{n\mu}{4\ddot{p}^{\mathrm{T}}\ddot{\omega}}\right)^2 + \left(\frac{1}{6}\right)^3}}\right), & \text{if } \frac{\ddot{p}^{\mathrm{T}}\ddot{\omega}}{2n\mu} > \frac{1}{6}. \end{cases} \quad (89)$$

It is worthwhile to note that for $\alpha < \breve{\alpha}$, $F'(\sin(\alpha)) < 0$, i.e., $F(\sin(\alpha))$ is a monotonic decreasing function of $\alpha \in [0, \breve{\alpha}]$.

We summarize the step size selection process as a simple algorithm as follows.

Algorithm 7.1 (Step Size Selection)

Data: $\sigma > 0$.

Step 1: Find the positive real solution of (84) to get $\sin(\bar{\alpha})$

Step 2: Find the smallest positive real solution of (88) to get $\sin(\acute{\alpha})$, the smallest positive real solution of (56) to get $\sin(\tilde{\alpha})$, and set $\sin(\check{\alpha}) = \max\{\sin(\tilde{\alpha}), \sin(\acute{\alpha})\}$.

Step 3: Calculate $\breve{\alpha}$ given by (89)

Step 4: The step size is obtained as $\sin(\alpha) = \min\{\sin(\bar{\alpha}), \sin(\check{\alpha}), \sin(\breve{\alpha})\}$.

7.4   The practical implementation

Therefore, Algorithm 5.1 can be implemented as follows.

Algorithm 7.2 (Arc-search path-following)

Data: $H \ge 0$, c, n, $\theta = 0.19$, $\epsilon > \sigma > 0$.

Step 0: Find initial point $(x^0, p^0, \omega^0) \in \mathcal{N}_2(\theta)$ using (82), κ using (81), and μ^0 using (83).

while $\kappa > \epsilon$

Step 1: Compute $(\dot{x}, \dot{p}, \dot{\omega})$ *and* $(\ddot{x}, \ddot{p}, \ddot{\omega})$ *using (27) and (28).*

Step 2: Select $\sin(\alpha)$ *using Algorithm 7.1. Update* $(x(\alpha), p(\alpha), \omega(\alpha))$ *and* $\mu(\alpha)$ *using (57) and (58).*

Step 3: Compute $(\Delta x, \Delta p, \Delta \omega)$ *using (63), update* $(x^{k+1}, p^{k+1}, \omega^{k+1})$ *and* μ^{k+1} *using (64) and (65).*

Step 4: Computer κ *using (81).*

Step 5: Set $k + 1 \to k$. *Go back to Step 1.*

end (while)

Remark 7.1 *The condition* $\mu > \sigma$ *guarantees that the equation (84) has a positive solution before terminate criterion is met.*

8   A design example

In this section, we will use the design example of [4] to demonstrate the effectiveness and efficiency of the proposed algorithm. The linear time-invariant system under consideration is given by

$$\mathbf{x}_{s+1} = \begin{bmatrix} 1 & h \\ -h & 1 \end{bmatrix} \mathbf{x}_s + \begin{bmatrix} 0 \\ h \end{bmatrix} \mathbf{u}_s = A\mathbf{x}_s + B\mathbf{u}_s, \quad s = 0, \dots, N-1, \tag{90}$$

with the initial state given by $\mathbf{x}_t = [15, 5]^{\mathrm{T}}$. The control constraints are

$$-1 \le \mathbf{u}_s \le 1, \quad s = 0, \dots, N-1. \tag{91}$$

The problem is to minimize

$$J = \min_{\mathbf{u}_0, \mathbf{u}_1, \cdots, \mathbf{u}_{N-1}} \frac{1}{2} \mathbf{x}_N^{\mathrm{T}} P \mathbf{x}_N + \frac{h}{2} \sum_{k=0}^{N-1} \left[\mathbf{x}_k^{\mathrm{T}} Q \mathbf{x}_k + R\mathbf{u}_k^2, \right] \tag{92}$$

where the matrices P, Q, and scalar R are given by

$$P = Q = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad R = 6.$$

This problem arises from discretization of the continuous-time problem of minimizing

$$\frac{1}{2} \mathbf{x}_T^{\mathrm{T}} P \mathbf{x}_T + \int_0^T \left[\mathbf{x}(t)^{\mathrm{T}} Q \mathbf{x}(t) + R\mathbf{u}(t)^2 \right] dt$$

subject to

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mathbf{u}(t) = A\mathbf{x}(t) + B\mathbf{u}(t),$$

27

and

$$1 \leq \mathbf{u}(t) \leq 1, \quad t \in [0, T],$$

where the interval $[0, T]$ is discretized into N intervals of length $h = \frac{T}{N}$.

In our implementation of Algorithm 7.2, $\epsilon = 10^{-8}$ and $\sigma = 10^{-10}$ are selected. For the simple design example, $T = 50$ and $N = 500$ are used. The original quadratic optimization problem has $Nr + Nm = 1500$ variables, $Nr = 1000$ equality constraints, and $2Nm = 1000$ inequality constraints. The reduced problem has $Nm = 500$ variables, no equality constraints, and $2Nm = 1000$ inequality constraints, a significantly simpler problem. The advantage will be even more significant if the state dimension r is significantly larger than the control dimension m. After 27 iterations, the algorithm converges. Using the optimal control inputs, we can calculate the state space response from (90). The control inputs and state space response are displayed in Figure 1.
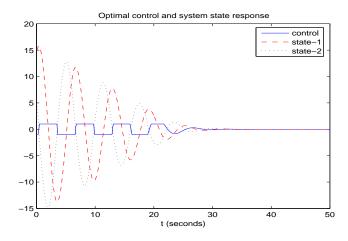
Figure 1: Optimal control with saturation constraint.

9   Conclusions

This paper proposes an arc-search interior-point algorithm for convex quadratic programming subject to box constraints that searches the optimizers along ellipses that approximate the central path. The saturation constrained LQR design is one such problem. The algorithm is proved to be polynomial with the complexity bound $O(\sqrt{n}\log(1/\epsilon))$. A constrained LQR design example from [4] is provided to demonstrate how the algorithm works. Preliminary test on this simple design problem shows that the proposed algorithm is promising. A MATLAB M-file implementation of Algorithm 7.2 is available from the author.

10   Appendix A: Proofs of Technical Lemmas

Proof of Lemma 5.2:

From (30), we have $\dot{x}^{\mathrm{T}}(\dot{\gamma} - \dot{\lambda}) = \dot{z}^{\mathrm{T}}\dot{\gamma} + \dot{y}^{\mathrm{T}}\dot{\lambda} = \dot{p}^{\mathrm{T}}\dot{\omega}$, $\ddot{x}^{\mathrm{T}}(\ddot{\gamma} - \ddot{\lambda}) = \ddot{z}^{\mathrm{T}}\ddot{\gamma} + \ddot{y}^{\mathrm{T}}\ddot{\lambda} = \ddot{p}^{\mathrm{T}}\ddot{\omega}$, $\ddot{x}^{\mathrm{T}}(\dot{\gamma} - \dot{\lambda}) = \ddot{p}^{\mathrm{T}}\dot{\omega}$,
and $\dot{x}^{\mathrm{T}}(\ddot{\gamma} - \ddot{\lambda}) = \dot{p}^{\mathrm{T}}\ddot{\omega}$. Pre-multiplying $\dot{x}^{\mathrm{T}}$ and $\ddot{x}^{\mathrm{T}}$ to (29) gives

$$\dot{x}^{\mathrm{T}}(\dot{\gamma} - \dot{\lambda}) = \dot{x}^{\mathrm{T}}H\dot{x},$$

$$\ddot{x}^{\mathrm{T}}(\ddot{\gamma} - \ddot{\lambda}) = \ddot{x}^{\mathrm{T}}H\ddot{x},$$

$$\ddot{x}^{\mathrm{T}}(\dot{\gamma} - \dot{\lambda}) = \ddot{x}^{\mathrm{T}}H\dot{x} = \dot{x}^{\mathrm{T}}H\ddot{x} = \dot{x}^{\mathrm{T}}(\ddot{\gamma} - \ddot{\lambda}).$$

(41) and (42) follow from the first two equations and the fact that H is positive definite. The last equation gives (43). Using (41), (42), and (43) gives

$$
\begin{aligned}
&(\dot{x}(1 - \cos(\alpha)) + \ddot{x}\sin(\alpha))^{\mathrm{T}}H(\dot{x}(1 - \cos(\alpha)) + \ddot{x}\sin(\alpha)) \\
=\ &(\dot{x}^{\mathrm{T}}H\dot{x})(1 - \cos(\alpha))^2 + 2(\dot{x}^{\mathrm{T}}H\ddot{x})\sin(\alpha)(1 - \cos(\alpha)) + (\ddot{x}^{\mathrm{T}}H\ddot{x})\sin^2(\alpha) \\
=\ &(\dot{x}^{\mathrm{T}}H\dot{x})(1 - \cos(\alpha))^2 + (\ddot{x}^{\mathrm{T}}H\ddot{x})\sin^2(\alpha) + (\ddot{x}^{\mathrm{T}}(\dot{\gamma} - \dot{\lambda}) + \dot{x}^{\mathrm{T}}(\ddot{\gamma} - \ddot{\lambda}))\sin(\alpha)(1 - \cos(\alpha)) \geq 0,
\end{aligned}
$$

which is the first inequality of (44). Using (41), (42), and (43) also gives

$$
\begin{aligned}
&(\dot{x}(1 - \cos(\alpha)) - \ddot{x}\sin(\alpha))^{\mathrm{T}}H(\dot{x}(1 - \cos(\alpha)) - \ddot{x}\sin(\alpha)) \\
=\ &(\dot{x}^{\mathrm{T}}H\dot{x})(1 - \cos(\alpha))^2 - 2(\dot{x}^{\mathrm{T}}H\ddot{x})\sin(\alpha)(1 - \cos(\alpha)) + (\ddot{x}^{\mathrm{T}}H\ddot{x})\sin^2(\alpha) \\
=\ &(\dot{x}^{\mathrm{T}}H\dot{x})(1 - \cos(\alpha))^2 + (\ddot{x}^{\mathrm{T}}H\ddot{x})\sin^2(\alpha) - (\ddot{x}^{\mathrm{T}}(\dot{\gamma} - \dot{\lambda}) + \dot{x}^{\mathrm{T}}(\ddot{\gamma} - \ddot{\lambda}))\sin(\alpha)(1 - \cos(\alpha)) \geq 0,
\end{aligned}
$$

which is the second inequality of (44). Replacing $\dot{x}(1 - \cos(\alpha))$ and $\ddot{x}\sin(\alpha)$ by $\dot{x}\sin(\alpha)$ and $\ddot{x}(1 - \cos(\alpha))$, and following the same method, we can obtain (45). ∎

Proof of Lemma 5.3:

From the last two rows of (27) or equivalently (31), we have

$$\Lambda\dot{y} + Y\dot{\lambda} = \Lambda Y e$$

$$\Gamma\dot{z} + Z\dot{\gamma} = \Gamma Z e.$$

Pre-multiplying $Y^{-\frac{1}{2}}\Lambda^{-\frac{1}{2}}$ on both sides of the first equality gives

$$Y^{-\frac{1}{2}}\Lambda^{\frac{1}{2}}\dot{y} + Y^{\frac{1}{2}}\Lambda^{-\frac{1}{2}}\dot{\lambda} = Y^{\frac{1}{2}}\Lambda^{\frac{1}{2}}e.$$

Pre-multiplying $Z^{-\frac{1}{2}}\Gamma^{-\frac{1}{2}}$ on both sides of the second equality gives

$$Z^{-\frac{1}{2}}\Gamma^{\frac{1}{2}}\dot{z} + Z^{\frac{1}{2}}\Gamma^{-\frac{1}{2}}\dot{\gamma} = Z^{\frac{1}{2}}\Gamma^{\frac{1}{2}}e. \tag{93}$$

Let $u = \begin{bmatrix} Y^{-\frac{1}{2}}\Lambda^{\frac{1}{2}}\dot{y} \\ Z^{-\frac{1}{2}}\Gamma^{\frac{1}{2}}\dot{z} \end{bmatrix}$, $v = \begin{bmatrix} Y^{\frac{1}{2}}\Lambda^{-\frac{1}{2}}\dot{\lambda} \\ Z^{\frac{1}{2}}\Gamma^{-\frac{1}{2}}\dot{\gamma} \end{bmatrix}$, and $w = \begin{bmatrix} Y^{\frac{1}{2}}\Lambda^{\frac{1}{2}}e \\ Z^{\frac{1}{2}}\Gamma^{\frac{1}{2}}e \end{bmatrix}$, use (30) and Lemma 5.2, we have $u^{\mathrm{T}}v = \dot{y}^{\mathrm{T}}\dot{\lambda} + \dot{z}^{\mathrm{T}}\dot{\gamma} = \dot{x}^{\mathrm{T}}(\dot{\gamma} - \dot{\lambda}) \geq 0$. Using Lemma 2.3 and (23), we have

$$\|u\|^2 + \|v\|^2 = \sum_{i=1}^{n}\left(\frac{\dot{y}_i^2\lambda_i}{y_i} + \frac{\dot{z}_i^2\gamma_i}{z_i}\right) + \sum_{i=1}^{n}\left(\frac{\dot{\lambda}_i^2 y_i}{\lambda_i} + \frac{\dot{\gamma}_i^2 z_i}{\gamma_i}\right) \leq \sum_{i=1}^{n}(y_i\lambda_i + z_i\gamma_i) = \sum_{i=1}^{2n} p_i\omega_i = 2n\mu.$$

Since $p_i > 0$ and $\omega_i > 0$, dividing both sides of the inequality by $\min_j p_i\omega_i$ and using (25) gives

$$\sum_{i=1}^{n}\left(\frac{\dot{y}_i^2}{y_i^2} + \frac{\dot{z}_i^2}{z_i^2}\right) + \sum_{i=1}^{n}\left(\frac{\dot{\gamma}_i^2}{\gamma_i^2} + \frac{\dot{\lambda}_i^2}{\lambda_i^2}\right) = \left\|\frac{\dot{p}}{p}\right\|^2 + \left\|\frac{\dot{\omega}}{\omega}\right\|^2 \leq \frac{2n\mu}{\min_j p_i\omega_i} \leq \frac{2n}{1-\theta}. \tag{94}$$

This proves (47). Combining (47) and Lemma 2.1 yields

$$\left\|\frac{\dot{p}}{p}\right\|^2\left\|\frac{\dot{\omega}}{\omega}\right\|^2 \leq \left(\frac{n}{(1-\theta)}\right)^2.$$

This leads to,

$$\left\|\frac{\dot{p}}{p}\right\|\left\|\frac{\dot{\omega}}{\omega}\right\| \leq \frac{n}{(1-\theta)}. \tag{95}$$

Therefore, using (25) and Cauchy-Schwarz inequality yields

$$\frac{\dot{p}^{\mathrm{T}}\dot{\omega}}{\mu} \leq \frac{|\dot{p}|^{\mathrm{T}}|\dot{\omega}|}{\mu} \leq (1+\theta)\frac{|\dot{p}|^{\mathrm{T}}|\dot{\omega}|}{\max_i p_i\omega_i} \leq (1+\theta)\left(\frac{|\dot{p}|}{p}\right)^{\mathrm{T}}\left(\frac{|\dot{\omega}|}{\omega}\right) \leq (1+\theta)\left\|\frac{\dot{p}}{p}\right\|\left\|\frac{\dot{\omega}}{\omega}\right\| \leq \frac{1+\theta}{1-\theta}n, \tag{96}$$

which is the second inequality of (49). From Lemma 5.2, $\dot{p}^{\mathrm{T}}\dot{\omega} = \dot{x}^{\mathrm{T}}(\dot{\gamma} - \dot{\lambda}) = \dot{x}^{\mathrm{T}}H\dot{x} \geq 0$, we have the first inequality of (49). ∎

Proof of Lemma 5.4:

Similar to the proof of Lemma 5.3, from (32), we have

$$\Lambda\ddot{y} + Y\ddot{\lambda} = -2\left(\dot{y} \circ \dot{\lambda}\right)$$
$$\iff Y^{-\frac{1}{2}}\Lambda^{\frac{1}{2}}\ddot{y} + Y^{\frac{1}{2}}\Lambda^{-\frac{1}{2}}\ddot{\lambda} = -2Y^{-\frac{1}{2}}\Lambda^{-\frac{1}{2}}\left(\dot{y} \circ \dot{\lambda}\right),$$

and

$$\Gamma\ddot{z} + Z\ddot{\gamma} = -2\left(\dot{z} \circ \dot{\gamma}\right)$$
$$\iff Z^{-\frac{1}{2}}\Gamma^{\frac{1}{2}}\ddot{z} + Z^{\frac{1}{2}}\Gamma^{-\frac{1}{2}}\ddot{\gamma} = -2Z^{-\frac{1}{2}}\Gamma^{-\frac{1}{2}}\left(\dot{z} \circ \dot{\gamma}\right).$$

Let $u = \begin{bmatrix} Y^{-\frac{1}{2}}\Lambda^{\frac{1}{2}}\ddot{y} \\ Z^{-\frac{1}{2}}\Gamma^{\frac{1}{2}}\ddot{z} \end{bmatrix}$, $v = \begin{bmatrix} Y^{\frac{1}{2}}\Lambda^{-\frac{1}{2}}\ddot{\lambda} \\ Z^{\frac{1}{2}}\Gamma^{-\frac{1}{2}}\ddot{\gamma} \end{bmatrix}$, and $w = \begin{bmatrix} -2Y^{-\frac{1}{2}}\Lambda^{-\frac{1}{2}}\left(\dot{y}\circ\dot{\lambda}\right) \\ -2Z^{-\frac{1}{2}}\Gamma^{-\frac{1}{2}}\left(\dot{z}\circ\dot{\gamma}\right) \end{bmatrix}$, using (30) and Lemma

5.2, $u^{\mathrm{T}}v = \ddot{y}^{\mathrm{T}}\ddot{\lambda} + \ddot{z}^{\mathrm{T}}\ddot{\gamma} = \ddot{x}^{\mathrm{T}}(\ddot{\gamma} - \ddot{\lambda}) \geq 0$. Using Lemma 2.3, we have

$$
\begin{aligned}
\|u\|^2 + \|v\|^2 &= \sum_{i=1}^{n}\left(\frac{\ddot{y}_i^2\lambda_i}{y_i} + \frac{\ddot{z}_i^2\gamma_i}{z_i}\right) + \sum_{i=1}^{n}\left(\frac{\ddot{\lambda}_i^2 y_i}{\lambda_i} + \frac{\ddot{\gamma}_i^2 z_i}{\gamma_i}\right) \\
&\leq \left\|-2Y^{-\frac{1}{2}}\Lambda^{-\frac{1}{2}}\left(\dot{y}\circ\dot{\lambda}\right)\right\|^2 + \left\|-2Z^{-\frac{1}{2}}\Gamma^{-\frac{1}{2}}\left(\dot{z}\circ\dot{\gamma}\right)\right\|^2 \\
&= 4\sum_{i=1}^{n}\left(\frac{\dot{y}_i^2}{y_i}\frac{\dot{\lambda}_i^2}{\lambda_i} + \frac{\dot{z}_i^2}{z_i}\frac{\dot{\gamma}_i^2}{\gamma_i}\right).
\end{aligned}
$$

Dividing both sides of the inequality by μ and using (25) gives

$$
\begin{aligned}
&(1-\theta)\left(\sum_{i=1}^{n}\left(\frac{\ddot{y}_i^2}{y_i^2} + \frac{\ddot{z}_i^2}{z_i^2}\right) + \sum_{i=1}^{n}\left(\frac{\ddot{\lambda}_i^2}{\lambda_i^2} + \frac{\ddot{\gamma}_i^2}{\gamma_i^2}\right)\right) \\
&= (1-\theta)\left(\left\|\frac{\ddot{p}}{p}\right\|^2 + \left\|\frac{\ddot{\omega}}{\omega}\right\|^2\right) \\
&\leq 4(1+\theta)\left(\sum_{i=1}^{n}\left(\frac{\dot{y}_i^2}{y_i^2}\frac{\dot{\lambda}_i^2}{\lambda_i^2} + \frac{\dot{z}_i^2}{z_i^2}\frac{\dot{\gamma}_i^2}{\gamma_i^2}\right)\right),
\end{aligned}
$$

in view of Lemma 5.3, this leads to

$$
\left\|\frac{\ddot{p}}{p}\right\|^2 + \left\|\frac{\ddot{\omega}}{\omega}\right\|^2 \leq 4\frac{1+\theta}{1-\theta}\left\|\frac{\dot{p}}{p}\circ\frac{\dot{\omega}}{\omega}\right\|^2 \leq 4\frac{1+\theta}{1-\theta}\left\|\frac{\dot{p}}{p}\right\|^2\left\|\frac{\dot{\omega}}{\omega}\right\|^2 \leq \frac{4(1+\theta)n^2}{(1-\theta)^3}. \tag{97}
$$

This proves (50). Combining (50) and Lemma 2.1 yields

$$
\left\|\frac{\ddot{p}}{p}\right\|^2\left\|\frac{\ddot{\omega}}{\omega}\right\|^2 \leq \left(\frac{2(1+\theta)n^2}{(1-\theta)^3}\right)^2.
$$

Using (25) and Cauchy-Schwarz inequality yields

$$
\frac{\ddot{p}^{\mathrm{T}}\ddot{\omega}}{\mu} \leq \frac{|\ddot{p}|^{\mathrm{T}}|\ddot{\omega}|}{\mu} \leq (1+\theta)\frac{|\ddot{p}|^{\mathrm{T}}|\ddot{\omega}|}{\max_i p_i\omega_i} \leq (1+\theta)\left(\frac{|\ddot{p}|}{p}\right)^{\mathrm{T}}\left(\frac{|\ddot{\omega}|}{\omega}\right) \leq (1+\theta)\left\|\frac{\ddot{p}}{p}\right\|\left\|\frac{\ddot{\omega}}{\omega}\right\| \leq \frac{2n^2(1+\theta)^2}{(1-\theta)^3},
$$

which is the second inequality of (52). Using (30) and Lemma 5.2, we have $\ddot{p}^{\mathrm{T}}\ddot{\omega} = \ddot{y}^{\mathrm{T}}\ddot{\lambda} + \ddot{z}^{\mathrm{T}}\ddot{\gamma} = \ddot{x}^{\mathrm{T}}(\ddot{\gamma}-\ddot{\lambda}) = \ddot{x}^{\mathrm{T}}H\ddot{x} \geq 0$. This proves the first inequality of (52). Finally, using (25), Cauchy-Schwarz inequality, (47), and (50) yields

$$
\begin{aligned}
\frac{\left|\dot{p}^{\mathrm{T}}\ddot{\omega}\right|}{\mu} &\leq \frac{|\dot{p}|^{\mathrm{T}}|\ddot{\omega}|}{\mu} \leq (1+\theta)\frac{|\dot{p}|^{\mathrm{T}}|\ddot{\omega}|}{\max_i p_i\omega_i} \leq (1+\theta)\left(\frac{|\dot{p}|}{p}\right)^{\mathrm{T}}\left(\frac{|\ddot{\omega}|}{\omega}\right) \\
&\leq (1+\theta)\left\|\frac{\dot{p}}{p}\right\|\left\|\frac{\ddot{\omega}}{\omega}\right\| \leq (1+\theta)\left(\frac{2n}{1-\theta}\right)^{\frac{1}{2}}\left(\frac{4(1+\theta)n^2}{(1+\theta)^3}\right)^{\frac{1}{2}} \leq \frac{(2n(1+\theta))^{\frac{3}{2}}}{(1-\theta)^2}.
\end{aligned}
$$

This proves the first inequality of (53). Replacing $\dot{p}$ by $\ddot{p}$ and $\ddot{\omega}$ by $\dot{\omega}$, then using the same reasoning, we can prove the second inequality of (53). ∎

Proof of Lemma 5.5:

Using (34), (36), (31), and (32), we have

$$y^{\mathrm{T}}(\alpha)\lambda(\alpha)$$

$$=\Big(y^{\mathrm{T}} - \dot{y}^{\mathrm{T}}\sin(\alpha) + \ddot{y}^{\mathrm{T}}(1-\cos(\alpha))\Big)\Big(\lambda - \dot{\lambda}\sin(\alpha) + \ddot{\lambda}(1-\cos(\alpha))\Big)$$

$$=y^{\mathrm{T}}\lambda - y^{\mathrm{T}}\dot{\lambda}\sin(\alpha) + y^{\mathrm{T}}\ddot{\lambda}(1-\cos(\alpha))$$

$$\quad - \dot{y}^{\mathrm{T}}\lambda\sin(\alpha) + \dot{y}^{\mathrm{T}}\dot{\lambda}\sin^2(\alpha) - \dot{y}^{\mathrm{T}}\ddot{\lambda}\sin(\alpha)(1-\cos(\alpha))$$

$$\quad + \ddot{y}^{\mathrm{T}}\lambda(1-\cos(\alpha)) - \ddot{y}^{\mathrm{T}}\dot{\lambda}\sin(\alpha)(1-\cos(\alpha)) + \ddot{y}^{\mathrm{T}}\ddot{\lambda}(1-\cos(\alpha))^2$$

$$=y^{\mathrm{T}}\lambda - (y^{\mathrm{T}}\dot{\lambda} + \lambda^{\mathrm{T}}\dot{y})\sin(\alpha) + (y^{\mathrm{T}}\ddot{\lambda} + \lambda^{\mathrm{T}}\ddot{y})(1-\cos(\alpha))$$

$$\quad - (\dot{y}^{\mathrm{T}}\ddot{\lambda} + \dot{\lambda}^{\mathrm{T}}\ddot{y})\sin(\alpha)(1-\cos(\alpha)) + \dot{y}^{\mathrm{T}}\dot{\lambda}\sin^2(\alpha) + \ddot{y}^{\mathrm{T}}\ddot{\lambda}(1-\cos(\alpha))^2$$

$$=y^{\mathrm{T}}\lambda(1-\sin(\alpha)) - 2\dot{y}^{\mathrm{T}}\dot{\lambda}(1-\cos(\alpha))$$

$$\quad - (\dot{y}^{\mathrm{T}}\ddot{\lambda} + \dot{\lambda}^{\mathrm{T}}\ddot{y})\sin(\alpha)(1-\cos(\alpha))$$

$$\quad + \dot{y}^{\mathrm{T}}\dot{\lambda}(1-\cos^2(\alpha)) + \ddot{y}^{\mathrm{T}}\ddot{\lambda}(1-\cos(\alpha))^2$$

$$=y^{\mathrm{T}}\lambda(1-\sin(\alpha)) + (\ddot{y}^{\mathrm{T}}\ddot{\lambda} - \dot{y}^{\mathrm{T}}\dot{\lambda})(1-\cos(\alpha))^2 - (\dot{y}^{\mathrm{T}}\ddot{\lambda} + \dot{\lambda}^{\mathrm{T}}\ddot{y})\sin(\alpha)(1-\cos(\alpha)). \tag{98}$$

Using (35), (37), (31), (32), and a similar derivation of (98), we have

$$z^{\mathrm{T}}(\alpha)\gamma(\alpha) = z^{\mathrm{T}}\gamma(1-\sin(\alpha)) + (\ddot{z}^{\mathrm{T}}\ddot{\gamma} - \dot{z}^{\mathrm{T}}\dot{\gamma})(1-\cos(\alpha))^2 - (\dot{z}^{\mathrm{T}}\ddot{\gamma} + \dot{\gamma}^{\mathrm{T}}\ddot{z})\sin(\alpha)(1-\cos(\alpha)). \tag{99}$$

Combining (98) and (99) gives

$$2n\mu(\alpha) = p^{\mathrm{T}}(\alpha)\omega(\alpha)$$

$$=y^{\mathrm{T}}(\alpha)\lambda(\alpha) + z^{\mathrm{T}}(\alpha)\gamma(\alpha)$$

$$=(y^{\mathrm{T}}\lambda + z^{\mathrm{T}}\gamma)(1-\sin(\alpha)) + (\ddot{y}^{\mathrm{T}}\ddot{\lambda} + \ddot{z}^{\mathrm{T}}\ddot{\gamma} - \dot{y}^{\mathrm{T}}\dot{\lambda} - \dot{z}^{\mathrm{T}}\dot{\gamma})(1-\cos(\alpha))^2$$

$$\quad - (\dot{y}^{\mathrm{T}}\ddot{\lambda} + \dot{z}^{\mathrm{T}}\ddot{\gamma} + \ddot{y}^{\mathrm{T}}\dot{\lambda} + \ddot{z}^{\mathrm{T}}\dot{\gamma})\sin(\alpha)(1-\cos(\alpha))$$

$$=(y^{\mathrm{T}}\lambda + z^{\mathrm{T}}\gamma)(1-\sin(\alpha)) + (\ddot{x}^{\mathrm{T}}(\ddot{\gamma} - \ddot{\lambda}) - \dot{x}^{\mathrm{T}}(\dot{\gamma} - \dot{\lambda}))(1-\cos(\alpha))^2 \qquad\text{use (30)}$$

$$\quad - (\dot{x}^{\mathrm{T}}(\ddot{\gamma} - \ddot{\lambda}) + \ddot{x}^{\mathrm{T}}(\dot{\gamma} - \dot{\lambda}))\sin(\alpha)(1-\cos(\alpha)) \tag{100}$$

$$\leq (y^{\mathrm{T}}\lambda + z^{\mathrm{T}}\gamma)\,(1-\sin(\alpha)) + (\ddot{x}^{\mathrm{T}}H\ddot{x} - \dot{x}^{\mathrm{T}}H\dot{x})(1-\cos(\alpha))^2 \qquad\text{use (44) in Lemma 5.2}$$

$$\quad + \dot{x}^{\mathrm{T}}H\dot{x}(1-\cos(\alpha))^2 + \ddot{x}^{\mathrm{T}}H\ddot{x}\sin^2(\alpha)$$

$$=(y^{\mathrm{T}}\lambda + z^{\mathrm{T}}\gamma)\,(1-\sin(\alpha)) + \ddot{x}^{\mathrm{T}}H\ddot{x}(1-\cos(\alpha))^2 + \ddot{x}^{\mathrm{T}}H\ddot{x}\sin^2(\alpha).$$

Dividing the both side by $2n$ proves the second inequality of the lemma. Combining (100) and (45) proves the first inequality of the lemma. ∎

32

Proof of Lemma 5.6:

From the second inequality of (54), we have

$$\mu(\alpha) - \mu \le \mu \sin(\alpha) \left(-1 + \frac{\ddot{x}^{\mathrm{T}} H \ddot{x}}{2n\mu} \sin(\alpha) + \frac{\ddot{x}^{\mathrm{T}} H \ddot{x}}{2n\mu} \sin^3(\alpha) \right).$$

Clearly, if $\frac{\ddot{x}^{\mathrm{T}} H \ddot{x}}{2n\mu} \le \frac{1}{2}$, for any $\alpha \in [0, \frac{\pi}{2}]$, the function

$$f(\alpha) := \left(-1 + \frac{\ddot{x}^{\mathrm{T}} H \ddot{x}}{2n\mu} \sin(\alpha) + \frac{\ddot{x}^{\mathrm{T}} H \ddot{x}}{2n\mu} \sin^3(\alpha) \right) \le 0,$$

and $\mu(\alpha) \le \mu$. If $\frac{\ddot{x}^{\mathrm{T}} H \ddot{x}}{2n\mu} > \frac{1}{2}$, using Lemma 2.5, the function f has one real solution $\sin(\alpha) \in (0, 1)$. The solution is given as

$$\sin(\hat{\alpha}) = \sqrt[3]{\frac{n\mu}{\ddot{x}^{\mathrm{T}} H \ddot{x}} + \sqrt{\left(\frac{n\mu}{\ddot{x}^{\mathrm{T}} H \ddot{x}} \right)^2 + \left(\frac{1}{3} \right)^3}} + \sqrt[3]{\frac{n\mu}{\ddot{x}^{\mathrm{T}} H \ddot{x}} - \sqrt{\left(\frac{n\mu}{\ddot{x}^{\mathrm{T}} H \ddot{x}} \right)^2 + \left(\frac{1}{3} \right)^3}}.$$

This proves the Lemma. ∎

Proof of Lemma 5.7:

Since $\sin(\tilde{\alpha})$ is the only positive real solution of (56) in $[0, 1]$ and $q(0) < 0$, substituting a_0, a_1, a_2, a_3 and a_4 into (56), we have, for all $\sin(\alpha) \le \sin(\tilde{\alpha})$,

$$\left(\left\| \ddot{p} \circ \ddot{\omega} - \dot{\omega} \circ \dot{p} - \frac{1}{2n} (\ddot{p}^{\mathrm{T}} \ddot{\omega} - \dot{\omega}^{\mathrm{T}} \dot{p}) e \right\| \right) \sin^4(\alpha) + \left(\left\| \dot{p} \circ \ddot{\omega} + \dot{\omega} \circ \ddot{p} - \frac{1}{2n} (\dot{p}^{\mathrm{T}} \ddot{\omega} + \dot{\omega}^{\mathrm{T}} \ddot{p}) e \right\| \right) \sin^3(\alpha)$$

$$\le - \left(2\theta \frac{\dot{p}^{\mathrm{T}} \dot{\omega}}{2n} \right) \sin^4(\alpha) - \left(2\theta \frac{\dot{p}^{\mathrm{T}} \dot{\omega}}{2n} \right) \sin^2(\alpha) + \theta \mu (1 - \sin(\alpha)). \tag{101}$$

33

Using (38), (39), (31), (32), (58), Lemma 2.2, (101), and the first inequality of (54), we have

$$
\begin{aligned}
&\left\| p(\alpha) \circ \omega(\alpha) - \mu(\alpha)e \right\| \\
={}&\left\| \left(p - \dot{p}\sin(\alpha) + \ddot{p}(1 - \cos(\alpha))\right) \circ \left(\omega - \dot{\omega}\sin(\alpha) + \ddot{\omega}(1 - \cos(\alpha))\right) - \mu(\alpha)e \right\| \\
={}&\left\| (p \circ \omega - \mu e)(1 - \sin(\alpha)) + \left(\ddot{p} \circ \ddot{\omega} - \dot{p} \circ \dot{\omega} - \frac{1}{2n}(\ddot{p}^{\mathrm{T}}\ddot{\omega} - \dot{p}^{\mathrm{T}}\dot{\omega})e\right)(1 - \cos(\alpha))^2 \right. \\
&\left. - \left(\dot{p} \circ \ddot{\omega} + \dot{\omega} \circ \ddot{p} - \frac{1}{2n}(\dot{p}^{\mathrm{T}}\ddot{\omega} + \ddot{p}^{\mathrm{T}}\dot{\omega})e\right)\sin(\alpha)(1 - \cos(\alpha)) \right\| \\
\leq{}&(1 - \sin(\alpha))\left\| p \circ \omega - \mu e \right\| + \left\|(\ddot{p} \circ \ddot{\omega} - \dot{p} \circ \dot{\omega} - \frac{1}{2n}(\ddot{p}^{\mathrm{T}}\ddot{\omega} - \dot{p}^{\mathrm{T}}\dot{\omega}))e\right\|(1 - \cos(\alpha))^2 \\
&+ \left\|(\dot{p} \circ \ddot{\omega} + \dot{\omega} \circ \ddot{p} - \frac{1}{2n}(\dot{p}^{\mathrm{T}}\ddot{\omega} + \ddot{p}^{\mathrm{T}}\dot{\omega})e\right\|\sin(\alpha)(1 - \cos(\alpha)) \qquad\qquad (102) \\
\leq{}&\theta\mu(1 - \sin(\alpha)) + \left\|(\ddot{p} \circ \ddot{\omega} - \dot{p} \circ \dot{\omega} - \frac{1}{2n}(\ddot{p}^{\mathrm{T}}\ddot{s} - \dot{p}^{\mathrm{T}}\dot{\omega}))e\right\|\sin^4(\alpha) + a_3\sin^3(\alpha) \\
\leq{}&2\theta\mu(1 - \sin(\alpha)) - \left(2\theta\frac{\dot{p}^{\mathrm{T}}\dot{\omega}}{2n}\right)(\sin^4(\alpha) + \sin^2(\alpha)) \\
\leq{}&2\theta\left(\mu(1 - \sin(\alpha)) - \left(\frac{\dot{x}^{\mathrm{T}}H\dot{x}}{2n}\right)\left((1 - \cos(\alpha))^2 + \sin^2(\alpha)\right)\right) \\
\leq{}&2\theta\mu(\alpha). \qquad\qquad (103)
\end{aligned}
$$

Hence, the point $(x(\alpha), p(\alpha), \omega(\alpha))$ satisfies the proximity condition for $\mathcal{N}_2(2\theta)$. To check the positivity condition $(p(\alpha), \omega(\alpha)) > 0$, note that the initial condition $(p, \omega) > 0$. It follows from (103) and Corollary 5.1 that, for $\sin(\alpha) \leq \sin(\bar{\alpha})$ and $\theta < 0.5$,

$$
p_i(\alpha)\omega_i(\alpha) \geq (1 - 2\theta)\mu(\alpha) > 0. \qquad\qquad (104)
$$

Therefore, we cannot have $p_i(\alpha) = 0$ or $\omega_i(\alpha) = 0$ for any index i when $\alpha \in [0, \sin^{-1}(\bar{\alpha})]$. This proves $(p(\alpha), \omega(\alpha)) > 0$. ■

Proof of Lemma 5.8:

Since

$$
\left\|\frac{\dot{p}}{p}\right\|^2 = \sum_{i=1}^{2n}\left(\frac{\dot{p}_i}{p_i}\right)^2, \quad \left\|\frac{\dot{\omega}}{\omega}\right\|^2 = \sum_{i=1}^{2n}\left(\frac{\dot{\omega}_i}{\omega_i}\right)^2,
$$

From Lemma 5.3 and (25), we have

$$\left(\frac{n}{1-\theta}\right)^2$$

$$\geq \;\; \left\|\frac{\dot{p}}{p}\right\|^2\left\|\frac{\dot{\omega}}{\omega}\right\|^2 = \left(\sum_{i=1}^{2n}\left(\frac{\dot{p}_i}{p_i}\right)^2\right)\left(\sum_{i=1}^{2n}\left(\frac{\dot{\omega}_i}{\omega_i}\right)^2\right)$$

$$\geq \;\; \sum_{i=1}^{2n}\left(\frac{\dot{p}_i}{p_i}\frac{\dot{\omega}_i}{\omega_i}\right)^2 = \left\|\frac{\dot{p}}{p}\circ\frac{\dot{\omega}}{\omega}\right\|^2$$

$$\geq \;\; \sum_{i=1}^{2n}\left(\frac{\dot{p}_i\dot{\omega}_i}{(1+\theta)\mu}\right)^2 = \frac{1}{(1+\theta)^2\mu^2}\left\|\dot{p}\circ\dot{\omega}\right\|^2,$$

i.e.,

$$\left\|\dot{p}\circ\dot{\omega}\right\|^2 \leq \left(\frac{1+\theta}{1-\theta}n\mu\right)^2$$

This proves (59). Using

$$\left\|\frac{\ddot{p}}{p}\right\|^2 = \sum_{i=1}^{2n}\left(\frac{\ddot{p}_i}{p_i}\right)^2, \quad \left\|\frac{\ddot{\omega}}{\omega}\right\|^2 = \sum_{i=1}^{2n}\left(\frac{\ddot{\omega}_i}{\omega_i}\right)^2,$$

and Lemma 5.4, then following the same procedure, it is easy to verify (60). From (47) and (50), we have

$$\left(\frac{2n}{(1-\theta)}\right)\left(\frac{4(1+\theta)n^2}{(1-\theta)^3}\right) \geq \left(\left\|\frac{\dot{p}}{p}\right\|^2 + \left\|\frac{\dot{\omega}}{\omega}\right\|^2\right)\left(\left\|\frac{\ddot{p}}{p}\right\|^2 + \left\|\frac{\ddot{\omega}}{\omega}\right\|^2\right)$$

$$\geq \;\; \left\|\frac{\ddot{p}}{p}\right\|^2\left\|\frac{\dot{\omega}}{\omega}\right\|^2 + \left\|\frac{\dot{p}}{p}\right\|^2\left\|\frac{\ddot{\omega}}{\omega}\right\|^2$$

$$= \;\; \left(\sum_{i=1}^{2n}\left(\frac{\ddot{p}_i}{p_i}\right)^2\right)\left(\sum_{i=1}^{2n}\left(\frac{\dot{\omega}_i}{\omega_i}\right)^2\right) + \left(\sum_{i=1}^{2n}\left(\frac{\dot{p}_i}{p_i}\right)^2\right)\left(\sum_{i=1}^{2n}\left(\frac{\ddot{\omega}_i}{\omega_i}\right)^2\right)$$

$$\geq \;\; \sum_{i=1}^{2n}\left(\frac{\ddot{p}_i\dot{\omega}_i}{p_i\omega_i}\right)^2 + \sum_{i=1}^{2n}\left(\frac{\dot{p}_i\ddot{\omega}_i}{p_i\omega_i}\right)^2$$

$$\geq \;\; \sum_{i=1}^{2n}\left(\frac{\ddot{p}_i\dot{\omega}_i}{(1+\theta)\mu}\right)^2 + \sum_{i=1}^{2n}\left(\frac{\dot{p}_i\ddot{\omega}_i}{(1+\theta)\mu}\right)^2$$

$$= \;\; \frac{1}{(1+\theta)^2\mu^2}\left(\left\|\ddot{p}\circ\dot{\omega}\right\|^2 + \left\|\dot{p}\circ\ddot{\omega}\right\|^2\right),$$

$$(105)$$

i.e.,

$$\left\|\ddot{p}\circ\dot{\omega}\right\|^2 + \left\|\dot{p}\circ\ddot{\omega}\right\|^2 \leq \frac{(2n)^3(1+\theta)^3}{(1-\theta)^4}\mu^2.$$

This proves the lemma. ∎

Proof of Lemma 5.9:

First notice that $q(\sin(\alpha))$ is a monotonic increasing function of $\sin(\alpha)$ for $\alpha \in [0, \frac{\pi}{2}]$ and $q(\sin(0)) < 0$,

35

therefore, we need only to show that $q(\frac{\theta}{\sqrt{n}}) < 0$ for $\theta \leq 0.22$. Using Lemma 2.6, we have

$$\left\| \dot{p} \circ \ddot{\omega} + \dot{\omega} \circ \ddot{p} - \frac{1}{2n}(\dot{p}^{\mathrm{T}}\ddot{\omega} + \dot{\omega}^{\mathrm{T}}\ddot{p})e \right\| \leq \left\| \dot{p} \circ \ddot{\omega} \right\| + \left\| \dot{\omega} \circ \ddot{p} \right\|,$$

$$\left\| \ddot{p} \circ \ddot{\omega} - \dot{\omega} \circ \dot{p} - \frac{1}{2n}(\ddot{p}^{\mathrm{T}}\ddot{\omega} - \dot{\omega}^{\mathrm{T}}\dot{p})e \right\| \leq \left\| \ddot{p} \circ \ddot{\omega} \right\| + \left\| \dot{\omega} \circ \dot{p} \right\|.$$

In view of Lemmas 5.8, 5.3, and 5.4, from (56), we have, for $\alpha \in [0, \frac{\pi}{2}]$,

$$q(\sin(\alpha)) \leq \left(\left\| \ddot{p} \circ \ddot{\omega} \right\| + \left\| \dot{\omega} \circ \dot{p} \right\| + 2\theta\frac{\dot{p}^{\mathrm{T}}\dot{\omega}}{2n} \right) \sin^4(\alpha) + \left(\left\| \dot{p} \circ \ddot{\omega} \right\| + \left\| \dot{\omega} \circ \ddot{p} \right\| \right) \sin^3(\alpha)$$

$$+ 2\theta\frac{\dot{p}^{\mathrm{T}}\dot{\omega}}{2n} \sin^2(\alpha) + \theta\mu\sin(\alpha) - \theta\mu$$

$$\leq \mu\bigg(\left(\frac{2(1+\theta)^2}{(1-\theta)^3}n^2 + \frac{n(1+\theta)}{(1-\theta)} + \frac{\theta(1+\theta)}{(1-\theta)} \right) \sin^4(\alpha) + 4\sqrt{2}\frac{(1+\theta)^{\frac{3}{2}}}{(1-\theta)^2}n^{\frac{3}{2}} \sin^3(\alpha)$$

$$+ \frac{\theta(1+\theta)}{(1-\theta)} \sin^2(\alpha) + \theta\sin(\alpha) - \theta \bigg).$$

Since $n \geq 1$ and $\theta > 0$, substituting $\sin(\alpha) = \frac{\theta}{\sqrt{n}}$ gives

$$q\left(\frac{\theta}{\sqrt{n}}\right) \leq \mu\bigg(\left(\frac{2(1+\theta)^2}{(1-\theta)^3}n^2 + \frac{n(1+\theta)}{(1-\theta)} + \frac{\theta(1+\theta)}{(1-\theta)} \right) \frac{\theta^4}{n^2} + 4\sqrt{2}\frac{(1+\theta)^{\frac{3}{2}}n^{\frac{3}{2}}}{(1-\theta)^2} \frac{\theta^3}{n^{\frac{3}{2}}}$$

$$+ \frac{\theta(1+\theta)}{(1-\theta)} \frac{\theta^2}{n} + \theta\frac{\theta}{\sqrt{n}} - \theta \bigg)$$

$$= \theta\mu\bigg(\frac{2\theta^3(1+\theta)^2}{(1-\theta)^3} + \frac{\theta^3(1+\theta)}{n(1-\theta)} + \frac{\theta^4(1+\theta)}{(1-\theta)n^2}$$

$$+ \frac{4\sqrt{2}\theta^2(1+\theta)^{\frac{3}{2}}}{(1-\theta)^2} + \frac{\theta^2(1+\theta)}{n(1-\theta)} + \frac{\theta}{\sqrt{n}} - 1 \bigg)$$

$$\leq \theta\mu\bigg(\frac{2\theta^3(1+\theta)^2}{(1-\theta)^3} + \frac{\theta^3(1+\theta)}{(1-\theta)} + \frac{\theta^4(1+\theta)}{(1-\theta)}$$

$$+ \frac{4\sqrt{2}\theta^2(1+\theta)^{\frac{3}{2}}}{(1-\theta)^2} + \frac{\theta^2(1+\theta)}{(1-\theta)} + \theta - 1 \bigg) := \theta\mu p(\theta). \tag{106}$$

Since $p(\theta)$ is monotonic increasing function of $\theta \in [0, 1)$, $p(0) < 0$, and it is easy to verify that $p(0.22) < 0$, this proves the lemma. ∎

Proof of Lemma 5.10:

Using Lemma 2.6, we have

$$0 \leq \left\| \Delta p \circ \Delta\omega - \frac{1}{2n}(\Delta p^{\mathrm{T}}\Delta\omega)e \right\|^2 \leq \|\Delta p \circ \Delta\omega\|^2. \tag{107}$$

Pre-multiplying $\left(P(\alpha)\Omega(\alpha) \right)^{-\frac{1}{2}}$ on the both sides of (66) yields

$$D\Delta\omega + D^{-1}\Delta p = \left(P(\alpha)\Omega(\alpha) \right)^{-\frac{1}{2}} \left(\mu(\alpha)e - P(\alpha)\Omega(\alpha)e \right).$$

Let $u = D\Delta\omega$, $v = D^{-1}\Delta p$, from (63), we have

$$u^{\mathrm{T}}v = \Delta p^{\mathrm{T}}\Delta\omega = \Delta y^{\mathrm{T}}\Delta\lambda + \Delta z^{\mathrm{T}}\Delta\gamma = \Delta x^{\mathrm{T}}(\Delta\gamma - \Delta\lambda) = \Delta x^{\mathrm{T}}H\Delta x \geq 0. \tag{108}$$

Use Lemma 2.4 and the assumption of $(x(\alpha), p(\alpha), \omega(\alpha)) \in \mathcal{N}_2(2\theta)$, we have

$$
\begin{aligned}
\left\|\Delta p \circ \Delta\omega\right\| &= \left\|u \circ v\right\| \leq 2^{-\frac{3}{2}}\left\|\left(P(\alpha)\Omega(\alpha)\right)^{-\frac{1}{2}}\left(\mu(\alpha)e - P(\alpha)\Omega(\alpha)e\right)\right\|^2 \\
&= 2^{-\frac{3}{2}}\sum_{i=1}^{2n}\frac{(\mu(\alpha) - p_i(\alpha)\omega_i(\alpha))^2}{p_i(\alpha)\omega_i(\alpha)} \\
&\leq 2^{-\frac{3}{2}}\frac{\|\mu(\alpha)e - p(\alpha)\circ\omega(\alpha)\|^2}{\min_i p_i(\alpha)\omega_i(\alpha)} \\
&\leq 2^{-\frac{3}{2}}\frac{(2\theta)^2\mu(\alpha)^2}{(1 - 2\theta)\mu(\alpha)} = 2^{\frac{1}{2}}\frac{\theta^2\mu(\alpha)}{(1 - 2\theta)}.
\end{aligned} \tag{109}
$$

Define $(p^{k+1}(t), \omega^{k+1}(t)) = (p(\alpha), \omega(\alpha)) + t(\Delta p, \Delta\omega)$. From (66) and (40), we have

$$p(\alpha)^{\mathrm{T}}\Delta\omega + \omega(\alpha)^{\mathrm{T}}\Delta p = 2n\mu - \sum_{i=1}^{2n}p_i(\alpha)\omega_i(\alpha) = 0. \tag{110}$$

Therefore,

$$\mu^{k+1}(t) = \frac{\left(p(\alpha) + t\Delta p\right)^{\mathrm{T}}\left(\omega(\alpha) + t\Delta\omega\right)}{2n} = \frac{p(\alpha)^{\mathrm{T}}\omega(\alpha) + t^2\Delta p^{\mathrm{T}}\Delta\omega}{2n} = \mu(\alpha) + t^2\frac{\Delta p^{\mathrm{T}}\Delta\omega}{2n}. \tag{111}$$

Since $\Delta p^{\mathrm{T}}\Delta\omega = \Delta x^{\mathrm{T}}H\Delta x \geq 0$, we conclude that $\mu^{k+1}(t) \geq \mu(\alpha)$. Using (111), (66), (107), and (109), we have

$$
\begin{aligned}
&\left\|p^{k+1}(t) \circ \omega^{k+1}(t) - \mu^{k+1}(t)e\right\| \\
=\ & \left\|(p(\alpha) + t\Delta p) \circ (\omega(\alpha) + t\Delta\omega) - \mu(\alpha)e - \frac{t^2}{2n}\left(\Delta p^{\mathrm{T}}\Delta\omega\right)e\right\| \\
=\ & \left\|p(\alpha) \circ \omega(\alpha) + t(\omega(\alpha) \circ \Delta p + p(\alpha) \circ \Delta\omega) + t^2\Delta p \circ \Delta\omega - \mu(\alpha)e - \frac{t^2}{2n}\left(\Delta p^{\mathrm{T}}\Delta\omega\right)e\right\| \\
=\ & \left\|p(\alpha) \circ \omega(\alpha) + t(\mu(\alpha)e - p(\alpha) \circ \omega(\alpha)) + t^2\Delta p \circ \Delta\omega - \mu(\alpha)e - \frac{t^2}{2n}\left(\Delta p^{\mathrm{T}}\Delta\omega\right)e\right\| \\
=\ & \left\|(1 - t)\left(p(\alpha) \circ \omega(\alpha) - \mu(\alpha)e\right) + t^2\left(\Delta p \circ \Delta\omega - \frac{1}{2n}\left(\Delta p^{\mathrm{T}}\Delta\omega\right)e\right)\right\| \\
\leq\ & (1 - t)(2\theta)\mu(\alpha) + t^2\frac{2^{\frac{1}{2}}\theta^2}{(1 - 2\theta)}\mu(\alpha) \\
\leq\ & \left((1 - t)(2\theta) + t^2\frac{2^{\frac{1}{2}}\theta^2}{(1 - 2\theta)}\right)\mu^{k+1} := f(t, \theta)\mu^{k+1}.
\end{aligned} \tag{112}
$$

Therefore, taking $t = 1$ gives $\left\|p^{k+1} \circ \omega^{k+1} - \mu^{k+1}e\right\| \leq \frac{2^{\frac{1}{2}}\theta^2}{(1-2\theta)}\mu^{k+1}$. It is easy to see that, for $\theta \leq 0.29$,

$$\frac{2^{\frac{1}{2}}\theta^2}{(1 - 2\theta)} = 0.2832 < \theta.$$

For $\theta \leq 0.29$ and $t \in [0,1]$, noticing $0 \leq f(t,\theta) \leq f(t,0.29) \leq 0.58(1-t) + 0.2832t^2 < 1$, and using Corollary 5.1, we have, for an additional condition $\sin(\alpha) \leq \sin^{-1}(\bar{\alpha})$,

$$
\begin{aligned}
p_i^{k+1}(t)\omega_i^{k+1}(t) &\geq (1 - f(t,\theta))\,\mu^{k+1}(t) \\
&= (1 - f(t,\theta))\left(\mu(\alpha) + \frac{t^2}{n}\Delta p^{\mathrm{T}}\Delta\omega\right) \\
&\geq (1 - f(t,\theta))\,\mu(\alpha) \\
&> 0,
\end{aligned}
\tag{113}
$$

Therefore, $(p^{k+1}(t), \omega^{k+1}(t)) > 0$ for $t \in [0,1]$, i.e., $(p^{k+1}, \omega^{k+1}) > 0$. This finishes the proof. $\blacksquare$

Proof of Lemma 5.11:

The first inequality of (67) follows from (108). Pre-multiplying both sides of (66) by $P^{-\frac{1}{2}}(\alpha)\Omega^{-\frac{1}{2}}(\alpha)$ gives

$$
P^{-\frac{1}{2}}(\alpha)\Omega^{\frac{1}{2}}(\alpha)\Delta p + P^{\frac{1}{2}}(\alpha)\Omega^{-\frac{1}{2}}(\alpha)\Delta\omega = P^{-\frac{1}{2}}(\alpha)\Omega^{-\frac{1}{2}}(\alpha)\Big(\mu(\alpha)e - P(\alpha)\Omega(\alpha)e\Big).
$$

Let $u = P^{-\frac{1}{2}}(\alpha)\Omega^{\frac{1}{2}}(\alpha)\Delta p$ and $v = P^{\frac{1}{2}}(\alpha)\Omega^{-\frac{1}{2}}(\alpha)\Delta\omega$, and $w = P^{-\frac{1}{2}}(\alpha)\Omega^{-\frac{1}{2}}(\alpha)\Big(\mu(\alpha)e - P(\alpha)\Omega(\alpha)e\Big)$, from (108), we have $u^{\mathrm{T}}v = \Delta p^{\mathrm{T}}\Delta\omega \geq 0$. Using Lemma 2.3 and the assumption of $(x(\alpha), p(\alpha), \omega(\alpha)) \in \mathcal{N}_2(2\theta)$, we have

$$
\begin{aligned}
\|u\|^2 + \|v\|^2 &= \sum_{i=1}^{2n}\left(\frac{(\Delta p_i)^2\omega_i(\alpha)}{p_i(\alpha)} + \frac{(\Delta\omega_i)^2 p_i(\alpha)}{\omega_i(\alpha)}\right) \\
&\leq \|w\|^2 = \sum_{i=1}^{2n}\frac{(\mu(\alpha) - p_i(\alpha)\omega_i(\alpha))^2}{p_i(\alpha)\omega_i(\alpha)} \\
&\leq \frac{\sum_{i=1}^{2n}(\mu(\alpha) - p_i(\alpha)\omega_i(\alpha))^2}{\min_i p_i(\alpha)\omega_i(\alpha)} \\
&\leq \frac{(2\theta)^2\mu^2(\alpha)}{(1-2\theta)\mu(\alpha)} = \frac{(2\theta)^2\mu(\alpha)}{(1-2\theta)}.
\end{aligned}
\tag{114}
$$

Dividing both sides by $\mu(\alpha)$ and using $p_i(\alpha)\omega_i(\alpha) \geq \mu(\alpha)(1-2\theta)$ yields

$$
\begin{aligned}
\sum_{i=1}^{2n}(1-2\theta)&\left(\frac{(\Delta p_i)^2}{p_i^2(\alpha)} + \frac{(\Delta\omega_i)^2}{\omega_i^2(\alpha)}\right) \\
&= (1-2\theta)\left(\left\|\frac{\Delta p}{p(\alpha)}\right\|^2 + \left\|\frac{\Delta\omega}{\omega(\alpha)}\right\|^2\right) \\
&\leq \frac{(2\theta)^2}{(1-2\theta)},
\end{aligned}
\tag{115}
$$

i.e.,

$$\left\|\frac{\Delta p}{p(\alpha)}\right\|^2 + \left\|\frac{\Delta \omega}{\omega(\alpha)}\right\|^2 \le \left(\frac{2\theta}{1-2\theta}\right)^2. \tag{116}$$

Invoking Lemma 2.1, we have

$$\left\|\frac{\Delta p}{p(\alpha)}\right\|^2 \cdot \left\|\frac{\Delta \omega}{\omega(\alpha)}\right\|^2 \le \frac{1}{4}\left(\frac{2\theta}{1-2\theta}\right)^4. \tag{117}$$

This gives

$$\left\|\frac{\Delta p}{p(\alpha)}\right\| \cdot \left\|\frac{\Delta \omega}{\omega(\alpha)}\right\| \le \frac{2\theta^2}{(1-2\theta)^2}. \tag{118}$$

Using Cauchy-Schwarz inequality, we have

$$\frac{(\Delta p)^{\mathrm{T}}(\Delta \omega)}{\mu(\alpha)}$$

$$\le \sum_{i=1}^{2n} \frac{|\Delta p_i||\Delta \omega_i|}{\mu(\alpha)}$$

$$\le (1+2\theta)\sum_{i=1}^{2n} \frac{|\Delta p_i|}{p_i(\alpha)}\frac{|\Delta \omega_i|}{\omega_i(\alpha)}$$

$$= (1+2\theta)\left|\frac{\Delta p}{p(\alpha)}\right|^{\mathrm{T}}\left|\frac{\Delta \omega}{\omega(\alpha)}\right|$$

$$\le (1+2\theta)\left\|\frac{\Delta p}{p(\alpha)}\right\| \cdot \left\|\frac{\Delta \omega}{\omega(\alpha)}\right\|$$

$$\le \frac{2\theta^2(1+2\theta)}{(1-2\theta)^2}. \tag{119}$$

Therefore,

$$\frac{(\Delta p)^{\mathrm{T}}(\Delta \omega)}{2n} \le \frac{\theta^2(1+2\theta)}{n(1-2\theta)^2}\mu(\alpha). \tag{120}$$

This proves the lemma. ∎

Proof of Lemma 5.13:

Using Lemmas 5.12, 5.5, 2.2, 5.2, 5.3, and 5.4, and noticing $\ddot{p}^{\mathrm{T}}\ddot{\omega} \ge 0$ and $\dot{p}^{\mathrm{T}}\dot{\omega} \ge 0$, we have

$$\mu^{k+1} \le \mu(\alpha)\left(1+\frac{\theta^2(1+2\theta)}{n(1-2\theta)^2}\right) = \mu(\alpha)\left(1+\frac{\delta_0}{n}\right) \tag{121a}$$

$$=\mu^k\left(1-\sin(\alpha)+\left(\frac{\ddot{p}^{\mathrm{T}}\ddot{\omega}}{2n\mu}-\frac{\dot{p}^{\mathrm{T}}\dot{\omega}}{2n\mu}\right)(1-\cos(\alpha))^2 - \left(\frac{\dot{p}^{\mathrm{T}}\ddot{\omega}}{2n\mu}+\frac{\dot{\omega}^{\mathrm{T}}\ddot{p}}{2n\mu}\right)\sin(\alpha)(1-\cos(\alpha))\right)\left(1+\frac{\delta_0}{n}\right)$$

$$\le\mu^k\left(1-\sin(\alpha)+\frac{\ddot{p}^{\mathrm{T}}\ddot{\omega}}{2n\mu}\sin^4(\alpha)+\left(\left|\frac{\dot{p}^{\mathrm{T}}\ddot{\omega}}{2n\mu}\right|+\left|\frac{\dot{\omega}^{\mathrm{T}}\ddot{p}}{2n\mu}\right|\right)\sin^3(\alpha)\right)\left(1+\frac{\delta_0}{n}\right)$$

$$\le\mu^k\left(1-\sin(\alpha)+\frac{n(1+\theta)^2}{(1-\theta)^3}\sin^4(\alpha)+\frac{2(2n)^{\frac{1}{2}}(1+\theta)^{\frac{3}{2}}}{(1-\theta)^2}\sin^3(\alpha)\right)\left(1+\frac{\delta_0}{n}\right) \tag{121b}$$

Substituting $\sin(\alpha) = \frac{\theta}{\sqrt{n}}$ into (121b) gives

$$\mu^{k+1} \leq \mu^k \left(1 - \frac{\theta}{\sqrt{n}} + \frac{n(1+\theta)^2}{(1-\theta)^3}\frac{\theta^4}{n^2} + \frac{2(2n)^{\frac{1}{2}}(1+\theta)^{\frac{3}{2}}}{(1-\theta)^2}\frac{\theta^3}{n^{\frac{3}{2}}}\right)\left(1 + \frac{\delta_0}{n}\right)$$

$$= \mu^k \left(1 - \frac{\theta}{\sqrt{n}} + \frac{\theta^4(1+\theta)^2}{n(1-\theta)^3} + \frac{2^{\frac{3}{2}}\theta^3(1+\theta)^{\frac{3}{2}}}{n(1-\theta)^2}\right)\left(1 + \frac{\delta_0}{n}\right)$$

$$= \mu^k \left(1 - \frac{\theta}{\sqrt{n}} + \frac{\delta_0}{n} + \frac{\theta^4(1+\theta)^2}{n(1-\theta)^3} + \frac{2^{\frac{3}{2}}\theta^3(1+\theta)^{\frac{3}{2}}}{n(1-\theta)^2} - \frac{\theta\delta_0}{n^{\frac{3}{2}}} + \frac{\delta_0}{n}\left[\frac{\theta^4(1+\theta)^2}{n(1-\theta)^3} + \frac{2^{\frac{3}{2}}\theta^3(1+\theta)^{\frac{3}{2}}}{n(1-\theta)^2}\right]\right)$$

$$= \mu^k \left(1 - \frac{\theta}{\sqrt{n}}\left[1 - \frac{\delta_0}{\sqrt{n}\theta} - \frac{\theta^3(1+\theta)^2}{\sqrt{n}(1-\theta)^3} - \frac{2^{\frac{3}{2}}\theta^2(1+\theta)^{\frac{3}{2}}}{\sqrt{n}(1-\theta)^2}\right] - \frac{\theta\delta_0}{n^{\frac{3}{2}}}\left[1 - \frac{\theta^3(1+\theta)^2}{\sqrt{n}(1-\theta)^3} - \frac{2^{\frac{3}{2}}\theta^2(1+\theta)^{\frac{3}{2}}}{\sqrt{n}(1-\theta)^2}\right]\right)$$

Since

$$1 - \frac{\theta^3(1+\theta)^2}{\sqrt{n}(1-\theta)^3} - \frac{2^{\frac{3}{2}}\theta^2(1+\theta)^{\frac{3}{2}}}{\sqrt{n}(1-\theta)^2} \geq 1 - \frac{\theta^3(1+\theta)^2}{(1-\theta)^3} - \frac{2^{\frac{3}{2}}\theta^2(1+\theta)^{\frac{3}{2}}}{(1-\theta)^2} := f(\theta),$$

where $f(\theta)$ is a monotonic decreasing function of θ, and for $\theta \leq 0.37$, $f(\theta) > 0$. Therefore, for $\theta \leq 0.37$,

$$\mu^{k+1} \leq \mu^k \left(1 - \frac{\theta}{\sqrt{n}}\left[1 - \frac{\delta_0}{\sqrt{n}\theta} - \frac{\theta^3(1+\theta)^2}{\sqrt{n}(1-\theta)^3} - \frac{2^{\frac{3}{2}}\theta^2(1+\theta)^{\frac{3}{2}}}{\sqrt{n}(1-\theta)^2}\right]\right)$$

$$= \mu^k \left(1 - \frac{\theta}{\sqrt{n}}\left[1 - \frac{\theta(1+2\theta)}{\sqrt{n}(1-2\theta)^2} - \frac{\theta^3(1+\theta)^2}{\sqrt{n}(1-\theta)^3} - \frac{2^{\frac{3}{2}}\theta^2(1+\theta)^{\frac{3}{2}}}{\sqrt{n}(1-\theta)^2}\right]\right) \tag{122}$$

Since

$$1 - \frac{\theta(1+2\theta)}{\sqrt{n}(1-2\theta)^2} - \frac{\theta^3(1+\theta)^2}{\sqrt{n}(1-\theta)^3} - \frac{2^{\frac{3}{2}}\theta^2(1+\theta)^{\frac{3}{2}}}{\sqrt{n}(1-\theta)^2} \geq 1 - \frac{\theta(1+2\theta)}{(1-2\theta)^2} - \frac{\theta^3(1+\theta)^2}{(1-\theta)^3} - \frac{2^{\frac{3}{2}}\theta^2(1+\theta)^{\frac{3}{2}}}{(1-\theta)^2} := g(\theta),$$

where $g(\theta)$ is a monotonic decreasing function of θ, and for $\theta \leq 0.19$, $g(\theta) > 0.0976 > 0$. For $\theta = 0.19$, $\theta g(\theta) > 0.0185$ and

$$\mu^{k+1} \leq \mu^k \left(1 - \frac{0.0185}{\sqrt{n}}\right).$$

This proves (69). ∎

References

[1] A. Alessio and A. Bemporad, A survey of explicit model predictive control, eds. L. Magni and D. M. Raimondo and F. Allgower, *Nonlinear model precidive control: toward new challenging applications*, Springer-Verlag, Berlin, 345-369, 2009.

[2] Alberto Bemporad and Manfred Morari and Vivek Dua and Efstratios Pistikopoulos, The explicit linear quadratic regulator for constrained systems, *Automatica*, Vol. 38, 3-20, 2002.

[3] A. B. Berkelaar and K. Roos and T. Terlaky, The optimal set and optimal partition approach to linear and quadratic programming, eds. H. Greenberg and T. Gal, *Recent Advances in Sensitivity Analysis and Parametric Programming*, Kluwer Publishers, Berlin, 1997.

[4] Dimitri P. Bertsekas, Projected Newton methods for optimization problems with simple constraints, *SIAM J. Control and Optimization*, Vol. 20, pp.221-246, 1982.

[5] Coralia Cartis and Nicholas I. M. Gould, Finding a point in the relative interior of a polyhedren, Technical Report, Computing Laboratory, Oxford University, Oxford, UK, 2007.

[6] M. P. do Carmo, Differential Geometry of Curves and Surfaces, Prentice-Hall, New Jersey, 1976.

[7] A.J. Goldman and A.W. Tucker, Theory of linear programming, eds. H.W. Kuhn and Tucker, *Linear Equalities and Related Systems*, Princeton University Press, Princeton, pp. 53-97, 1956.

[8] O. Guler and Y. Ye, Convergence behavior of interior-point algorithms, *Mathematical Programming*, Vol. 60, pp. 215-228, 1993.

[9] C. G. Han and P. Pardalos and Yinyu Ye, Computational aspects of an interior point algorithm for quadratic programming problem with box constraints, eds. T.F. Coleman and Y. Li, *Large-Scale Numerical Optimization*, SIAM Publications, Philadelphia, PA, pp. 92-112, 1990.

[10] Don Herbison-Evans, Solving Quartics and Cubics for Graphics, Technical Report, Basser Department of Computer Science, University of Sydney, Sydney, Australia, 1994.

[11] Tingshu Hu and Zongli Lin, *Control systems with actuator saturation: analysis and design*, Birkhauser, Boston, 2001.

[12] S. Mehrotra, On the implementation of a primal-dual interior point method, *SIAM Journal on Optimization*, Vol. 2, pp. 575-601, 1992.

[13] S. Mizuno and M. Todd and Y. Ye, On adaptive step primal-dual interior-point algorithms for linear programming, *Mathematics of Operations Research*, Vol. 18, pp. 964-981, 1993.

[14] R. Monteiro and I. Adler, Interior path following primal-dual algorithms. Part I: linear programming, *Mathematical Programming*, Vol. 44, pp. 27-41, 1989.

[15] Andrei D. Polyanin and Alexander V. Manzhirov, *Handbook of Mathematics For Engineers and Scientists*, Chapman & Hall/CRC, Noca Raton, FL, 2007.

[16] S. J. Qin and T. A. Badgwell, A survey of industrial model predictive control technology, *Control Engineering Practice*, Vol. 11, pp. 733-764, 2003.

[17] C. V. Rao and S. J. Wright and J. B. Rawling, Application of interior-point methods to model predictive control, *Journal of Optimization Theory and Applications*, Vol. 99, pp. 723-757, 1998.

[18] P. M. Stoltz and S. Sivapiragasam and T. Anthony, Satellite orbit-raising using LQR control with fixed thrusters, *Advances in the Astronautical Sciences*, Vol. 98, pp. 109-120, 1998.

[19] Y. Wang and S. Boyd, Fast Model Predictive Control Using Online Optimization, *IEEE Transactions on control systems technology*, Vol. 18, pp. 267-278, 2010.

[20] S. Wright, *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia, 1997.

[21] S. J. Wright, Interior point methods for optimal control of discrete time systems, *Journal of Optimization Theory and Applications*, Vol. 77, pp. 161-187, 1993.

[22] Y. Yang, Arc-Search Path-Following Interior-Point Algorithms for Linear Programming, *Optimization Online*, August, 2009.

[23] Y. Yang, A Polynomial Arc-Search Interior-Point Algorithm for Linear Programming, *Optimization Online*, November, 2010.

[24] Y. Yang, A Polynomial Arc-Search Interior-Point Algorithm for Convex Qudratic Programming, *European Journal of Operational Research*, Vol. 215, pp. 25-38, 2011.

[25] Luca Zaccarian and Andrew R. Teel, *Modern Anti-windup Synthesis: Control augmentation for actuator saturation*, Princeton University Press, Princeton, 2011.

[26] Y. Zhang, Solving Large-Scale Linear Programs by Interior-Point Methods Under the MATLAB Environment, Technical Report, Department of Mathematics and Statistics, University of Maryland, Baltimore County, Marland, 1996.