MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Infeasibility Detection in Alternating Direction Method of Multipliers for Convex Quadratic Programs

Raghunathan, A. U.; Di Cairano, S.

TR2014-109 December 2014

Abstract

We investigate the infeasibility detection in the alternating direction method of multipliers (ADMM) when minimizing a convex quadratic objective subject to linear equalities and simple bounds. The ADMM formulation consists of alternating between an equality constrained quadratic program (QP) and a projection onto the bounds. We show that: (i) the sequence of iterates generated by ADMM diverges, (ii) the divergence is restricted to the component of the multipliers along the range space of the constraints and (iii) the primal iterates converge to a minimizer of the Euclidean distance between the subspace defined by equality constraints and the convex set defined by bounds. In addition, we derive the optimal value for the step size parameter in the ADMM algorithm that maximizes the rate of convergence of the primal iterates and dual iterates along the null space. In fact, the optimal step size parameter for the infeasible instances is identical to that for the feasible instances. The theoretical results allow us to specify a practical termination condition for infeasibility and the performance of such criterion is demonstrated in a model predictive control application.

IEEE Conference on Decision and Control

Infeasibility Detection in Alternating Direction Method of Multipliers for Convex Quadratic Programs

Arvind U. Raghunathan[1] and Stefano Di Cairano[1]

***Abstract*— We investigate the infeasibility detection in the alternating direction method of multipliers (ADMM) when minimizing a convex quadratic objective subject to linear equalities and simple bounds. The ADMM formulation consists of alternating between an equality constrained quadratic program (QP) and a projection onto the bounds. We show that: (i) the sequence of iterates generated by ADMM diverges, (ii) the divergence is restricted to the component of the multipliers along the range space of the constraints and (iii) the primal iterates converge to a minimizer of the Euclidean distance between the subspace defined by equality constraints and the convex set defined by bounds. In addition, we derive the optimal value for the step size parameter in the ADMM algorithm that maximizes the rate of convergence of the primal iterates and dual iterates along the null space. In fact, the optimal step size parameter for the infeasible instances is identical to that for the feasible instances. The theoretical results allow us to specify a practical termination condition for infeasibility and the performance of such criterion is demonstrated in a model predictive control application.**

I. INTRODUCTION

In this paper, we consider the solution of QPs of the form:

$$\begin{aligned} \min_{\boldsymbol{y}} \ & \frac{1}{2}\boldsymbol{y}^T\boldsymbol{Q}\boldsymbol{y} + \boldsymbol{q}^T\boldsymbol{y} \\ \text{s.t. } & \boldsymbol{A}\boldsymbol{y} = \boldsymbol{b} \\ & \boldsymbol{y} \in \boldsymbol{\mathcal{Y}} \end{aligned} \tag{1}$$

where $\boldsymbol{y}, \boldsymbol{q} \in \mathbb{R}^n$, $\boldsymbol{Q} \succeq 0$ is a symmetric, positive semidefinite matrix, $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ with $m < n$, $\boldsymbol{b} \in \mathbb{R}^m$ and $\boldsymbol{\mathcal{Y}} = [\boldsymbol{y}^{\min}, \boldsymbol{y}^{\max}]$ where $-\infty \le \boldsymbol{y}_i^{\min} < \boldsymbol{y}_i^{\max} \le \infty$ for $i = 1,\ldots,n$. The assumption on $\boldsymbol{\mathcal{Y}}$ is imposed for computational reasons although the results developed in this paper apply to general convex sets as well.

ADMM algorithms were first proposed by Gabay and Mercier [1]. For a recent survey article refer to Boyd et al [2]. A number of recent papers that have studied convergence of ADMM include [3]–[8]. All of the above cited papers consider the convergence of the algorithm under the assumption that the problem is feasible. Ecsktein and Bertsekas [9] show that for infeasible convex problems at least one of the primal and dual sequences generated by ADMM diverges. Aside from [9], there has been little attention to the behavior of the ADMM algorithm when the problem is infeasible.

In this paper we investigate the behavior of ADMM algorithm on infeasible instances of convex QPs with equalities and bounds. The ADMM formulation is based on our previous work [8] where we split the QP (1) into two blocks, an equality constrained QP and a projection onto the set defined by $\boldsymbol{\mathcal{Y}}$. In [8] we established that our ADMM algorithm converges 2-step Q-linearly to a solution when the QP is feasible and derived the optimal step size parameter. In this work, we consider the detection of infeasibility for the QPs in (1) when using the ADMM formulation of [8]. In particular, we show that sequence of iterates generated by the algorithm does not have a limit point. However, the divergence of the iterates is restricted to the component of the multipliers in the range space of the equality constraints. The iterates for the equality and projection problem converge to minimizers of the Euclidean distance between the subspace defined by equality constraints $\boldsymbol{A}\boldsymbol{y} = \boldsymbol{b}$ and the convex set $\boldsymbol{\mathcal{Y}}$. The choice of optimal step size parameter is identical to that for the feasible case derived in [8]. This shows that the proposed ADMM algorithm converges at an optimal rate for both feasible and infeasible QPs. Furthermore, the analysis allows to define a criterion for identifying infeasible QPs and terminating before the maximum number of iterations is reached.

The rest of the paper is organized as follows. Section II provides some background on the linear spaces, projection operator and the notion of infeasibility of the QP in (1). The ADMM algorithm is described in Section III. We provide a characterization of the limiting behavior of the ADMM algorithm on infeasible QPs in Section IV. Convergence of the algorithm to the limiting sequence and derivation of the optimal step size parameter is shown in Section V. Section VI presents termination conditions and numerical results on infeasibility detection in QPs arising from Model Predictive Control. Conclusions and future work are discussed in Section VII.

Notation: We denote by $\mathbb{R}, \mathbb{R}_+$ the set of reals and set of non-negative reals, respectively. All vectors are assumed to be column vectors. For a vector $x \in \mathbb{R}^n$, x^T denotes its transpose and for two vectors x, y, $(x, y) = [x^T \ y^T]^T$. For a matrix $A \in \mathbb{R}^{n\times n}$, $\rho(A)$ denotes the spectral radius, $\lambda_i(A)$ denotes the eigenvalues and $\lambda_{\min}(A), \lambda_{\max}(A)$ denote the minimum and maximum eigenvalues. For a symmetric matrix A, $A \succ 0$ ($A \succeq 0$) denotes positive (semi)definiteness. We denote by $\boldsymbol{I}_n \in \mathbb{R}^{n\times n}$ the identity matrix. The notation $\lambda \perp x \in \boldsymbol{\mathcal{Y}}$ denotes the inequality $\lambda^T(x' - x) \ge 0,\ \forall x' \in \boldsymbol{\mathcal{Y}}$, which is also called the *variational inequality*. We use $\|\cdot\|$ to denote the 2-norm for vectors and matrices. A sequence $\{x^k\} \subset \mathbb{R}^n$ converging to x^* is said to converge at a Q-linear rate if $\|x^{k+1} - x^*\| \le r\|x^k - x^*\|$ where $0 < r < 1$. We denote by $\{x^k\} \to \bar{x}$ the convergence of the sequence to $\bar{x}$.

[1] A.U. Raghunathan and S. Di Cairano are with Mitsubishi Electric Research Laboratories, Cambridge, MA 02139. raghunathan,dicairano@merl.com

II. Background

We make the following assumptions on the QP in (1).

Assumption 1: The set $\mathcal{Y}$ is non-empty, $\mathcal{Y} \neq \emptyset$.

Assumption 2: The matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ has full row rank equal to m.

Assumption 3: The hessian is positive definite on the null space of the equality constraints, i.e., $\boldsymbol{Z}^T \boldsymbol{Q} \boldsymbol{Z} \succ 0$ where $\boldsymbol{Z} \in \mathbb{R}^{n \times (n-m)}$ is a basis for the null space of $\boldsymbol{A}$.

A. Range and Null Spaces

We will denote by $\boldsymbol{R} \in \mathbb{R}^{n \times m}$ an orthonormal basis for the range space of $\boldsymbol{A}^T$ and by $\boldsymbol{Z} \in \mathbb{R}^{n \times (n-m)}$ an orthonormal basis for the null space of $\boldsymbol{A}$. Then,

$$\boldsymbol{R}^T \boldsymbol{R} = \boldsymbol{I}_m, \;\; \boldsymbol{Z}^T \boldsymbol{Z} \;=\; \boldsymbol{I}_{n-m}, \tag{2a}$$

$$\boldsymbol{R}^T \boldsymbol{Z} \;=\; 0 \tag{2b}$$

$$\boldsymbol{R}\boldsymbol{R}^T + \boldsymbol{Z}\boldsymbol{Z}^T \;=\; \boldsymbol{I}_n. \tag{2c}$$

where the first follows from the the orthonormality of basis matrices, the second from orthogonality of the bases, and the final one follows from the columns of $\boldsymbol{R}, \boldsymbol{Z}$ spanning $\mathbb{R}^n$.

B. Projection on to a Convex Set

Given a convex set $\mathcal{Y} \subseteq \mathbb{R}^n$ we will denote by $\mathbb{P}_{\mathcal{Y}} : \mathbb{R}^n \to \mathcal{Y}$ the projection operator which is defined as the solution of the following strictly convex program,

$$\mathbb{P}_{\mathcal{Y}}(\boldsymbol{y}) := \arg\min_{\boldsymbol{w} \in \mathcal{Y}} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{w}\|^2 \tag{3}$$

The optimality conditions of (3) can be simplified as,

$$\left.\begin{array}{r} \mathbb{P}_{\mathcal{Y}}(\boldsymbol{y}) - \boldsymbol{y} - \boldsymbol{\lambda} = 0 \\ \boldsymbol{\lambda} \perp \mathbb{P}_{\mathcal{Y}}(\boldsymbol{y}) \in \mathcal{Y} \end{array}\right\} \implies (\mathbb{P}_{\mathcal{Y}}(\boldsymbol{y}) - \boldsymbol{y}) \perp \mathbb{P}_{\mathcal{Y}}(\boldsymbol{y}) \in \mathcal{Y}. \tag{4}$$

C. Infeasible QP

Under Assumptions 1 and 2 QP (1) is infeasible iff

$$\{\boldsymbol{y} | \boldsymbol{A}\boldsymbol{y} = \boldsymbol{b}\} \cap \mathcal{Y} = \emptyset. \tag{5}$$

Further, there exist $\boldsymbol{y}^\circ$ such that $\boldsymbol{A}\boldsymbol{y}^\circ = \boldsymbol{b}$, $\boldsymbol{w}^\circ \in \mathcal{Y}$,

$$\begin{aligned} (\boldsymbol{y}^\circ, \boldsymbol{w}^\circ) = \arg\min_{\boldsymbol{y},\boldsymbol{w}} \; & \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{w}\|^2 \\ \text{s.t. } & \boldsymbol{A}\boldsymbol{y} = \boldsymbol{b}, \; \boldsymbol{w} \in \mathcal{Y} \end{aligned} \tag{6}$$

where $\boldsymbol{y}^\circ \neq \boldsymbol{w}^\circ$. It is easily seen from the optimality conditions of (6) that

$$\begin{aligned} & \boldsymbol{w}^\circ - \boldsymbol{y}^\circ \in \text{range}(\boldsymbol{A}^T), \; \boldsymbol{\lambda}^\circ = \boldsymbol{w}^\circ - \boldsymbol{y}^\circ, \\ & \boldsymbol{\lambda}^\circ \perp \boldsymbol{w}^\circ \in \mathcal{Y} \end{aligned} \tag{7}$$

where λ° is a multiplier for the $\boldsymbol{w} \in \mathcal{Y}$. It is easy to show that $\boldsymbol{A}\boldsymbol{y} = \alpha \boldsymbol{b} + (1-\alpha)\boldsymbol{A}\boldsymbol{w}^\circ$, for any $0 < \alpha < 1$ is a hyperplane separating the linear subspace defined by equalities and the set $\mathcal{Y}$. Further, from the optimality conditions it is clear that $(\boldsymbol{y}^\circ + \boldsymbol{Z}\boldsymbol{y}_{\boldsymbol{Z}}, \boldsymbol{w}^\circ + \boldsymbol{Z}\boldsymbol{y}_{\boldsymbol{Z}}, \boldsymbol{\lambda}^\circ)$ is a KKT point of (6) provided $\boldsymbol{y}_{\boldsymbol{Z}} \in \mathbb{R}^{n-m} \neq 0$ is such that $\boldsymbol{w}^\circ + \boldsymbol{Z}\boldsymbol{y}_{\boldsymbol{Z}} \in \mathcal{Y}$. In other words, the point $(\boldsymbol{y}^\circ, \boldsymbol{w}^\circ)$ is only unique along the range space of $\boldsymbol{A}^T$. Observe that $(\boldsymbol{y}^\circ + \boldsymbol{Z}\boldsymbol{y}_{\boldsymbol{Z}}, \boldsymbol{w}^\circ + \boldsymbol{Z}\boldsymbol{y}_{\boldsymbol{Z}})$ are minimizers of Euclidean distance between the hyperplane, $\boldsymbol{A}\boldsymbol{y} = \boldsymbol{b}$ and the convex set, $\mathcal{Y}$.

III. ADMM Algorithm

Consider the following reformulation of the QP in (1),

$$\begin{aligned} \min_{\boldsymbol{y},\boldsymbol{w}} \; & \frac{1}{2}\boldsymbol{y}^T \boldsymbol{Q} \boldsymbol{y} + \boldsymbol{q}^T \boldsymbol{y} \\ \text{s.t. } & \boldsymbol{A}\boldsymbol{y} = \boldsymbol{b}, \boldsymbol{w} \in \mathcal{Y} \\ & \boldsymbol{y} = \boldsymbol{w}. \end{aligned} \tag{8}$$

Proceeding as in [8], define the augmented lagrangian as,

$$L_\beta(\boldsymbol{y}, \boldsymbol{w}, \boldsymbol{\lambda}) := \frac{1}{2}\boldsymbol{y}^T \boldsymbol{Q} \boldsymbol{y} + \boldsymbol{q}^T \boldsymbol{y} + \frac{\beta}{2}\|\boldsymbol{y} - \boldsymbol{w} - \boldsymbol{\lambda}\|^2 \tag{9}$$

where $\beta > 0$ is the penalty parameter and $\beta\boldsymbol{\lambda}$ is the multiplier for $\boldsymbol{y} = \boldsymbol{w}$. We propose to solve (8) by an ADMM algorithm with steps:

$$\begin{aligned} \boldsymbol{y}^{k+1} &= \arg\min_{\boldsymbol{y}} \; L_\beta(\boldsymbol{y}, \boldsymbol{w}^k, \boldsymbol{\lambda}^k) \text{ s.t. } \boldsymbol{A}\boldsymbol{y} = \boldsymbol{b} \\ &= \boldsymbol{M}(\boldsymbol{w}^k + \boldsymbol{\lambda}^k - \tilde{q}) + \boldsymbol{N}\boldsymbol{b} \qquad (10a) \\ \boldsymbol{w}^{k+1} &= \arg\min_{\boldsymbol{w}} \; L_\beta(\boldsymbol{y}^{k+1}, \boldsymbol{w}, \boldsymbol{\lambda}^k) \text{ s.t. } \boldsymbol{w} \in \mathcal{Y} \\ &= \mathbb{P}_{\mathcal{Y}}(\boldsymbol{y}^{k+1} - \boldsymbol{\lambda}^k) \qquad (10b) \\ \boldsymbol{\lambda}^{k+1} &= \boldsymbol{\lambda}^k + \boldsymbol{w}^{k+1} - \boldsymbol{y}^{k+1} \qquad (10c) \end{aligned}$$

where $\boldsymbol{M} := \boldsymbol{Z}\left(\boldsymbol{Z}^T(\boldsymbol{Q}/\beta + \boldsymbol{I}_n)\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}^T$, $\boldsymbol{N} := (\boldsymbol{I}_n - \boldsymbol{M}\boldsymbol{Q}/\beta)\boldsymbol{R}(\boldsymbol{A}\boldsymbol{R})^{-1}$, and $\tilde{\boldsymbol{q}} = \boldsymbol{q}/\beta$. Observe that if the step-size parameter β is fixed, $\boldsymbol{M}, \boldsymbol{N}\boldsymbol{b}$ for the iterations in (10) can be computed just once, possibly (long) before the iterations in (10) being executed. Substituting (10a) in (10b), (10c) and simplifying,

$$\begin{aligned} \boldsymbol{w}^{k+1} &= \mathbb{P}_{\mathcal{Y}}(\boldsymbol{v}^k) \\ \boldsymbol{\lambda}^{k+1} &= (\mathbb{P}_{\mathcal{Y}} - \boldsymbol{I}_n)(\boldsymbol{v}^k) \end{aligned} \tag{11}$$

where $(\mathbb{P}_{\mathcal{Y}} - \boldsymbol{I}_n)(\boldsymbol{v}^k)$ is shorthand for $(\mathbb{P}_{\mathcal{Y}}(\boldsymbol{v}^k) - \boldsymbol{v}^k)$ and,

$$\boldsymbol{v}^k = \boldsymbol{M}\boldsymbol{w}^k + (\boldsymbol{M} - \boldsymbol{I}_n)\boldsymbol{\lambda}^k - \boldsymbol{M}\tilde{\boldsymbol{q}} + \boldsymbol{N}\boldsymbol{b}. \tag{12}$$

The update step (11) is of the form in (4) with $\boldsymbol{y}$ replaced by $\boldsymbol{v}^k$ and hence,

$$\boldsymbol{\lambda}^{k+1} \perp \boldsymbol{w}^{k+1} \in \mathcal{Y}. \tag{13}$$

IV. Limit Sequence of ADMM

Suppose QP (1) is infeasible, then there does not exist any fix points for (10). If that were not the case, then from (10c) there should exist $\boldsymbol{y}, \boldsymbol{w}$ such that $\boldsymbol{y} = \boldsymbol{w}$, $\boldsymbol{A}\boldsymbol{y} = \boldsymbol{b}$ and $\boldsymbol{w} \in \mathcal{Y}$ which violates the assumption of QP (1) being infeasible. We will show that (10) generates a sequence in which only the sequence of multipliers $\{\boldsymbol{\lambda}^k\}$ diverge. In particular, $\{\boldsymbol{\lambda}\}^k$ diverges along a direction in range($\boldsymbol{R}$). The iterates $\{(\boldsymbol{y}^k, \boldsymbol{w}^k)\}$ converge to a point $(\boldsymbol{y}^\circ + \boldsymbol{y}^{\boldsymbol{Q}}, \boldsymbol{w}^\circ + \boldsymbol{y}^{\boldsymbol{Q}})$ minimizing the Euclidean distance between the hyperplane and convex set, where $\boldsymbol{y}^{\boldsymbol{Q}} \in \text{range}(\boldsymbol{Z})$, possibly equal to 0, is such that $\boldsymbol{w}^\circ + \boldsymbol{y}^{\boldsymbol{Q}} \in \mathcal{Y}$. Notably, the limit point for $\{(\boldsymbol{y}^k, \boldsymbol{w}^k)\}$ is independent of β. First we characterize $\boldsymbol{y}^{\boldsymbol{Q}}$.

Lemma 1: Suppose Assumptions 1-3 hold and the QP (1) is infeasible. Then, there exists $\boldsymbol{y}^{\boldsymbol{Q}} \in \text{range}(\boldsymbol{Z})$, $\boldsymbol{\lambda}^{\boldsymbol{Q}} \in \mathbb{R}^n$, with $\boldsymbol{y}^{\boldsymbol{Q}}$, $\boldsymbol{Z}^T\boldsymbol{\lambda}^{\boldsymbol{Q}}$ unique, such that

$$\begin{aligned} \boldsymbol{Z}^T\boldsymbol{Q}(\boldsymbol{y}^\circ + \boldsymbol{y}^{\boldsymbol{Q}}) + \boldsymbol{Z}^T\boldsymbol{q} - \boldsymbol{Z}^T\boldsymbol{\lambda}^{\boldsymbol{Q}} = 0 \\ \boldsymbol{\lambda}^{\boldsymbol{Q}} \perp (\boldsymbol{w}^\circ + \boldsymbol{y}^{\boldsymbol{Q}}) \in \mathcal{Y}. \end{aligned} \tag{14}$$

Furthermore, $(\boldsymbol{\lambda^Q}+\gamma\boldsymbol{\lambda}^\circ)\ \forall\ \gamma \geq 0$ is also a solution to (14).

Proof: Since $\boldsymbol{y}^Q \in \text{range}(\boldsymbol{Z})$, let $\boldsymbol{y^Q}=\boldsymbol{Z}\boldsymbol{y^Q_Z}$ for some $\boldsymbol{y^Q_Z} \in \mathbb{R}^{n-m}$. Substituting this in (14) and simplifying, it is easy to show that (14) are the optimality conditions for,

$$\begin{aligned}\min_{\boldsymbol{y^Q_Z}}\ & \frac{1}{2}(\boldsymbol{y^Q_Z})^T(\boldsymbol{Z}^T\boldsymbol{Q}\boldsymbol{Z})\boldsymbol{y^Q_Z}+(\boldsymbol{Z}^T\boldsymbol{q}+\boldsymbol{Z}^T\boldsymbol{Q}\boldsymbol{y}^\circ)^T\boldsymbol{y^Q_Z} \\ \text{s.t. } & \boldsymbol{w}^\circ+\boldsymbol{Z}\boldsymbol{y^Q_Z}\in\boldsymbol{\mathcal{Y}}.\end{aligned} \tag{15}$$

The strict convexity of the QP (15) follows from Assumption 3 and this guarantees uniqueness of $\boldsymbol{y^Q_Z}$, if one exists. Weak Slater's condition [10] holds for the QP (15) since the constraints in $\boldsymbol{\mathcal{Y}}$ are affine and $\boldsymbol{y^Q_Z}=0$ is a feasible point. The satisfaction of convexity and weak Slater's condition by QP (15) implies that strong duality holds for (15) and the claim on existence of $\boldsymbol{y^Q_Z},\boldsymbol{\lambda^Q}$ holds. The uniqueness of $\boldsymbol{y^Q}$ follows from uniqueness of $\boldsymbol{y^Q_Z}$ and full column rank of $\boldsymbol{Z}$. The uniqueness of $\boldsymbol{Z}^T\boldsymbol{\lambda^Q}$ follows from the first equation of (14) and uniqueness of $\boldsymbol{y^Q}$.

To prove the remaining claim, consider the choice of $(\boldsymbol{\lambda^Q}+\gamma\boldsymbol{\lambda}^\circ)$ as a solution to (14). Satisfaction of the first equation in (14) follows from $\boldsymbol{\lambda}^\circ \in \text{range}(\boldsymbol{R})$ by (7) and $\boldsymbol{Z}^T\boldsymbol{R}=0$ by (2b). As for the variational inequality in (14),

$$\begin{aligned}&(\boldsymbol{\lambda^Q}+\gamma\boldsymbol{\lambda}^\circ)^T(\boldsymbol{w}'-(\boldsymbol{w}^\circ+\boldsymbol{y^Q}))\\ =\ & \underbrace{(\boldsymbol{\lambda^Q})^T(\boldsymbol{w}'-(\boldsymbol{w}^\circ+\boldsymbol{y^Q}))}_{\geq 0}+\underbrace{\gamma(\boldsymbol{\lambda}^\circ)^T(\boldsymbol{w}'-\boldsymbol{w}^\circ)}_{\geq 0}\\ & -\underbrace{\gamma(\boldsymbol{\lambda}^\circ)^T\boldsymbol{y^Q}}_{=0}\geq 0\ \forall\ \boldsymbol{w}'\in\boldsymbol{\mathcal{Y}}\end{aligned}$$

where the first term is non-negative by variational inequality in (14), the second term is non-negative by variational inequality in (7) and the last term vanishes since $\boldsymbol{\lambda}^\circ \in \text{range}(\boldsymbol{R})$ and $\boldsymbol{y^Q}\in\text{range}(\boldsymbol{Z})$, proving the claim. ■

The next lemma establishes some properties of the ADMM iterate sequence.

Lemma 2: Let $\{(\boldsymbol{y}^k,\boldsymbol{w}^k,\boldsymbol{\lambda}^k)\}$ be generated by the ADMM algorithm (10). Then the following statements are equivalent.

$$\{\boldsymbol{w}^k\}\to\bar{\boldsymbol{w}},\{\boldsymbol{Z}^T\boldsymbol{\lambda}^k\}\to\bar{\boldsymbol{\lambda}}_{\boldsymbol{Z}},\{\boldsymbol{R}^T(\boldsymbol{\lambda}^{k+1}-\boldsymbol{\lambda}^k)\}\to\bar{\boldsymbol{\lambda}}_{\boldsymbol{R}} \tag{16}$$

$$\{\boldsymbol{w}^k\}\to\bar{\boldsymbol{w}},\{\boldsymbol{y}^k\}\to\bar{\boldsymbol{y}},\bar{\boldsymbol{y}}\neq\bar{\boldsymbol{w}} \tag{17}$$

for some $\bar{\boldsymbol{\lambda}}_{\boldsymbol{R}}\neq 0$. Further, if (16) (or (17)) hold then,

$$\left\{\frac{(\boldsymbol{\lambda}^k)^T(\boldsymbol{w}^k-\boldsymbol{y}^k)}{\|\boldsymbol{\lambda}^k\|\|\boldsymbol{w}^k-\boldsymbol{y}^k\|}\right\}\to 1. \tag{18}$$

Proof: Suppose (16) holds. From (10a)

$$\{\boldsymbol{y}^k\}\to\bar{\boldsymbol{y}}:=\boldsymbol{M}(\bar{\boldsymbol{w}}+\boldsymbol{Z}\bar{\boldsymbol{\lambda}}_{\boldsymbol{Z}}-\tilde{\boldsymbol{q}})+\boldsymbol{N}\boldsymbol{b}.$$

from (2c) and $\boldsymbol{M}\boldsymbol{R}=0$. From (16), it must hold by (10c) that $\{\boldsymbol{Z}^T(\boldsymbol{w}^k-\boldsymbol{y}^k)\}\to 0$, $\{\boldsymbol{R}^T(\boldsymbol{w}^k-\boldsymbol{y}^k)\}\not\to 0$. Thus, $\bar{\boldsymbol{y}}\neq\bar{\boldsymbol{w}}$.

Suppose (17) holds. From (10a) and using (17) we have that $\{\boldsymbol{Z}^T\boldsymbol{\lambda}^k\}$ must converge. Since, $\bar{\boldsymbol{y}}\neq\bar{\boldsymbol{w}}$ and $\{\boldsymbol{Z}^T\boldsymbol{\lambda}^k\}\to\bar{\boldsymbol{\lambda}}_{\boldsymbol{Z}}$, then by (10c) this implies that $\{\boldsymbol{R}^T(\boldsymbol{\lambda}^{k+1}-\boldsymbol{\lambda}^k)\}\to\bar{\boldsymbol{\lambda}}_{\boldsymbol{R}}$ for some $\bar{\boldsymbol{\lambda}}_{\boldsymbol{R}}\in\mathbb{R}^m$.

To show (18), consider the following decomposition,

$$\frac{\boldsymbol{w}^k-\boldsymbol{y}^k}{\|\boldsymbol{w}^k-\boldsymbol{y}^k\|}=\alpha_1^k\boldsymbol{\zeta}+\boldsymbol{\nu}_1^k,\ \frac{\boldsymbol{\lambda}^k}{\|\boldsymbol{\lambda}^k\|}=\alpha_2^k\boldsymbol{\zeta}+\boldsymbol{\nu}_2^k \tag{19}$$

where $\boldsymbol{\zeta}=(\bar{\boldsymbol{w}}-\bar{\boldsymbol{y}})/\|\bar{\boldsymbol{w}}-\bar{\boldsymbol{y}}\|$, and $\boldsymbol{\zeta}^T\boldsymbol{\nu}_i^k=0$, for $i=1,2$. Further, from (16),(17) we have that $\{\alpha_i^k\}\to 1$, $\{\boldsymbol{\nu}_i^k\}\to 0$ for $i=1,2$. Substituting (19) in (18),

$$\begin{aligned}&\frac{(\boldsymbol{\lambda}^k)^T(\boldsymbol{w}^k-\boldsymbol{y}^k)}{\|\boldsymbol{\lambda}^k\|\|\boldsymbol{w}^k-\boldsymbol{y}^k\|}=(\alpha_1^k\boldsymbol{\zeta}+\boldsymbol{\nu}_1^k)^T(\alpha_2^k\boldsymbol{\zeta}+\boldsymbol{\nu}_2^k)\\ &=\alpha_1^k\alpha_2^k+\boldsymbol{\zeta}^T(\alpha_1^k\boldsymbol{\nu}_1^k+\alpha_2^k\boldsymbol{\nu}_2^k)+(\boldsymbol{\nu}_1^k)^T\boldsymbol{\nu}_2^k\\ &\geq\alpha_1^k\alpha_2^k-\alpha_1^k\|\boldsymbol{\nu}_1^k\|-\alpha_2^k\|\boldsymbol{\nu}_2^k\|-\|\boldsymbol{\nu}_1^k\|\|\boldsymbol{\nu}_2^k\|\end{aligned}$$

where the second equality follows from expanding terms and using $\|\zeta\|=1$, while the last inequality is obtained from the Cauchy-Schwarz inequality. The result in (18) follows from the limit of the sequence of $\alpha_i^k,\boldsymbol{\nu}_i^k$. ■

Using Lemmas 1 and 2 we can state the limiting behavior of the ADMM iterations (10) when the QP (1) is infeasible.

Theorem 1: Suppose Assumptions 1-3 hold. Then, the following are true.

(i) If QP (1) is infeasible then, $\{(\boldsymbol{y}^\circ+\boldsymbol{y^Q},\boldsymbol{w}^\circ+\boldsymbol{y^Q},\hat{\boldsymbol{\lambda}}^k)\}$ is a sequence satisfying (10) for $k\geq k'$ sufficiently large with, $\boldsymbol{y^Q},\boldsymbol{\lambda^Q}$ as defined in (14) and,

$$\hat{\boldsymbol{\lambda}}^k=\frac{1}{\beta}(\boldsymbol{\lambda^Q}+(k-\gamma_1)\boldsymbol{\lambda}^\circ),\gamma_1\leq k'. \tag{20}$$

(ii) If the ADMM algorithm (10) generates $\{(\boldsymbol{y}^k,\boldsymbol{w}^k,\boldsymbol{\lambda}^k)\}$ satisfying (17) then, the QP (1) is infeasible. Further, $\bar{\boldsymbol{y}}=\boldsymbol{y}^\circ+\boldsymbol{y^Q}$, $\bar{\boldsymbol{w}}=\boldsymbol{w}^\circ+\boldsymbol{w^Q}$ and $\boldsymbol{\lambda}^k$ satisifes (20).

Proof: Consider the claim in (i). For proving that (10a) holds, we need to show that,

$$\boldsymbol{y}^\circ+\boldsymbol{y^Q}-\boldsymbol{M}(\boldsymbol{w}^\circ+\boldsymbol{y^Q}+\hat{\boldsymbol{\lambda}}^k-\tilde{\boldsymbol{q}})-\boldsymbol{N}\boldsymbol{b}=0. \tag{21}$$

Multiplying the left hand side of (21) by $\boldsymbol{R}^T$, using $\boldsymbol{R}^T\boldsymbol{M}=0$, $\boldsymbol{R}^T\boldsymbol{y^Q}=0$ and simplifying,

$$\boldsymbol{R}^T\boldsymbol{y}^\circ-(\boldsymbol{A}\boldsymbol{R})^{-1}\boldsymbol{b}=(\boldsymbol{A}\boldsymbol{R})^{-1}(\boldsymbol{A}\boldsymbol{R}\boldsymbol{R}^T\boldsymbol{y}^\circ-\boldsymbol{b})=0 \tag{22}$$

where the last equality follows from (7). Multiplying the left hand side of (21) by $\boldsymbol{Z}^T$, from $\boldsymbol{Z}^T\boldsymbol{M}=\hat{\boldsymbol{M}}\boldsymbol{Z}^T$ where $\hat{\boldsymbol{M}}=(\boldsymbol{Z}^T\boldsymbol{Q}\boldsymbol{Z}/\beta+\boldsymbol{I}_{n-m})^{-1}$, $\boldsymbol{Z}^T\boldsymbol{N}\boldsymbol{b}=-(\hat{\boldsymbol{M}}\boldsymbol{Z}^T\boldsymbol{Q}/\beta)\boldsymbol{R}\boldsymbol{R}^T(\boldsymbol{y}^\circ+\boldsymbol{y^Q})$ we obtain,

$$\begin{aligned}&\boldsymbol{Z}^T(\boldsymbol{y}^\circ+\boldsymbol{y^Q})-\hat{\boldsymbol{M}}\boldsymbol{Z}^T(\boldsymbol{w}^\circ+\boldsymbol{y^Q}+\hat{\boldsymbol{\lambda}}^k-\tilde{\boldsymbol{q}})\\ &\quad+\hat{\boldsymbol{M}}\boldsymbol{Z}^T(\boldsymbol{Q}/\beta)\boldsymbol{R}\boldsymbol{R}^T(\boldsymbol{y}^\circ+\boldsymbol{y^Q})\\ &=\hat{\boldsymbol{M}}\Big((\boldsymbol{Z}^T\boldsymbol{Q}\boldsymbol{Z}/\beta+\boldsymbol{I}_{n-m})\boldsymbol{Z}^T(\boldsymbol{y}^\circ+\boldsymbol{y^Q})\\ &\quad-\boldsymbol{Z}^T\Big((\boldsymbol{w}^\circ+\boldsymbol{y^Q}+\hat{\boldsymbol{\lambda}}^k-\tilde{\boldsymbol{q}})+(\boldsymbol{Q}/\beta)\boldsymbol{R}\boldsymbol{R}^T(\boldsymbol{y}^\circ+\boldsymbol{y^Q})\Big)\Big)\\ &=\hat{\boldsymbol{M}}\Big(\boldsymbol{Z}^T(\boldsymbol{Q}/\beta)(\boldsymbol{y}^\circ+\boldsymbol{y^Q})+\boldsymbol{Z}^T(\boldsymbol{y}^\circ+\boldsymbol{y^Q})\\ &\quad-\boldsymbol{Z}^T(\boldsymbol{w}^\circ+\boldsymbol{y^Q}+\boldsymbol{\lambda^Q}-\tilde{\boldsymbol{q}})\Big)\\ &=(\hat{\boldsymbol{M}}/\beta)\Big(\boldsymbol{Z}^T\boldsymbol{Q}(\boldsymbol{y}^\circ+\boldsymbol{y^Q})+\boldsymbol{Z}^T\boldsymbol{q}-\boldsymbol{Z}^T\boldsymbol{\lambda^Q}\Big)=0\end{aligned} \tag{23}$$

where the first equality follows simply by removing $\hat{\boldsymbol{M}}$ as the common multiplicative factor, the second equality follows from (2c), the third equality from (7), (20) and the final equality from (14). Combining (22) and (23) shows that the said sequence satisfies (21). To prove that (10b) holds consider for any $\boldsymbol{w}' \in \boldsymbol{\mathcal{Y}}$,

$$\begin{aligned}
&\left(\boldsymbol{w}^\circ + \boldsymbol{y}^{\boldsymbol{Q}} - \boldsymbol{y}^\circ - \boldsymbol{y}^{\boldsymbol{Q}} + \hat{\boldsymbol{\lambda}}^k\right)^T \left(\boldsymbol{w}' - \boldsymbol{w}^\circ - \boldsymbol{y}^{\boldsymbol{Q}}\right) \\
&= \left(\boldsymbol{w}^\circ - \boldsymbol{y}^\circ + \hat{\boldsymbol{\lambda}}^k\right)^T \left(\boldsymbol{w}' - \boldsymbol{w}^\circ - \boldsymbol{y}^{\boldsymbol{Q}}\right) \\
&= -\frac{1}{\beta}\left(\boldsymbol{\lambda}^{\boldsymbol{Q}} + (k+1)\boldsymbol{\lambda}^\circ\right)^T \left(\boldsymbol{w}' - \boldsymbol{w}^\circ - \boldsymbol{y}^{\boldsymbol{Q}}\right) \\
&= \frac{1}{\beta}\left(\boldsymbol{\lambda}^{\boldsymbol{Q}} + (k-\gamma_1+1)\boldsymbol{\lambda}^\circ\right)^T \left(\boldsymbol{w}' - \boldsymbol{w}^\circ - \boldsymbol{y}^{\boldsymbol{Q}}\right) \geq 0
\end{aligned} \tag{24}$$

where the second equality follows from (7) and (20), and the inequality follows from Lemma 1 by noting that $\gamma = (k - \gamma_1 + 1) \geq 0$. Thus, $\boldsymbol{w}^\circ + \boldsymbol{w}^{\boldsymbol{Q}} = \mathbb{P}_{\boldsymbol{\mathcal{Y}}}\left(\boldsymbol{y}^\circ + \boldsymbol{y}^{\boldsymbol{Q}} - \boldsymbol{\lambda}^k\right)$ holds and the sequence in the claim satisfies (10b). Finally, the definition of $\boldsymbol{\lambda}^k$ in (20) implies that (10c) holds, and thus (i) is proved.

Consider the claim in part (ii). From (18) we have that for any $\epsilon > 0$ there exists k_ϵ such that for all $k \geq k_\epsilon$,

$$\frac{(\boldsymbol{\lambda}^k)^T(\boldsymbol{w}^k - \boldsymbol{y}^k)}{\|\boldsymbol{w}^k - \boldsymbol{y}^k\|^2} \geq (1-\epsilon)\frac{\|\boldsymbol{\lambda}^k\|}{\|\boldsymbol{w}^k - \boldsymbol{y}^k\|}. \tag{25}$$

From which we have that,

$$\boldsymbol{\lambda}^k = \alpha^k(\boldsymbol{w}^k - \boldsymbol{y}^k) + \boldsymbol{\mu}^k, \tag{26a}$$

$$\alpha^k = \frac{(\boldsymbol{\lambda}^k)^T(\boldsymbol{w}^k - \boldsymbol{y}^k)}{\|\boldsymbol{w}^k - \boldsymbol{y}^k\|^2} \geq (1-\epsilon)\frac{\|\boldsymbol{\lambda}^k\|}{\|\boldsymbol{w}^k - \boldsymbol{y}^k\|}, \tag{26b}$$

$$\|\boldsymbol{\mu}^k\| \leq \sqrt{1-(1-\epsilon)^2}\|\boldsymbol{\lambda}^k\|. \tag{26c}$$

Then for all $\boldsymbol{w}' \in \boldsymbol{\mathcal{Y}}$ we have that,

$$\begin{aligned}
&(\boldsymbol{w}^k - \boldsymbol{y}^k)^T(\boldsymbol{w}' - \boldsymbol{w}^k) \\
&= \frac{1}{\alpha^k}\underbrace{(\boldsymbol{\lambda}^k)^T(\boldsymbol{w}' - \boldsymbol{w}^k)}_{\geq 0} - \frac{1}{\alpha^k}(\boldsymbol{\mu}^k)^T(\boldsymbol{w}' - \boldsymbol{w}^k) \\
&\geq -\frac{\sqrt{1-(1-\epsilon)^2}}{1-\epsilon}\|\boldsymbol{w}^k - \boldsymbol{y}^k\|\|\boldsymbol{w}' - \boldsymbol{w}^k\|
\end{aligned} \tag{27}$$

where the inequality follows from (13), the Cauchy-Schwarz inequality and the substitution of (26b) and (26c). Hence,

$$\begin{aligned}
&\lim_{k\to\infty}\frac{(\boldsymbol{w}^k - \boldsymbol{y}^k)^T(\boldsymbol{w}' - \boldsymbol{w}^k)}{\|\boldsymbol{w}^k - \boldsymbol{y}^k\|\|\boldsymbol{w}' - \boldsymbol{w}^k\|} \geq 0 \;\forall\, \boldsymbol{w}' \in \boldsymbol{\mathcal{Y}} \\
&\implies \frac{(\bar{\boldsymbol{w}} - \bar{\boldsymbol{y}})^T(\boldsymbol{w}' - \bar{\boldsymbol{w}})}{\|\bar{\boldsymbol{w}} - \bar{\boldsymbol{y}}\|\|\boldsymbol{w}' - \bar{\boldsymbol{w}}\|} \geq 0 \;\forall\, \boldsymbol{w}' \in \boldsymbol{\mathcal{Y}}.
\end{aligned} \tag{28}$$

and $(\bar{\boldsymbol{w}} - \bar{\boldsymbol{y}}) \perp \bar{\boldsymbol{w}} \in \boldsymbol{\mathcal{Y}}$. Since $\boldsymbol{A}\bar{\boldsymbol{y}} = \boldsymbol{b}$, $\bar{\boldsymbol{w}} \in \boldsymbol{\mathcal{Y}}$ we have that $(\bar{\boldsymbol{y}}, \bar{\boldsymbol{w}})$ satisfies (7) and hence, the QP (1) is infeasible. From uniqueness of the range space component in (6), $\boldsymbol{R}^T\bar{\boldsymbol{y}} = \boldsymbol{R}^T\boldsymbol{y}^\circ$, $\boldsymbol{R}^T\bar{\boldsymbol{w}} = \boldsymbol{R}^T\boldsymbol{w}^\circ$ and also $\boldsymbol{Z}^T\bar{\boldsymbol{w}} = \boldsymbol{Z}^T\bar{\boldsymbol{y}}$. From the update steps in the ADMM (10) we have that,

$$\begin{aligned}
\boldsymbol{Z}^T\left(\boldsymbol{Q}\left(\boldsymbol{y}^\circ + \boldsymbol{Z}\boldsymbol{Z}^T(\bar{\boldsymbol{y}} - \boldsymbol{y}^\circ)\right) + \boldsymbol{q} - \beta\boldsymbol{\lambda}^k\right) = 0, \\
\boldsymbol{\lambda}^k \perp \boldsymbol{w}^\circ + \boldsymbol{Z}\boldsymbol{Z}^T(\bar{\boldsymbol{w}} - \boldsymbol{w}^\circ) \in \boldsymbol{\mathcal{Y}},
\end{aligned} \tag{29}$$

for all k sufficiently large, where first equation follows by replacing $\boldsymbol{y}^{\boldsymbol{Q}}$, $\boldsymbol{\lambda}^{\boldsymbol{Q}}$ by $\boldsymbol{Z}\boldsymbol{Z}^T(\bar{\boldsymbol{y}} - \boldsymbol{y}^\circ)$, $\beta\boldsymbol{\lambda}^k$, respectively, in (23), and the second condition follows from (13). The conditions in (29) are the conditions in (14) and hence, Lemma 1 applies to yield that $\boldsymbol{Z}\boldsymbol{Z}^T(\bar{\boldsymbol{y}} - \boldsymbol{y}^\circ) = \boldsymbol{Z}\boldsymbol{Z}^T(\bar{\boldsymbol{w}} - \boldsymbol{w}^\circ) = \boldsymbol{y}^{\boldsymbol{Q}}$, $\boldsymbol{Z}^T\boldsymbol{\lambda}^k = \boldsymbol{Z}^T\boldsymbol{\lambda}^{\boldsymbol{Q}}$. Thus, $\bar{\boldsymbol{y}} = \boldsymbol{y}^\circ + \boldsymbol{y}^{\boldsymbol{Q}}$, $\bar{\boldsymbol{w}} = \boldsymbol{w}^\circ + \boldsymbol{y}^{\boldsymbol{Q}}$, $\boldsymbol{\lambda}^k$ satisfies (20) and the claim holds. ■

Observe that in (20) the range space component $\boldsymbol{\lambda}^{\boldsymbol{Q}}$ is not unique. The ADMM iterations only specify that $\hat{\boldsymbol{\lambda}}^{k+1} - \hat{\boldsymbol{\lambda}}^k = \boldsymbol{\lambda}^\circ$, and hence that (20) holds for some constant γ_1.

V. Convergence of the Algorithm

First, we recall some results on the projection operator.

Lemma 3 (Lemma 3 [8]): For any $v, v' \in \mathbb{R}^n$:

(i) $(\mathbb{P}_{\boldsymbol{\mathcal{Y}}}(v) - \mathbb{P}_{\boldsymbol{\mathcal{Y}}}(v'))^T((\boldsymbol{I}_n - \mathbb{P}_{\boldsymbol{\mathcal{Y}}})(v) - (\boldsymbol{I}_n - \mathbb{P}_{\boldsymbol{\mathcal{Y}}})(v')) \geq 0$

(ii) $\|(\mathbb{P}_{\boldsymbol{\mathcal{Y}}}(v), (\boldsymbol{I}_n - \mathbb{P}_{\boldsymbol{\mathcal{Y}}})(v)) - (\mathbb{P}_{\boldsymbol{\mathcal{Y}}}(v'), (\boldsymbol{I}_n - \mathbb{P}_{\boldsymbol{\mathcal{Y}}})(v'))\| \leq \|v - v'\|$

(iii) $\|(2\mathbb{P}_{\boldsymbol{\mathcal{Y}}} - \boldsymbol{I}_n)(v) - (2\mathbb{P}_{\boldsymbol{\mathcal{Y}}} - \boldsymbol{I}_n)(v')\| \leq \|v - v'\|$

The following result on spectral radius of $\boldsymbol{M}$ is also useful.

Lemma 4 (Lemma 4 [8]): Suppose Assumptions 2 and 3 hold. Then, $\rho(\boldsymbol{Z}^T\boldsymbol{M}\boldsymbol{Z}) < 1$ and $\rho(\boldsymbol{M}) < 1$.

Next, we introduce some properties satisfied by the iterates (11). The proofs, which are not shown for the sake of brevity, can be obtained using Lemmas 3 and 4.

Lemma 5: Suppose Assumptions 1–3 hold. Then, the sequence $\{(\boldsymbol{w}^k, \boldsymbol{\lambda}^k)\}$ produced by (11) is such that:

(i) $\|\boldsymbol{v}^{k+1} - \boldsymbol{v}^k\| \leq \|(\boldsymbol{w}^{k+1}, \boldsymbol{\lambda}^{k+1}) - (\boldsymbol{w}^k, \boldsymbol{\lambda}^k)\|$

(ii) $\|(\boldsymbol{w}^{k+1}, \boldsymbol{\lambda}^{k+1}) - (\boldsymbol{w}^k, \boldsymbol{\lambda}^k)\| \leq \|\boldsymbol{v}^k - \boldsymbol{v}^{k-1}\|$

(iii) $\|(\boldsymbol{w}^{k+1}, \boldsymbol{\lambda}^{k+1}) - (\boldsymbol{w}^k, \boldsymbol{\lambda}^k)\| \leq \|(\boldsymbol{w}^k, \boldsymbol{\lambda}^k) - (\boldsymbol{w}^{k-1}, \boldsymbol{\lambda}^{k-1})\|$

(iv) $\|\boldsymbol{v}^{k+1} - \boldsymbol{v}^k\| \leq \|\boldsymbol{v}^k - \boldsymbol{v}^{k-1}\|$.

Lemma 6: Suppose Assumptions 1- 3 hold and define,

$$\boldsymbol{u}^k = (2\mathbb{P}_{\boldsymbol{\mathcal{Y}}} - \boldsymbol{I}_n)(\boldsymbol{v}^k) - (2\mathbb{P}_{\boldsymbol{\mathcal{Y}}} - \boldsymbol{I}_n)(\boldsymbol{v}^{k-1}). \tag{30}$$

Then, inequality in Lemma 5(iv) holds strictly if $\boldsymbol{Z}^T\boldsymbol{u}^k \neq 0$.

We omit the proof for brevity and refer the reader to Theorem 3 in [8] where a similar inequality is proved.

The following result characterizes the limit behavior of iterates for infeasible QP (1) in terms of the sequence $\{\boldsymbol{v}^k\}$.

Lemma 7: Suppose Assumptions 1- 3 hold. Then,

$$\boldsymbol{v}^{k+1} - \boldsymbol{v}^k = \boldsymbol{v}^k - \boldsymbol{v}^{k-1} \neq 0 \tag{31}$$

holds iff $\boldsymbol{y}^l = \boldsymbol{y}^\circ + \boldsymbol{y}^{\boldsymbol{Q}}$, $\boldsymbol{w}^l = \boldsymbol{w}^\circ + \boldsymbol{y}^{\boldsymbol{Q}}$ and $\boldsymbol{\lambda}^l$ satisfies (20) $\forall\, l \geq k-1$.

Proof: The *if* part follows trivially for the given choice of $(\boldsymbol{w}^k, \boldsymbol{\lambda}^k)$. Consider the *only if* part. We cannot have $\boldsymbol{Z}^T\boldsymbol{u}^k \neq 0$ since that will violate (31) by Lemma 6. Hence, $\boldsymbol{Z}^T\boldsymbol{u}^k = 0 \implies \boldsymbol{R}\boldsymbol{R}^T\boldsymbol{u}^k = \boldsymbol{u}^k$. Using (12),(30)

$$\begin{aligned}
\boldsymbol{v}^{k+1} - \boldsymbol{v}^k &= \frac{1}{2}\left(-\boldsymbol{u}^k + \boldsymbol{v}^k - \boldsymbol{v}^{k-1}\right) \\
&= (\boldsymbol{I}_n - \mathbb{P}_{\boldsymbol{\mathcal{Y}}})(\boldsymbol{v}^k) - (\boldsymbol{I}_n - \mathbb{P}_{\boldsymbol{\mathcal{Y}}})(\boldsymbol{v}^{k-1}) \\
&= \boldsymbol{v}^k - \boldsymbol{v}^{k-1} \neq 0
\end{aligned} \tag{32}$$

where the last equality from (31). Combining this with Lemma 5(i) and 5(ii), yields that $\mathbb{P}_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{v}^k) = \boldsymbol{w}^{k+1} = \bar{\boldsymbol{w}}$ $\forall\, k$. Furthermore convergence of $\boldsymbol{w}^k$ and $\boldsymbol{Z}^T\boldsymbol{u}^k = 0$ yields that $\boldsymbol{Z}^T\boldsymbol{\lambda}^k = \bar{\boldsymbol{\lambda}}_{\boldsymbol{Z}}$ $\forall\, k$. The update steps for (10a) yield that $\boldsymbol{y}^k = \bar{\boldsymbol{y}}$ $\forall\, k$. Further, from (32) we have that $\boldsymbol{\lambda}^k \neq \boldsymbol{\lambda}^{k-1}$

which implies that $\bar{\boldsymbol{y}} \neq \bar{\boldsymbol{w}}$. Thus, the sequence of iterates satisfy conditions in Theorem 1(ii) and the claim follows. ■

Theorem 2: Suppose Assumptions 1-3 hold, $\beta > 0$ and QP (1) is infeasible. Then, (i) $\{\boldsymbol{v}^k\}$ converges Q-linearly to a sequence satisfying (31) and (ii) $\{(\boldsymbol{w}^k, \boldsymbol{\lambda}^k)\}_{k\geq 2}$ converges 2-step Q-linearly to a sequence defined in Theorem 1(i).

Proof: Infeasibility of QP (1) ensures that $\boldsymbol{y}^\circ$, $\boldsymbol{w}^\circ$, $\boldsymbol{y}^{\boldsymbol{Q}}$, $\boldsymbol{\lambda}^\circ$, $\boldsymbol{\lambda}^{\boldsymbol{Q}}$ are well-defined. From Lemma 6 we have that $\{\|\boldsymbol{v}^k - \boldsymbol{v}^{k-1}\|\}$ decreases monotonically until (31) holds. Hence, the claim on $\{\boldsymbol{v}^k\}$ is proved. The result on $\{(\boldsymbol{w}^k, \boldsymbol{\lambda}^k)\}$ follows from Lemma 3(ii), the result on monotonic decrease of $\{\|\boldsymbol{v}^k - \boldsymbol{v}^{k-1}\|\}$ and Lemma 3(i). ■

From Lemma 6 it is clear that rate of convergence is influenced by the components of $\boldsymbol{u}^k$ along the null space of the constraints. We can affect the contraction resulting from the null space component by choosing β^* to minimize $\frac{\|2\boldsymbol{ZMZ}^T - \boldsymbol{I}_{n-m}\|}{2} + \frac{1}{2}$ where the eigenvalues of $\boldsymbol{ZMZ}^T$ satisfy $\lambda(\boldsymbol{ZMZ}^T) = \lambda((\boldsymbol{Z}^T(\boldsymbol{Q}/\beta + \boldsymbol{I}_n)\boldsymbol{Z})^{-1}) = \beta/(\beta + \lambda(\boldsymbol{Z}^T\boldsymbol{QZ}))$. Thus, the optimal choice for the step size is given by,

$$\beta^* = \arg\min_{\beta>0} \max_i \left(\left| \frac{\beta}{\beta + \lambda_i(\boldsymbol{Z}^T\boldsymbol{QZ})} - \frac{1}{2} \right| + \frac{1}{2} \right).$$

Theorem 3: Suppose Assumptions 1-3 hold. Then, the optimal step-size for the class of convex QPs in (1) to converge to the limiting sequence in Theorem 1(i) is

$$\beta^* = \sqrt{\lambda_{\min}(\boldsymbol{Z}^T\boldsymbol{QZ})\lambda_{\max}(\boldsymbol{Z}^T\boldsymbol{QZ})}. \tag{33}$$

The choice of optimal step size parameter and proof is identical to that for the feasible case derived in [8], so the proof is omitted here. Thus, the choice of the step size for the proposed ADMM algorithm results in optimal convergence rate for both feasible and infeasible QPs.

VI. Numerical Results in MPC Applications

A. Practial Termination Condition

Based on Theorem 1, we propose the satisfaction of the following conditions for detecting infeasibility:

$$\max(\beta\|\boldsymbol{w}^k - \boldsymbol{w}^{k-1}\|, \|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^{k-1}\|) > \epsilon_o \tag{34a}$$

$$\frac{\max(\|\boldsymbol{y}^k - \boldsymbol{y}^{k-1}\|, \beta\|\boldsymbol{w}^k - \boldsymbol{w}^{k-1}\|)}{\max(\beta\|\boldsymbol{w}^k - \boldsymbol{w}^{k-1}\|, \|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^{k-1}\|)} \leq \epsilon_r \tag{34b}$$

$$\frac{(\boldsymbol{\lambda}^k)^T(\boldsymbol{w}^k - \boldsymbol{y}^k)}{\|\boldsymbol{\lambda}^k\|\|\boldsymbol{w}^k - \boldsymbol{y}^k\|} \geq 1 - \epsilon_a \tag{34c}$$

$$\boldsymbol{\lambda}^k \circ (\boldsymbol{w}^k - \boldsymbol{y}^k) \geq 0 \text{ or } \frac{\|\Delta\boldsymbol{v}^k - \Delta\boldsymbol{v}^{k-1}\|}{\|\boldsymbol{v}^k\|} \leq \epsilon_v \tag{34d}$$

where, $0 \leq \epsilon_o, \epsilon_r, \epsilon_a, \epsilon_v \ll 1$, $\circ$ represents the component-wise multiplication operation and $\Delta\boldsymbol{v}^k = \boldsymbol{v}^k - \boldsymbol{v}^{k-1}$. The left hand side (34a) is the error criterion used for termination in feasible QPs [7], [8]. Condition (34a) requires that the optimality conditions are not satisfied to a tolerance of ϵ_o, while (34b) requires that the change in $\boldsymbol{y}, \boldsymbol{w}$ iterates be much smaller than the change in the $\boldsymbol{w}, \boldsymbol{\lambda}$ iterates. In the case of a feasible QP all the iterates converge and nothing specific can be said about this ratio. However, as shown in Theorem 1 the multiplier iterates change by a constant vector in the case of an infeasible QP. Hence, we expect the ratio in (34b) to be small in the infeasibile case while (34a) is large. The condition (34c) checks for the satisfaction of (18) to a tolerance of ϵ_a. The first condition in (34d) checks that each component of $\boldsymbol{\lambda}^k$ and $\boldsymbol{w}^k - \boldsymbol{y}^k$ have the same sign. In a sense, this is a stricter requirement of the angle condition (34c). In our numerical experiments we have observed that the satisfaction of this condition can be quite slow to converge when the iterates are far from a solution. In such instances, we have also observed that, the quantity $\|\boldsymbol{v}^k\|$ has actually diverged to a large value. To remedy this we also monitor the ratio of $\|\Delta\boldsymbol{v}^k - \Delta\boldsymbol{v}^{k-1}\|$ (which converges to 0, refer Lemma 7) to $\|\boldsymbol{v}^k\|$ ($\|\boldsymbol{v}^k\| \to \infty$). This ratio is expected to converge to 0 on infeasible instances. We recommend following parameter setting: $\epsilon_o = 10^{-6}, \epsilon_r = 10^{-3}, \epsilon_a = 10^{-3}, \epsilon_v = 10^{-4}$. While these values have worked well on a large number of problems, these constants might have to be modifed depending on scaling of the problem.

B. Numerical Example

In this section we present some numerical results on the infeasibility detection for the ADMM algorithm (10) applied to the quadratic programs arising in constrained linear model predictive control (MPC) [11]. MPC operates by repeatedly solving at any sampling time $t \in \mathbb{Z}_{0+}$ the finite horizon optimal control problem

$$\min_{U_t} \quad \|x_{N|t}\|^2_{P_\mathrm{M}} + \sum_{i=0}^{N-1} \left(\|x_{i|t}\|^2_{Q_\mathrm{M}} + \|u_{i|t}\|^2_{R_\mathrm{M}}\right) \tag{35a}$$
$$\text{s.t.} \quad x_{i+1|t} = A_s x_{i|t} + B_s u_{i|t} \tag{35b}$$
$$z_{i|t} = C_s x_{i|t} + D_s u_{i|t} \tag{35c}$$
$$x_{\min} \leq x_{i|t} \leq x_{\max},\ i = 1 \ldots N \tag{35d}$$
$$u_{\min} \leq u_{i|t} \leq u_{\max},\ i = 1 \ldots N \tag{35e}$$
$$z_{\min} \leq z_{i|t} \leq z_{\max},\ i = 1 \ldots N \tag{35f}$$
$$x_{0|t} = x(t) \tag{35g}$$

where $\|a\|^2_Q = a^TQa$, $U_t = [u_{0|t}, \ldots, u_{N-1|t}]$, and $P_\mathrm{M}, Q_\mathrm{M} \succeq 0$, $R_\mathrm{M} \succ 0$. The input applied to the plant is selected from the the optimal solution of (35), $U^*(t)$, as $u(t) = u^*_{0|t}$.

The optimal control problem (35) can be formulated as a family (for varying $x(t)$) of quadratic programs (1) where $\boldsymbol{y}^T = [x^T u^T z^T]$, the equality constraints are defined by the system state and output equations (35b), (35c), the set $\boldsymbol{\mathcal{Y}}$ is a box defined by (35d), (35e), (35f), and the cost function is obtained from (35a). Assumption 1 is satisfied for admissible control systems, Assumption 2 is always satisfied and Assumption 3 is satisfied if $R_\mathrm{M} \succ 0$. The initialization (35g) amounts to changing part of the equality constraint vector, i.e., $\boldsymbol{b}^T = [x(t)^T\ \tilde{b}^T]^T$, at the different sampling instants. All the following numerical tests are executed in Matlab, and we have verified the solution provided by the algorithm in (10)

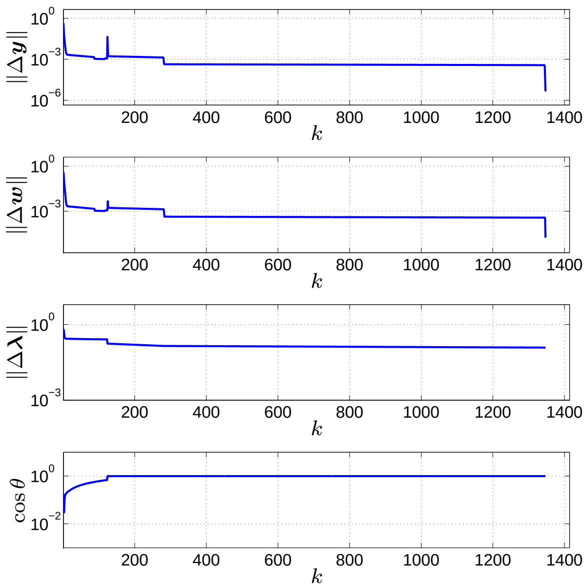

Fig. 1. Example: infeasible QP of MPC of a spacecraft with flexible appendage subject to constraints: evolution of the norm of the ADMM iterates variables/variable differences.

by multiple algorithms, including `quadprog` [12] and the active set QP solver in the `NAG` [13] package.

The example is an MPC controller for a 1-axis re-orientation maneuver for a spacecraft with a flexible appendage, such as a lightweight solar panel. The system is modeled as a 4^{th} order system, see, e.g., [14], which is augmented with two additional states, one for the reference angle to be tracked, and one for the incremental representation of the input. The constraints include upper and lower bound on the torque for controlling the angular rotation of the spacecraft bus, $|u| \leq 0.1$, and bounds on the maximal deflection of the flexible appendage, $|\chi| \leq 0.47$, where χ[m] is the lateral displacement of the tip of the flexible appendage, from the position it holds while at rest.

The MPC controller is designed with sampling period $T_s = 0.25$s, $N = 8$, $R_{\rm M} = 4$, $Q_{\rm M} = P_{\rm M}$, $Q_{\rm M} = C'_M Q_y C_M$, where C_M is a matrix such that $C_M x = (\phi - r_\phi)$, ϕ is the angle of the spacecraft bus, r_ϕ is the reference for such angle, x is the full state of the model, $Q_y = 1$. It may take in the order of $10^3 - 10^4$ iterations for solving the feasible QPs in this example, depending on the initial conditions.

We set as initial condition $x(0) = \begin{bmatrix} 0 & -0.45 & 0.5 & 0 \end{bmatrix}^T$, which correspond to a situation where the appendage is significantly bent in one direction, with a relatively larger velocity toward the opposite direction. While the MPC problem is feasible for $t = 0$, $t = 1$, and for these steps algorithm (10) correctly compute the solution, due to the initial condition and the limited actuation range at $t = 2$ the problem becomes infeasible. In Figure 1 we show the evolution of $\|\Delta \boldsymbol{w}^k\| = \|\boldsymbol{w}^k - \boldsymbol{w}^{k-1}\|$, $\|\Delta \boldsymbol{y}^k\| = \|\boldsymbol{y}^k - \boldsymbol{y}^{k-1}\|$, $\|\Delta \boldsymbol{\lambda}^k\| = \|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^{k-1}\| = \|\boldsymbol{y}^{k-1} - \boldsymbol{w}^{k-1}\|$, and $\cos\theta^k$, where $\cos\theta^k = \frac{(\boldsymbol{\lambda}^k)^T(\boldsymbol{w}^k - \boldsymbol{y}^k)}{\|\boldsymbol{\lambda}^k\|\|\boldsymbol{w}^k - \boldsymbol{y}^k\|}$, during the iterations of algorithm (10) at $t = 2$. The algorithm terminates due to $\|\Delta \boldsymbol{w}^k\|$ and $\|\Delta \boldsymbol{y}^k\|$ being small and $\cos\theta \approx 1$, which is correctly detecting infeasibility. At termination $\|\Delta \boldsymbol{\lambda}\| > 0$, i.e., $\bar{\boldsymbol{w}} \neq \bar{\boldsymbol{y}}$, which detects infeasibility.

VII. Conclusion

We have presented infeasibilty detection in alternating direction method of multipliers for a class of convex QPs. In particular we have established that the divergence of iterates is restricted to the multipliers along the range space of the equality constraints. Furthermore, we have derived the optimal step size parameter that maximizes the rate of convergence for the bounded components of the iterate sequence. In particular, the choice of the optimal step size parameter in the infeasible case is shown to be identical to that of the feasible case derived in [8]. Thus, the proposed step size parameter is guaranteed to work well for both feasible and infeasible instanced of QPs. We have confirmed the approach by numerical results in example applications to the QPs arising from constrained linear MPC.

References

[1] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems in finite-element approximations," *Computers and Mathematics with Applications*, vol. 2, pp. 17–40, 1976.

[2] S. Boyd, N. Parikh, E. Chu, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[3] D. Boley, "Linear convergence of the alternating direction method of multipliers," Dept. of Comp. Sci. and Eng., U. Minnesota, Tech. Rep. TR 12-009, 2012.

[4] M. Hong and Z.-Q. Luo, "On the Linear Convergence of the Alternating Direction Method of Multipliers ," arXiv, Tech. Rep. arXiv:1208.3922, 2013.

[5] R. D. C. Monteiro and B. F. Svaiter, "Iteration-complexity of block-decomposition algorithms and the alternating minimization augmented lagrangian method," *SIAM Journal on Optimization*, vol. 23, no. 1, pp. 475–507, 2013.

[6] E. Ghadimi, A. Teixeira, I. Shames, and M. Johansson, "Optimal parameter selection for the alternating direction method of multipliers (ADMM): quadratic problems," arXiv, Tech. Rep. arXiv:1306.2454v1, 2013.

[7] A. U. Raghunathan and S. Di Cairano, "Alternating Direction Method of Multipliers (ADMM) for Strictly Convex Programs : Optimal Parameter Selection," in *Proceedings of American Control Conference*, 2014, pp. 4324–4329.

[8] ——, "Optimal step-size selection in alternating direction method of multipliers for convex quadratic programs and model predictive control ," in *Proceedings of Symposium on Mathematical Theory of Networks and Systems*, 2014, pp. 807–814.

[9] J. Eckstein and D. P. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming: Series A*, vol. 55, no. 3, pp. 293–318, 1992.

[10] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[11] C. Garcia, D. Prett, and M. Morari, "Model predictive control: theory and practice–a survey," *Automatica(Oxford)*, vol. 25, no. 3, pp. 335–348, 1989.

[12] The Mathworks Inc., *Matlab Optimization Toolbox*. Natick, MA: The Mathworks Inc., 2011, http://www.mathworks.com.

[13] NAG Ltd., `e04nf`, *E04 Minimizing or Maximizing a Function*, in *NAG Toolbox for MATLAB*. Oxford, U.K.: NAG Ltd., 2011, http://www.nag.co.uk.

[14] S. Di Cairano and I. V. Kolmanovsky, "Further developments and applications of network reference governor for constrained systems,"