Improving Fast Dual Ascent for MPC - Part II: The Embedded Case \star

Pontus Giselsson*

* Electrical Engineering, Stanford University (e-mail: pontusg@stanford.edu).

Abstract:

Recently, several authors have suggested the use of first order methods, such as fast dual ascent and the alternating direction method of multipliers, for embedded model predictive control. The main reason is that they can be implemented using simple arithmetic operations only. However, a known limitation of gradient-based methods is that they are sensitive to ill-conditioning of the problem data. In this paper, we present a fast dual gradient method for which the sensitivity to ill-conditioning is greatly reduced. This is achieved by approximating the negative dual function with a quadratic upper bound with different curvature in different directions in the algorithm, as opposed to having the same curvature in all directions as in standard fast gradient methods. The main contribution of this paper is a characterization of the set of matrices that can be used to form such a quadratic upper bound to the negative dual function. We also describe how to choose a matrix from this set to get an improved approximation of the dual function, especially if it is ill-conditioned, compared to the approximation used in standard fast dual gradient methods. This can give a significantly improved performance as illustrated by a numerical evaluation on an ill-conditioned AFTI-16 aircraft model.

1. INTRODUCTION

Several authors including O'Donoghue et al. (2013); Jerez et al. (2013); Richter et al. (2013); Patrinos and Bemporad (2014) have recently proposed first order optimization methods as appropriate for embedded model predictive control. In O'Donoghue et al. (2013); Jerez et al. (2013), the alternating direction method of multipliers (ADMM, see Boyd et al. (2011)) were used and high computational speeds were reported when implemented on embedded hardware. In Richter et al. (2013); Patrinos and Bemporad (2014), the optimal control problems arising in model predictive were solved using different formulations of fast dual gradient methods. In Richter et al. (2013), the equality constraints, i.e. the dynamic constraints, are dualized and a diagonal cost and box constraints are assumed. The resulting dual problem is solved using a fast gradient method. In Patrinos and Bemporad (2014), the same splitting as in O'Donoghue et al. (2013); Jerez et al. (2013) is used, but a fast gradient method is used to solve the resulting problem as opposed to ADMM in O'Donoghue et al. (2013); Jerez et al. (2013). In this paper, we will show how to improve and generalize the fast dual gradient methods presented in Richter et al. (2013); Patrinos and Bemporad (2014).

Fast gradient methods as used in Richter et al. (2013); Patrinos and Bemporad (2014) have been around since the early 80's when the seminal paper Nesterov (1983) was published. However, fast gradient methods did not render much attention before the mid 00's, after which an increasing interest has emerged. Several extensions and generalizations of the fast gradient method have been proposed, e.g. in Nesterov (2003, 2005). In Beck and Teboulle (2009), the method was generalized to allow for minimization of composite objective functions. Further, a unified framework for fast gradient methods and their generalizations were presented in Tseng (2008). To use fast gradient methods for composite minimization, one objective term should be convex and differentiable with a Lipschitz continuous gradient, while the other should be proper, closed, and convex. The former condition is equivalent to the existence of a quadratic upper bound to the function, with the same curvature in all directions. The curvature is specified by the Lipschitz constant to the gradient. In fast gradient methods, the quadratic upper bound serves as an approximation of the function to be minimized, since the bound is minimized in every iteration of the algorithm. If the quadratic upper bound does not well approximate the function to be minimized, slow convergence properties are expected. By instead allowing for a quadratic upper bound with different curvature in different directions, as in generalized fast gradient methods Zuo and Lin (2011), the bound can closer approximate the function to be minimized. For an appropriate choice of non-uniform quadratic upper bound, this can significantly improve the performance of the algorithm.

In (Nesterov, 2005, Theorem 1), a Lipschitz constant to the gradient of the dual function to strongly convex problems is presented. This result quantifies the curvature of

^{*} During the preparation of this paper, the author was a member of the LCCC Linnaeus Center at Lund University. Financial support from the Swedish Research Council for the author's Postdoctoral studies at Stanford University is gratefully acknowledged. Further, Eric Chu is gratefully acknowledged for constructive feedback and Alexander Domahidi is gratefully acknowledged for suggesting the AFTI-16 control problem as a benchmark and providing FORCES code for the same.

a uniform quadratic upper bound to the negative dual function. This result was improved in (Richter et al., 2013, Theorem 7) when the primal cost is restricted to being quadratic. Using these quadratic upper bounds, with the same curvature in all directions, as dual function approximation in a fast dual gradient method, may result in slow convergence rates. Especially for ill-conditioned problems where the upper bound does not well approximate the negative dual function. In this paper, the main result is a new characterization of the set of matrices that can be used to describe quadratic upper bounds to the negative dual function. This result generalizes and improves previous results in Nesterov (2005); Richter et al. (2013). We also show how to appropriately choose a matrix from this set to get a quadratic upper bound that well approximates the negative dual function. Since in the proposed method, the dual function approximation is better that in standard fast dual gradient methods used in Richter et al. (2013); Patrinos and Bemporad (2014), better convergence rate properties are expected.

In model predictive control, much offline computational effort can be devoted to improve the online execution time of the solver. This is done, e.g., in explicit MPC, see Bemporad et al. (2002), where the explicit parametric solution is computed beforehand, and found through a look-up table online. In this paper, the offline computational effort is devoted to choose a matrix that describes the quadratic upper bound to the negative dual function. The computed matrix is the same in all samples in the controller and can therefore be computed offline. The algorithm is evaluated on a pitch control problem in an AFTI-16 aircraft that has previously been studied in Kapasouris et al. (1990); Bemporad et al. (1997). This is a challenging problem for first order methods since it is very ill-conditioned. The numerical evaluation shows that the method presented in this paper outperforms other first-order methods presented in O'Donoghue et al. (2013); Jerez et al. (2013); Richter et al. (2013); Patrinos and Bemporad (2014) with one to three orders of magnitude. Also, the numerical evaluation shows that a C implementation of our algorithm outperform FORCES, Domahidi et al. (2012), which is a C codegenerator for MPC problems using a tailored interior point method, and the general commercial QP-solver MOSEK.

This paper extends the conference publication Giselsson (2014b), and is the second of a series of two papers on improving duality-based optimization in MPC, with Giselsson (2014a) being the first.

2. PRELIMINARIES AND NOTATION

2.1 Notation

We denote by \mathbb{R} , \mathbb{R}^n , $\mathbb{R}^{m \times n}$, the sets of real numbers, vectors, and matrices. $\mathbb{S}^n \subseteq \mathbb{R}^{n \times n}$ is the set of symmetric matrices, and $\mathbb{S}_{++}^n \subseteq \mathbb{S}^n$, $[\mathbb{S}_{+}^n] \subseteq \mathbb{S}^n$, are the sets of positive [semi] definite matrices. Further, $L \succeq M$ and $L \succ M$ where $L, M \in \mathbb{S}^n$ denotes $L - M \in \mathbb{S}_{+}^n$ and $L - M \in \mathbb{S}_{+}^n$ respectively. We also use notation $\langle x, y \rangle = x^T y$, $\langle x, y \rangle_H = x^T H y$, $\|x\|_2 = \sqrt{x^T x}$, and $\|x\|_H = \sqrt{x^T H x}$. Finally, $I_{\mathcal{X}}$ denotes the indicator function for the set \mathcal{X} , i.e. $I_{\mathcal{X}}(x) \triangleq \begin{cases} 0, & x \in \mathcal{X} \\ \infty, & \text{else} \end{cases}$.

$2.2 \ Preliminaries$

In this section, we introduce generalizations of well used concepts. We generalize the notion of strong convexity as well as the notion of Lipschitz continuity of the gradient of convex functions. We also define conjugate functions and state a known result on dual properties of a function and its conjugate.

For differentiable and convex functions $f : \mathbb{R}^n \to \mathbb{R}$ that have a Lipschitz continuous gradient with constant L, we have that

$$\|\nabla f(x_1) - \nabla f(x_2)\|_2 \le L \|x_1 - x_2\|_2 \tag{1}$$

holds for all $x_1, x_2 \in \mathbb{R}^n$. This is equivalent to that

$$f(x_1) \le f(x_2) + \langle \nabla f(x_2), x_1 - x_2 \rangle + \frac{L}{2} \|x_1 - x_2\|_2^2 \quad (2)$$

holds for all $x_1, x_2 \in \mathbb{R}^n$ (Nesterov, 2003, Theorem 2.1.5). In this paper, we allow for a generalized version of the quadratic upper bound (2) to f, namely that

$$f(x_1) \le f(x_2) + \langle \nabla f(x_2), x_1 - x_2 \rangle + \frac{1}{2} \|x_1 - x_2\|_{\mathbf{L}}^2 \quad (3)$$

holds for all $x_1, x_2 \in \mathbb{R}^n$ where $\mathbf{L} \in \mathbb{S}^n_+$. The bound (2) is obtained by setting $\mathbf{L} = LI$ in (3).

Remark 1. For concave functions f, i.e. where -f is convex, the Lipschitz condition (1) is equivalent to that the following quadratic lower bound

$$f(x_1) \ge f(x_2) + \langle \nabla f(x_2), x_1 - x_2 \rangle - \frac{L}{2} \|x_1 - x_2\|_2^2 \quad (4)$$

holds for all $x_1, x_2 \in \mathbb{R}^n$. The generalized counterpart naturally becomes that

$$f(x_1) \ge f(x_2) + \langle \nabla f(x_2), x_1 - x_2 \rangle - \frac{1}{2} \|x_1 - x_2\|_{\mathbf{L}}^2$$
(5)

holds for all $x_1, x_2 \in \mathbb{R}^n$.

Next, we state a Lemma on equivalent characterizations of the condition (3).

Lemma 2. Assume that $f : \mathbb{R}^n \to \mathbb{R}$ is convex and differentiable. The condition that

$$f(x_1) \le f(x_2) + \langle \nabla f(x_2), x_1 - x_2 \rangle + \frac{1}{2} \|x_1 - x_2\|_{\mathbf{L}}^2 \quad (6)$$

holds for some $\mathbf{L} \in \mathbb{S}^n_+$ and all $x_1, x_2 \in \mathbb{R}^n$ is equivalent to that

$$\langle \nabla f(x_1) - \nabla f(x_2), x_1 - x_2 \rangle \le ||x_1 - x_2||_{\mathbf{L}}^2.$$
(7)
holds for all $x_1, x_2 \in \mathbb{R}^n$.

Proof. To show the equivalence, we introduce the function $g(x) := \frac{1}{2}x^T \mathbf{L}x - f(x)$. According to (Nesterov, 2003, Theorem 2.1.3) and since g is differentiable, $g : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if ∇g is monotone. The function g is convex if and only if

$$g(x_1) \ge g(x_2) + \langle \nabla g(x_2), x_1 - x_2 \rangle = = \frac{1}{2} x_2^T \mathbf{L} x_2 - f(x_2) + \langle \mathbf{L} x_2 - \nabla f(x_2), x_1 - x_2 \rangle = -f(x_2) - \langle \nabla f(x_2), x_1 - x_2 \rangle - \frac{1}{2} ||x_1 - x_2||_{\mathbf{L}}^2 + \frac{1}{2} x_1^T \mathbf{L} x_1.$$

Noting that $g(x_1) = \frac{1}{2} x_1^T \mathbf{L} x_1 - f(x_1)$ gives the negated version of (6).

Monotonicity of ∇g is equivalent to

$$D \leq \langle \nabla g(x_1) - \nabla g(x_2), x_1 - x_2 \rangle$$

= $\langle \mathbf{L}x_1 - \nabla f(x_1) - \mathbf{L}x_2 + \nabla f(x_2), x_1 - x_2 \rangle$
= $\|x_1 - x_2\|_{\mathbf{L}}^2 - \langle \nabla f(x_1) - \nabla f(x_2), x_1 - x_2 \rangle.$

Rearranging the terms gives (7). This concludes the proof.

Next, we state the corresponding result for concave functions.

Corollary 3. Assume that $f : \mathbb{R}^n \to \mathbb{R}$ is concave and differentiable. The condition that

$$f(x_1) \ge f(x_2) + \langle \nabla f(x_2), x_1 - x_2 \rangle - \frac{1}{2} \|x_1 - x_2\|_{\mathbf{L}}^2 \quad (8)$$

holds for some $\mathbf{L} \in \mathbb{S}^n_+$ and all $x_1, x_2 \in \mathbb{R}^n$ is equivalent to that

$$\langle \nabla f(x_1) - \nabla f(x_2), x_2 - x_1 \rangle \le ||x_1 - x_2||_{\mathbf{L}}^2.$$
(9)
holds for all $x_1, x_2 \in \mathbb{R}^n$.

Proof. The proof follows directly from -f being convex and applying Lemma 2.

The standard definition of a differentiable and strongly convex function $f : \mathbb{R}^n \to \mathbb{R}$ is that it satisfies

$$f(x_1) \ge f(x_2) + \langle \nabla f(x_2), x_1 - x_2 \rangle + \frac{\sigma}{2} \|x_1 - x_2\|_2^2 \quad (10)$$

for any $x_1, x_2 \in \mathbb{R}^n$, where the modulus $\sigma \in \mathbb{R}_{++}$ describes a lower bound on the curvature of the function. In this paper, the definition (10) is generalized to allow for a quadratic lower bound with different curvature in different directions.

Definition 4. A differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is strongly convex with matrix H if and only if

$$f(x_1) \ge f(x_2) + \langle \nabla f(x_2), x_1 - x_2 \rangle + \frac{1}{2} \|x_1 - x_2\|_H^2$$

holds for all $x_1, x_2 \in \mathbb{R}^n$, where $H \in \mathbb{S}^n_{++}$.

Remark 5. The traditional definition of strong convexity (10) is obtained from Definition 4 by setting $H = \sigma I$.

Lemma 6. Assume that $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable and strongly convex with matrix H. The condition that

$$f(x_1) \ge f(x_2) + \langle \nabla f(x_2), x_1 - x_2 \rangle + \frac{1}{2} \|x_1 - x_2\|_H^2 \quad (11)$$

holds for all $x_1, x_2 \in \mathbb{R}^n$ is equivalent to that

$$\langle \nabla f(x_1) - \nabla f(x_2), x_1 - x_2 \rangle \ge ||x_1 - x_2||_H^2$$
 (12)
holds for all $x_1, x_2 \in \mathbb{R}^n$.

Proof. To show the equivalence, we introduce the function $g(x) := f(x) - \frac{1}{2}x^T Hx$ and proceed similarly to in the proof of Lemma (2). According to (Nesterov, 2003, Theorem 2.1.3) and since g is differentiable, $g : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if ∇g is monotone. The function g is convex if and only if

$$\begin{split} g(x_1) &\geq g(x_2) + \langle \nabla g(x_2), x_1 - x_2 \rangle = \\ &= f(x_2) - \frac{1}{2} x_2^T H x_2 + \langle \nabla f(x_2) - H x_2, x_1 - x_2 \rangle \\ &= f(x_2) + \langle \nabla f(x_2), x_1 - x_2 \rangle + \frac{1}{2} \|x_1 - x_2\|_H^2 - \frac{1}{2} x_1^T H x_1. \end{split}$$

Noting that $g(x_1) = f(x_1) - \frac{1}{2} x_1^T H x_1$ gives (11).

Monotonicity of ∇g is equivalent to

$$0 \le \langle \nabla g(x_1) - \nabla g(x_2), x_1 - x_2 \rangle = \langle \nabla f(x_1) - Hx_1 - \nabla f(x_2) + \mathbf{L}x_2, x_1 - x_2 \rangle = \langle \nabla f(x_1) - \nabla f(x_2), x_1 - x_2 \rangle - \|x_1 - x_2\|_{H^1}^2.$$

Rearranging the terms gives (12). This concludes the proof.

The condition (11) is a quadratic lower bound on the function value, while the condition (3) is a quadratic upper

bound on the function value. These two properties are linked through the conjugate function

$$f^{\star}(y) \triangleq \sup_{x} \left\{ y^T x - f(x) \right\}.$$

More precisely, we have the following result.

Proposition 7. Assume that $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is closed, proper, and strongly convex with modulus σ on the relative interior of its domain. Then the conjugate function f^* is convex and differentiable, and $\nabla f^*(y) = x^*(y)$, where $x^*(y) = \arg \max_x \{y^T x - f(x)\}$. Further, ∇f^* is Lipschitz continuous with constant $L = \frac{1}{\sigma}$.

A straight-forward generalization is given by the chain-rule and was proven in (Nesterov, 2005, Theorem 1) (which also proves the less general Proposition 7).

Corollary 8. Assume that $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is closed, proper, and strongly convex with modulus σ on the relative interior of its domain. Further, define $g^*(y) \triangleq f^*(Ay)$. Then g^* is convex and differentiable, and $\nabla g^*(y) = A^T x^*(Ay)$, where $x^*(Ay) = \arg \max_x \{(Ay)^T x - f(x)\}$. Further, ∇g^* is Lipschitz continuous with constant $L = \frac{\|A\|_2^2}{2}$.

For the case when $f(x) = \frac{1}{2}x^T H x + g^T x$, i.e. f is a quadratic, a tighter Lipschitz constant to $\nabla g^*(y) = \nabla f^*(Ay)$ was provided in (Richter et al., 2013, Theorem 7), namely $L = ||AH^{-1}A^T||_2$.

3. PROBLEM FORMULATION

We consider optimization problems of the form

minimize
$$f(x) + h(x) + g(Bx)$$
 (13)
subject to $Ax = b$

where $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{p \times n}$, $b \in \mathbb{R}^m$. We assume that the following assumption holds throughout the paper: Assumption 9.

- (a) The function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable and strongly convex with matrix H.
- (b) The extended valued functions $h : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ and $g : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, are proper, closed, and convex.
- (c) $A \in \mathbb{R}^{m \times n}$ has full row rank.

Remark 10. Examples of functions that satisfy Assumption 9(a) and 9(b) are $f(x) = \frac{1}{2}x^T H x + g^T x$ with $H \in \mathbb{S}^{n}_{++}$ for Assumption 9(a), and $g = I_{\mathcal{X}}, g = \|\cdot\|_1, g = I^*_{\mathcal{X}}$, or g = 0 for Assumption 9(b). If Assumption 9(c) is not satisfied, redundant equality constraints can, without affecting the solution of (13), be removed to satisfy the assumption.

The optimization problem (13) can equivalently be written as

minimize
$$f(x) + h(x) + g(y)$$
 (14)
subject to $Ax = b$
 $Bx = y$

We introduce dual variables $\lambda \in \mathbb{R}^m$ for the equality constraints Ax = b and dual variables $\mu \in \mathbb{R}^p$ for the equality constraints Bx = y. This gives the following Lagrange dual problem

$$\sup_{\lambda,\mu} \inf_{x,y} \left\{ f(x) + h(x) + \lambda^T (Ax - b) + g(y) + \mu^T (Bx - y) \right\}$$
$$= \sup_{\lambda,\mu} \left[-\sup_x \left\{ (-A^T \lambda - B^T \mu)^T x - f(x) - h(x) \right\} - b^T \lambda - \sup_y \left\{ \mu^T y - g(y) \right\} \right]$$

$$= \sup_{\lambda,\mu} \left\{ -F^{\star}(-A^{T}\lambda - B^{T}\mu) - b^{T}\lambda - g^{\star}(\mu) \right\}$$
(15)

where F^* is the conjugate function to F := f + h and g^* is the conjugate function to g. For ease of exposition, we introduce $\nu = (\lambda, \mu) \in \mathbb{R}^{m+p}$, $C = [A^T B^T]^T \in \mathbb{R}^{(m+p) \times n}$, and $c = (b, 0) \in \mathbb{R}^{m+p}$ and the following function

$$d(\nu) := -F^{\star}(-C^{T}\nu) - c^{T}\nu = -F^{\star}(-A^{T}\lambda - B^{T}\mu) - b^{T}\lambda.$$
(16)

This implies that the dual problem (15) can be written as

m

aximize
$$d(\nu) - g^{\star}([0 \ I]\nu).$$
 (17)

To evaluate the function d, an optimization problem is solved. The minimand to this problem is denoted by

$$x^{\star}(\nu) := \arg\min_{x} \left\{ F(x) + \nu^{T} C x \right\}$$
(18)
$$= \arg\min_{x} \left\{ f(x) + h(x) + \lambda^{T} A x + \mu^{T} B x \right\}.$$

From Corollary 8 we get that the function d is concave and differentiable with gradient

$$\nabla d(\nu) = Cx^{\star}(\nu) - c$$

and that ∇d is Lipschitz continuous with constant $L = \|C\|_2^2 / \lambda_{\min}(H)$, i.e., that

$$\|\nabla d(\nu_1) - \nabla d(\nu_2)\|_2 \le L \|\nu_1 - \nu_2\|_2 \tag{19}$$

holds for all $\nu_1, \nu_2 \in \mathbb{R}^{m+p}$. As stated in Remark 1, (19) is equivalent to that the following quadratic lower bound to the concave function d holds for all $\nu_1, \nu_2 \in \mathbb{R}^{m+p}$

$$d(\nu_1) \ge d(\nu_2) + \langle \nabla d(\nu_2), \nu_1 - \nu_2 \rangle - \frac{L}{2} \|\nu_1 - \nu_2\|_2^2$$

In the following section we will show that the function d satisfies the following tighter condition

$$d(\nu_1) \ge d(\nu_2) + \langle \nabla d(\nu_2), \nu_1 - \nu_2 \rangle - \frac{1}{2} \|\nu_1 - \nu_2\|_{\mathbf{L}}^2 \quad (20)$$

for all $\nu_1, \nu_2 \in \mathbb{R}^{m+p}$ and $\mathbf{L} \in \mathbb{S}^{m+p}_+$ that satisfies $\mathbf{L} \succeq CH^{-1}C^T$.

4. DUAL FUNCTION PROPERTIES

To show that the function d as defined in (16) satisfies (20), we need the following lemma.

Lemma 11. Suppose that Assumption 9 holds. Then

$$x^{\star}(\nu_1) - x^{\star}(\nu_2) \|_H \le \|\nu_1 - \nu_2\|_{CH^{-1}C}$$

holds for all $\nu_1, \nu_2 \in \mathbb{R}^{m+p}$, where $x^*(\nu)$ is defined in (18).

Proof. We first show that

$$\langle \nabla f(x^{\star}(\nu_{1})) - \nabla f(x^{\star}(\nu_{2})), x^{\star}(\nu_{1}) - x^{\star}(\nu_{2}) \rangle \leq \langle C^{T}(\nu_{1} - \nu_{2}), x^{\star}(\nu_{2}) - x^{\star}(\nu_{1}) \rangle.$$
 (21)

First order optimality conditions of (18) with ν_1 and ν_2 respectively are

$$0 \in \nabla f(x^{\star}(\nu_1)) + \partial h(x^{\star}(\nu_1)) + C^T \nu_1, \qquad (22)$$

$$0 \in \nabla f(x^{\star}(\nu_2)) + \partial h(x^{\star}(\nu_2)) + C^T \nu_2.$$
(23)

We denote by $\xi(x^*(\nu_1)) \in \partial h(x^*(\nu_1))$ and $\xi(x^*(\nu_2)) \in \partial h(x^*(\nu_2))$ the sub-gradients that give equalities in (22) and (22) respectively. This gives

$$0 = \nabla f(x^{\star}(\nu_1)) + \xi(x^{\star}(\nu_1)) + C^T \nu_1, \qquad (24)$$

$$0 = \nabla f(x^{\star}(\nu_2)) + \xi(x^{\star}(\nu_2)) + C^T \nu_2.$$
 (25)

Taking the scalar product of (24) with $x^*(\nu_2) - x^*(\nu_1)$ and the scalar product of (25) with $x^*(\nu_1) - x^*(\nu_2)$, and summing gives

$$\begin{aligned} \langle \nabla f(x^{\star}(\nu_{1})) - \nabla f(x^{\star}(\nu_{2})), x^{\star}(\nu_{1}) - x^{\star}(\nu_{2}) \rangle + \\ + \langle C^{T}(\nu_{1} - \nu_{2}), x^{\star}(\nu_{1}) - x^{\star}(\nu_{2}) \rangle = \\ = \langle \xi(x^{\star}(\nu_{1})) - \xi(x^{\star}(\nu_{2})), x^{\star}(\nu_{2}) - x^{\star}(\nu_{1}) \rangle \leq 0 \end{aligned}$$

where the inequality holds since sub-differentials of proper, closed, and convex functions are (maximal) monotone mappings, see (Rockafellar, 1970 §24). This implies that (21) holds.

Further

$$\begin{aligned} \|x^{\star}(\nu_{1}) - x^{\star}(\nu_{2})\|_{H}^{2} &\leq \\ &\leq \langle \nabla f(x^{\star}(\nu_{1})) - \nabla f(x^{\star}(\nu_{2})), x^{\star}(\nu_{1}) - x^{\star}(\nu_{2}) \rangle \\ &\leq \langle C^{T}(\nu_{1} - \nu_{2}), x^{\star}(\nu_{2}) - x^{\star}(\nu_{1}) \rangle \\ &= \langle H^{-1/2}C^{T}(\nu_{1} - \nu_{2}), H^{1/2}(x^{\star}(\nu_{2}) - x^{\star}(\nu_{1})) \rangle \\ &\leq \|H^{-1/2}C^{T}(\nu_{1} - \nu_{2})\|_{2} \|x^{\star}(\nu_{2}) - x^{\star}(\nu_{1})\|_{H} \end{aligned}$$

where the first inequality comes from Lemma 6, the second from (21), and the final inequality is due to Cauchy Schwarz inequality. This implies that

$$\|x^{\star}(\nu_{1}) - x^{\star}(\nu_{2})\|_{H} \le \|\nu_{1} - \nu_{2}\|_{CH^{-1}C^{T}}$$

which concludes the proof.

Now we are ready to state the main theorem of this section. Theorem 12. The function d defined in (16) is concave, differentiable and satisfies

$$d(\nu_1) \ge d(\nu_2) + \langle \nabla d(\nu_2), \nu_1 - \nu_2 \rangle - \frac{1}{2} \|\nu_1 - \nu_2\|_{\mathbf{L}}^2 \quad (26)$$

for every $\nu_1, \nu_2 \in \mathbb{R}^{m+p}$ and $\mathbf{L} \in \mathbb{S}^{m+p}_+$ that satisfies $\mathbf{L} \succeq CH^{-1}C^T$.

Proof. Concavity and differentiability is deduced from Danskin's Theorem, see (Bertsekas, 1999, Proposition B.25).

To show (26), we have for any $\nu_1, \nu_2 \in \mathbb{R}^{m+p}$ that

$$\begin{split} \langle \nabla d(\nu_1) - \nabla d(\nu_2), \nu_2 - \nu_1 \rangle &= \\ &= \langle Cx^*(\nu_1) - c - Cx^*(\nu_2) + c, \nu_2 - \nu_1 \rangle \\ &= \langle x^*(\nu_1) - x^*(\nu_2), C^T(\nu_2 - \nu_1) \rangle \\ &= \langle x^*(\nu_1) - x^*(\nu_2), H^{-1}C^T(\nu_2 - \nu_1) \rangle_H \\ &\leq \|x^*(\nu_1) - x^*(\nu_2)\|_H \|H^{-1}C^T(\nu_2 - \nu_1)\|_H \\ &\leq \|H^{-1}C^T(\nu_2 - \nu_1)\|_H^2 \\ &= (\nu_2 - \nu_1)^T C H^{-1}C^T(\nu_2 - \nu_1) \\ &= \|\nu_2 - \nu_1\|_{CH^{-1}C^T}^2 \end{split}$$

where the first inequality is due to Cauchy-Schwarz inequality and the second comes from Lemma 11. Applying Corollary 3 gives the result.

Next, we show that if f is a strongly convex quadratic function and h satisfies certain conditions, then Theorem 12 gives the best possible bound of the form (26).

Proposition 13. Assume that $f(x) = \frac{1}{2}x^THx + \zeta^T x$ with $H \in \mathbb{S}^n_{++}$ and $\zeta \in \mathbb{R}^n$ and that there exists a set

 $\mathcal{X} \subseteq \mathbb{R}^n$ with non-empty interior on which h is linear, i.e. $h(x) = \xi_{\mathcal{X}}^T x + \theta_{\mathcal{X}}$ for all $x \in \mathcal{X}$. Further, assume that there exists $\tilde{\nu}$ such that $x^*(\tilde{\nu}) \in \operatorname{int}(\mathcal{X})$. Then for any matrix $\mathbf{L} \not\succeq CH^{-1}C^T$, there exist ν_1 and ν_2 such that (26) does not hold.

Proof. Since $x^*(\widetilde{\nu}) \in \operatorname{int}(\mathcal{X})$ we get for all $\nu_{\epsilon} \in \mathcal{B}_{\epsilon}^{m+p}(0)$, where the radius ϵ is small enough, that $x^*(\widetilde{\nu}) - H^{-1}C^T\nu_{\epsilon} \in \mathcal{X}$. Introducing $x_{\epsilon} = -H^{-1}C^T\nu_{\epsilon}$, we get from the optimality conditions to (18) (that specifies $x^*(\nu)$) that

$$0 = Hx^{*}(\widetilde{\nu}) + \zeta + \xi_{\mathcal{X}} + C^{T}\widetilde{\nu}$$

= $H(x^{*}(\widetilde{\nu}) + x_{\epsilon}) + \zeta + \xi_{\mathcal{X}} + C^{T}(\widetilde{\nu} + \nu_{\epsilon})$
= $H(x^{*}(\widetilde{\nu}) + x_{\epsilon}) + \zeta + h'(x^{*}(\widetilde{\nu}) + x_{\epsilon}) + C^{T}(\widetilde{\nu} + \nu_{\epsilon})$

where $h'(x^{\star}(\tilde{\nu}) \in \partial h(x^{\star}(\tilde{\nu}) \text{ and } x^{\star}(\tilde{\nu}) + x_{\epsilon} \in \mathcal{X} \text{ is used in the last step. This implies that } x^{\star}(\tilde{\nu} + \nu_{\epsilon}) = x^{\star}(\tilde{\nu}) + x_{\epsilon} \text{ and consequently that } x^{\star}(\tilde{\nu} + \nu_{\epsilon}) \in \mathcal{X} \text{ for any } \nu_{\epsilon} \in \mathcal{B}_{\epsilon}^{m+p}(0).$ Thus, for any $\nu \in \tilde{\nu} \oplus \mathcal{B}_{\epsilon}^{m+p}(0)$ we get

$$d(\nu) = \min_{x} \frac{1}{2} x^{T} H x + \zeta^{T} x + h(x) + \nu^{T} (Cx - c)$$

= $\min_{x} \frac{1}{2} x^{T} H x + \zeta^{T} x + \xi^{T} x + \nu^{T} (Cx - c)$
= $-\frac{1}{2} \nu^{T} C H^{-1} C^{T} \nu + \xi^{T} \nu + \theta$

where $\xi \in \mathbb{R}^n$ and $\theta \in \mathbb{R}$ collects the linear and constant terms respectively. Since on the set $\tilde{\nu} \oplus \mathcal{B}_{\epsilon}^{m+p}(0)$, d is a quadratic with Hessian $CH^{-1}C^T$, it is straight-forward to verify that (26) holds with equality for all $\nu_1, \nu_2 \in \tilde{\nu} \oplus \mathcal{B}_{\epsilon}^{m+p}(0)$ if $\mathbf{L} = CH^{-1}C^T$. Thus, since $\tilde{\nu} \oplus \mathcal{B}_{\epsilon}^{m+p}(0)$ has non-empty interior, we can for any matrix $\mathbf{L} \succeq CH^{-1}C^T$ find $\nu_1, \nu_2 \in \tilde{\nu} \oplus \mathcal{B}_{\epsilon}^{m+p}(0)$ such that

$$\|\nu_1 - \nu_2\|_{CH^{-1}C^T} \ge \|\nu_1 - \nu_2\|_{\mathbf{L}}.$$

This implies that for any $\mathbf{L} \not\succeq CH^{-1}C^T$ there exist $\nu_1, \nu_2 \in \widetilde{\nu} \oplus \mathcal{B}_{\epsilon}^{m+p}(0)$ such that

$$d(\nu_1) = d(\nu_2) + \langle \nabla d(\nu_2), \nu_1 - \nu_2 \rangle - \frac{1}{2} \|\nu_1 - \nu_2\|_{CH^{-1}C^T}$$

$$\leq d(\nu_2) + \langle \nabla d(\nu_2), \nu_1 - \nu_2 \rangle - \frac{1}{2} \|\nu_1 - \nu_2\|_{\mathbf{L}}$$

This concludes the proof.

Proposition 13 shows that the bound in Theorem 12 is indeed the best obtainable bound of the form (26) if fis a quadratic and h specifies the stated assumptions. Examples of functions that satisfy the assumptions on h in Proposition 13 include linear functions, indicator functions of closed convex constraint sets with non-empty interior, and the 1-norm. However, indicator functions for affine subspaces do not satisfy the the assumptions of Proposition 13 since their interiors are empty (except for the trivial sub-space \mathbb{R}^n). In the following proposition we will present a result that shows how Theorem 26 can be improved in that case.

Proposition 14. Assume that $f(x) = \frac{1}{2}x^T H x + \zeta^T x$ with $H \in \mathbb{S}^n_{++}$ and $\zeta \in \mathbb{R}^n$, and that $h = I_{Ax=b}$. Then (26) holds for all $\mathbf{L} \in \mathbb{S}^{m+p}_+$ such that $\mathbf{L} \succeq CH^{-1/2}(I-M)H^{-1/2}C^T$ where $M = H^{-1/2}A^T(AH^{-1}A^T)^{-1}AH^{-1/2}$. Further, for any matrix $\mathbf{L} \succeq CH^{-1/2}(I-M)H^{-1/2}C^T$ there exist $\nu_1, \nu_2 \in \mathbb{R}^{m+p}$ such that (26) does not hold.

Proof. We have

$$d(\nu) = -F^{\star}(-C^{T}\nu) - c^{T}\nu$$

= $-\sup_{x} \left(-\nu^{T}Cx - f(x) - h(x)\right) - c^{T}\nu$
= $\inf_{x} \left(\nu^{T}Cx + \frac{1}{2}x^{T}Hx + \zeta^{T}x + I_{Ax=b}(x)\right) - c^{T}\nu$
(27)

since F = f + h. The solution $x^{\star}(\nu)$ to the minimization problem satisfies the following KKT-equations

$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^*(\nu) \\ \lambda^*(\nu) \end{bmatrix} = \begin{bmatrix} -C^T \nu - \zeta \\ b \end{bmatrix}$$
(28)

where $\lambda^{\star}(\nu)$ are dual variables corresponding to the equality constraints. We have

$$x^{\star}(\nu) = -H^{-1}(A^T\lambda^{\star}(\nu) + C^T\nu + \zeta).$$

Inserting this into the second set of equations in (28) gives

$$-AH^{-1}(A^T\lambda^*(\nu) + C^T\nu + \zeta) = b.$$

Since by assumption A has full row rank and H in positive definite, $AH^{-1}A^{T}$ is invertible. Introducing the notation $H_{A} = AH^{-1}A^{T}$, this implies that

$$\lambda^{\star}(\nu) = -H_A^{-1}(AH^{-1}(C^T\nu + \zeta) + b)$$

which in turn implies that

$$\begin{split} x^{\star}(\nu) &= H^{-1}(A^{T}H_{A}^{-1}(AH^{-1}(C^{T}\nu+\zeta)+b) - C^{T}\nu-\zeta) \\ &= H^{-1}(A^{T}H_{A}^{-1}AH^{-1}-I)(C^{T}\nu+\zeta) + H^{-1}A^{T}H_{A}^{-1}b \\ &= -H^{-1/2}(I-M)H^{-1/2}(C^{T}\nu+\zeta) + H^{-1}A^{T}H_{A}^{-1}b. \end{split}$$

Insertion of this into (27) gives after straight-forward computations that

$$d(\nu) = -\frac{1}{2}\nu^T C H^{-1/2} (I - M) H^{-1/2} C^T \nu + \xi^T \nu + \theta$$

where $\xi \in \mathbb{R}^{m+p}$ and $\theta \in \mathbb{R}$ collect the linear and constant terms respectively. This implies that d is a concave quadratic function with negative Hessian $CH^{-1}(I - M)H^{-1/2}C^T$. For concave quadratic functions, it is straight-forward to verify that (26) holds with equality for all $\nu_1, \nu_2 \in \mathbb{R}^{m+p}$ if **L** is chosen as the negative Hessian, i.e. $\mathbf{L} = CH^{-1/2}(I - M)H^{-1/2}C^T$. This further implies, that for any $\mathbf{L} \succeq CH^{-1/2}(I - M)H^{-1/2}C^T$ there exist $\nu_1, \nu_2 \in \mathbb{R}^{m+p}$ such that (26) does not hold. This concludes the proof.

For the preceding result to hold, it is actually sufficient to assume that f is strongly convex on the null-space of A since this results in an unique solution of $x^*(\nu)$. The corresponding result is stated in the following proposition. *Proposition 15.* Assume that $f(x) = \frac{1}{2}x^T Hx + \zeta^T x$ with $H \in \mathbb{S}^n_+$ and $\zeta \in \mathbb{R}^n$, and that $h = I_{Ax=b}$. Further assume $x^T Hx > 0$ whenever $x \neq 0$ and Ax = 0, i.e. that H is positive definite on the null-space of A. Then (26) holds for all $\mathbf{L} \in \mathbb{S}^{m+p}_+$ such that $\mathbf{L} \succeq CK_{11}C^T$ where

$$\begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} = \begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix}^{-1}.$$
 (29)

Further, for any matrix $\mathbf{L} \succeq CK_{11}C^T$ there exist $\nu_1, \nu_2 \in \mathbb{R}^{m+p}$ such that (26) does not hold.

Proof. Since *H* is positive definite on the null-space of *A*, the KKT-matrix in (28) is invertible and $\begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}$ exists, see (Boyd and Vandenberghe, 2004, p. 523). Equation (29) implies that the solution the the KKT-system (28) is given by

$$\begin{bmatrix} x^{\star}(\nu) \\ \lambda^{\star}(\nu) \end{bmatrix} = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \begin{bmatrix} -C^T \nu - \zeta \\ b \end{bmatrix}.$$

That is, $x^{\star}(\nu) = -K_{11}(C^T\nu + \zeta) + K_{12}b$. Inserting this into (27) gives

$$d(\nu) = -\frac{1}{2}\nu^{T}C(2K_{11} - K_{11}HK_{11})C^{T}\nu + \xi^{T}\nu + \theta$$

= $-\frac{1}{2}\nu^{T}CK_{11}C^{T}\nu + \xi^{T}\nu + \theta$

where again $\xi \in \mathbb{R}^{m+p}$ and $\theta \in \mathbb{R}$ collect the linear and constant terms, and where $K_{11}HK_{11} = K_{11}$ is used in the second equality. This identity follows from the upper left block of $\begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}$ and using $K_{11}^T A = K_{11}A = AK_{11} = 0$, where $AK_{11} = 0$ follows from the lower left block of $\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$. This implies that d is a concave and quadratic function with negative Hessian $CK_{11}C^T$, which implies that (26) holds with equality for any $\nu_1, \nu_2 \in \mathbb{R}^{m+p}$ if $\mathbf{L} = CK_{11}C^T$. This further implies, that for any $\mathbf{L} \not\succeq CK_{11}C^T$ there exist $\nu_1, \nu_2 \in \mathbb{R}^{m+p}$ such that (26) does not hold. This concludes the proof.

Remark 16. In the model predictive control context, the preceding result implies that the quadratic cost matrix associated with inputs should be positive definite, while the quadratic cost matrix associated with the states need only be positive semi-definite.

5. FAST DUAL GRADIENT METHODS

In this section, we will describe generalized fast gradient methods and show how they can be applied to solve the dual problem (15). Generalized fast gradient methods can be applied to solve problems of the form

minimize
$$\ell(x) + \psi(x)$$
 (30)

where $x \in \mathbb{R}^n, \psi$: $\mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is proper, closed and convex, ℓ : $\mathbb{R}^n \to \mathbb{R}$ is convex, differentiable, and satisfies

$$\ell(x_1) \le \ell(x_2) + \langle \nabla \ell(x_2), x_1 - x_2 \rangle + \frac{1}{2} \|x_1 - x_2\|_{\mathbf{L}}^2 \quad (31)$$

for all $x_1, x_2 \in \mathbb{R}^n$ and some $\mathbf{L} \in \mathbb{S}^n_{++}$. Before we state the algorithm, we define the generalized prox operator

$$\operatorname{prox}_{\psi}^{\mathbf{L}}(x) := \arg\min_{u} \left\{ \psi(y) + \frac{1}{2} \|y - x\|_{\mathbf{L}}^2 \right\}$$
(32)

and note that

$$\operatorname{prox}_{\psi}^{\mathbf{L}}\left(x-\mathbf{L}^{-1}\nabla\ell(x)\right) =$$

$$= \arg\min_{y}\left\{\frac{1}{2}\|y-x+\mathbf{L}^{-1}\nabla\ell(x)\|_{\mathbf{L}}^{2}+\psi(y)\right\}$$

$$= \arg\min_{y}\left\{\ell(x)+\langle\nabla\ell(x),y-x\rangle+\frac{1}{2}\|y-x\|_{\mathbf{L}}^{2}+\psi(y)\right\}.$$
(33)

The generalized fast gradient method is stated below.

Algorithm 1. Generalized fast gradient method

Set: $y^1 = x^0 \in \mathbb{R}^n, t^1 = 1$ For $k \ge 1$ $x^k = \operatorname{pros}_{\psi}^{\mathbf{L}} \left(y^k - \mathbf{L}^{-1} \nabla \ell(y^k) \right)$ $t^{k+1} = \frac{1 + \sqrt{1 + 4(t^k)^2}}{2}$ $y^{k+1} = x^k + \left(\frac{t^{k-1}}{t^{k+1}} \right) (x^k - x^{k-1})$

The standard fast gradient method as presented in Beck and Teboulle (2009) is obtained by setting $\mathbf{L} = LI$ in Algorithm 1, where L is the Lipschitz constant to $\nabla \ell$. The main step of the fast gradient method is to perform a prox-step, i.e., to minimize (33) which can be seen as an approximation of the function $\ell + \psi$. For the standard fast gradient method, ℓ is approximated with a quadratic upper bound that has the same curvature, described by L, in all directions. If this quadratic upper bound is a bad approximation of the function to be minimized, slow convergence rate properties are expected. The generalization to allow for a matrix **L** in the algorithm allows for quadratic upper bounds with different curvature in different directions. This enables for quadratic upper bounds that much better approximate the function ℓ and consequently gives improved convergence rate properties.

The generalized fast gradient method has a convergence rate of (see Zuo and Lin (2011))

$$\ell_{\psi}(x^{k}) - \ell_{\psi}(x^{\star}) \le \frac{2\|x^{\star} - x^{0}\|_{\mathbf{L}}^{2}}{(k+1)^{2}}$$
(34)

where $\ell_{\psi} := \ell + \psi$. The convergence rate of the standard fast gradient method as given in Beck and Teboulle (2009), is obtained by setting $\mathbf{L} = LI$ in (34).

The objective here is to apply the generalized fast gradient method to solve the dual problem (15). By introducing $\tilde{g}(\nu) = g^*([0 \ I]\nu)$, the dual problem (15) can be expressed $\max_{\nu} d(\nu) - \tilde{g}(\nu)$, where d is defined in (16). As shown in Theorem 12, the function -d satisfies the properties required to apply generalized fast gradient methods. Namely that (31) holds for any $\mathbf{L} \in \mathbb{S}^{m+p}_+$ such that $\mathbf{L} \succeq CH^{-1}C^T$. Further, since g is a closed, proper, and convex function so is g^* , see (Rockafellar, 1970, Theorem 12.2), and by (Rockafellar, 1970, Theorem 5.7) so is \tilde{g} . This implies that generalized fast gradient methods, i.e. Algorithm 1, can be used to solve the dual problem (15). We set $-d = \ell$ and $\tilde{g} = \psi$, and restrict $\mathbf{L} = \text{blkdiag}(\mathbf{L}_{\lambda}, \mathbf{L}_{\mu})$ to get the following algorithm.

$$\begin{split} & \text{Set: } z^1 = \lambda^0 \in \mathbb{R}^m, v^1 = \mu^0 \in \mathbb{R}^p, t^1 = 1 \\ & \text{For } k \ge 1 \\ & y^k = \arg\min_x \left\{ f(x) + h(x) + (z^k)^T A x + (v^k)^T B x \right\} \\ & \lambda^k = z^k + \mathbf{L}_{\lambda}^{-1} (A y^k - b) \\ & \mu^k = \operatorname{prox}_{g^\star}^{\mathbf{L}_{\mu}} (v^k + \mathbf{L}_{\mu}^{-1} B y^k) \\ & t^{k+1} = \frac{1 + \sqrt{1 + 4(t^k)^2}}{2} \\ & z^{k+1} = \lambda^k + \left(\frac{t^k - 1}{t^{k+1}} \right) (\lambda^k - \lambda^{k-1}) \\ & v^{k+1} = \mu^k + \left(\frac{t^k - 1}{t^{k+1}} \right) (\mu^k - \mu^{k-1}) \end{split}$$

where y^k is the primal variable at iteration k that is used to help compute the gradient $\nabla d(\nu^k)$ where $\nu^k = (z^k, v^k)$. To arrive at the λ^k and μ^k iterations, we let $\xi^k = (\lambda^k, \mu^k)$, and note that

$$\begin{aligned} \xi^{k} &= \operatorname{prox}_{\widetilde{g}}^{\mathbf{L}} \left(\nu^{k} + \mathbf{L}^{-1} \nabla d(\nu^{k}) \right) \tag{35} \\ &= \arg\min_{\nu} \left\{ \frac{1}{2} \| \nu - \nu^{k} - \mathbf{L}^{-1} \nabla d(\nu^{k}) \|_{\mathbf{L}}^{2} + g^{\star}([0 \ I] \nu) \right\} \\ &= \left[\arg\min_{z} \left\{ \frac{1}{2} \| z - z^{k} - \mathbf{L}_{\lambda}^{-1} \nabla_{z} d(\nu^{k}) \|_{\mathbf{L}_{\lambda}}^{2} \right\} \\ \arg\min_{v} \left\{ \frac{1}{2} \| v - v^{k} - \mathbf{L}_{\mu}^{-1} \nabla_{v} d(\nu^{k}) \|_{\mathbf{L}_{\mu}}^{2} + g^{\star}(v) \right\} \right] \\ &= \left[z^{k} + \mathbf{L}_{\lambda}^{-1} (Ay^{k} - b) \\ \operatorname{prox}_{g^{\star}}^{\mu} (v^{k} + \mathbf{L}_{\mu}^{-1} By^{k}) \right]. \end{aligned}$$

In the following proposition we state the convergence rate properties of Algorithm 2.

Proposition 17. Suppose that Assumption 9 holds. If $\mathbf{L} = \text{blkdiag}(\mathbf{L}_{\lambda}, \mathbf{L}_{\mu}) \in \mathbb{S}^{m+p}_{++}$ is chosen such that $\mathbf{L} \succeq CH^{-1}C^{T}$. Then Algorithm 2 converges with the rate

$$D(\nu^{\star}) - D(\nu^{k}) \le \frac{2 \left\| \nu^{\star} - \nu^{0} \right\|_{\mathbf{L}}^{2}}{(k+1)^{2}}, \forall k \ge 1$$
(36)

where $D = d - \tilde{g}$ and k is the iteration number.

Proof. Algorithm 2 is Algorithm 1 applied to solve the dual problem (15). The convergence rate of Algorithm 1 is given by (34) provided that the function to be minimized a sum of one convex, differentiable function that satisfies (31) and one closed, proper, and convex function, see Zuo and Lin (2011). The discussion preceding the presentation of Algorithm 2 shows that the dual function to be optimized satisfies these properties for any $\mathbf{L} \in \mathbb{S}^{m+p}_{++}$ that satisfies $\mathbf{L} \succeq CH^{-1}C^T$. This concludes the proof.

Remark 18. If $h = I_{Ax=b}$, the requirement on **L** in Proposition 17 changes according to the results presented in Proposition 14 and Proposition 15.

Remark 19. By forming a specific running average of previous primal variables, it is possible to prove a O(1/k)convergence rate for the distance to the primal variable optimum and a $O(1/k^2)$ convergence rate for the worst case primal infeasibility, see Patrinos and Bemporad (2014).

For some choices of conjugate functions g^* , $\operatorname{prox}_{g^*}^{L_{\mu}}(x)$ can be difficult to evaluate. For standard prox operators (given by $\operatorname{prox}_{g^*}^I(x)$), Moreau decomposition (Rockafellar, 1970, Theorem 31.5) states that

$$\operatorname{prox}_{a^{\star}}^{I}(x) + \operatorname{prox}_{a}^{I}(x) = x.$$

In the following proposition, we will generalize this result to hold for the generalized prox-operator used here.

Proposition 20. Assume that $g : \mathbb{R}^n \to \mathbb{R}$ is a proper, closed, and convex function. Then

$$\operatorname{prox}_{g^{\star}}^{\mathbf{L}}(x) + \mathbf{L}^{-1} \operatorname{prox}_{g}^{\mathbf{L}^{-1}}(\mathbf{L}x) = x$$

every $x \in \mathbb{R}^{n}$ and any $\mathbf{L} \in \mathbb{S}_{+}^{n}$.

for

Proof. Optimality conditions for the prox operator (32) give that $y = \operatorname{prox}_{g^{\star}}^{\mathbf{L}}(x)$ if and only if

$$0 \in \partial g^{\star}(y) + \mathbf{L}(y - x)$$

Introducing $v = \mathbf{L}(x - y)$ gives $v \in \partial g^{\star}(y)$ which is equivalent to $y \in \partial g(v)$ (Rockafellar, 1970, Corollary 23.5.1). Since $y = x - \mathbf{L}^{-1}v$ we have

$$0 \in \partial g(v) + (\mathbf{L}^{-1}v - x)$$

which is the optimality condition for $v = \operatorname{prox}_{g}^{\mathbf{L}^{-1}}(\mathbf{L}x)$. This concludes the proof. Remark 21. If $g = I_{\mathcal{X}}$ where $I_{\mathcal{X}}$ is the indicator function, then g^* is the support function. Evaluating the prox operator (32) with g^* being a support function is difficult. However, through Proposition 20, this can be rewritten to only require the a projection operation onto the set \mathcal{X} . If \mathcal{X} is a box constraint and **L** is diagonal, then the projection becomes a max-operation and hence very cheap to implement.

Remark 22. We are not restricted to have one auxiliary term g only. We can have any number of auxiliary terms g_i that all decompose according to the computations in (35), i.e., we get one prox-operation in the algorithm for every auxiliary term g_i .

6. CHOOSING THE L-MATRIX

From Theorem 12 and Proposition 15, we get that the Lmatrix used in the quadratic lower bound in the algorithm should satisfy $\mathbf{L} \succeq CPC^T$, where $P = H^{-1}$ or $P = K_{11}$ depending on if the assumptions in Theorem 12 or Proposition 15 are satisfied. To get as fast convergence as possible, the approximation of the function d used in the algorithm should as accurately as possible resemble the function d itself. In view of Theorem 12 and Proposition 15, we want L to be a close as possible to CPC^T . Letting $\mathbf{L} = (D^T D)^{-1}$, we propose to achieve this by minimizing the condition number of $DCPC^TD^T$, subject to $I \succeq DCPC^TD^T$. If there are no structural constraints on \mathbf{L} and if CPC^T has full rank, then minimizing the condition number of $DCPC^TD^T$ gives $\mathbf{L} = (D^TD)^{-1} = CPC^T$. However, this situation is quite uncommon. First, we often have structural constraints on L that need to be taken into account. The most common such structural constraint is diagonal \mathbf{L} , since for separable g, the complexity of computing $\operatorname{prox}_{g}^{\mathbf{L}}(x)$ is not increased compared to using $\mathbf{L} = LI$. Sometimes, block-diagonal \mathbf{L} can be used, or in rare cases, full matrices L. All these structural constraints - diagonal, block-diagonal, and full - can be represented as follows: let \mathbb{L} be a set of pairs (i, j) for which \mathbf{L}_{ij} may be non-zero, then

$$\mathcal{L} = \{ \mathbf{L} \in \mathbb{S}^{m+p}_{++} \mid \mathbf{L} = (D^T D)^{-1}, \\ D \in \mathbb{R}^{(m+p) \times (m+p)} \text{ invertible,} \}$$

$$\mathbf{L}_{ii} = [\mathbf{L}^{-1}]_{ii} = [D^{-1}]_{ii} = 0 \text{ if } (i, j) \notin \mathbb{L}$$

For instance, letting $\mathbb{L} = \{(1, 1), (2, 2), \dots, (m + p, m + p)\}$ restricts $\mathbf{L} \in \mathcal{L}$ to be diagonal. A second issue that hinders the choice of $\mathbf{L} = CPC^T$, is that \mathbf{L} is restricted to be positive definite, while CPC^T is positive definite only if C has full row rank and if P is positive definite. When CPC^T is not positive definite, we instead propose to minimize the ratio between the largest and smallest nonzero eigenvalues (since the eigenvalues that are zero cannot be changed). Letting $\lambda_1(DCPC^TD^T)$ be the largest nonzero eigenvalue of $DCPC^TD^T$ and $\lambda_r(DCPC^TD^T)$ be the smallest non-zero eigenvalue of $DCPC^TD^T$ (where if r = m + p all eigenvalues are non-zero), the proposed optimization problems can be written as

$$D = \arg\min_{(D^T D)^{-1} \in \mathcal{L}} \frac{\lambda_1 (D C P C^T D^T)}{\lambda_r (D C P C^T D^T)}.$$
 (37)

Next we will show how to solve (37) in the following three cases, which include all problem instances we will encounter:

 $\begin{array}{ll} (\text{C1}) & CPC^T \in \mathbb{S}^{(m+p)}_{++} \\ (\text{C2}) & QC^TCQ^T \in \mathbb{S}^q_{++} \text{ where } P = Q^TQ \text{ and } Q \in \mathbb{R}^{q \times n} \\ (\text{C3}) & \operatorname{rank}(QC^TCQ^T) = \operatorname{rank}(CPC^T) < \min(m+p,q) \end{array}$

Before we present how to compute the optimal preconditioner in each of the three cases, we state the following lemma.

Lemma 23. For any matrix $A \in \mathbb{R}^{m \times n}$, the non-zero eigenvalues of $A^T A$ equals the non-zero eigenvalues of AA^T .

Proof. Without loss of generality, we assume that $m \leq n$ and that rank $(A) = q \leq m$. Let $A = U\Sigma V^T$, be the singular value decomposition of A, where $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ are orthonormal, and

$$\Sigma = \begin{bmatrix} s_1 & & \\ & \ddots & \\ & & s_q \\ & & & 0 \end{bmatrix} \quad 0 \end{bmatrix} \in \mathbb{R}^{m \times n}.$$

This implies that $AA^T = U\Sigma V^T V\Sigma^T U^T = U(\Sigma\Sigma^T)U^T$, or equivalently that $(AA^T)U = U(\Sigma\Sigma^T)$, and that $A^TA = V\Sigma^T U^T U\Sigma V^T = V\Sigma^T \Sigma V^T$, or equivalently that $(A^TA)V = V(\Sigma^T\Sigma)$. That is, the eigenvalues to AA^T are given by the diagonal entries of $\Sigma\Sigma^T$ and the eigenvalues to A^TA are given by the diagonal entries of $\Sigma^T\Sigma$, i.e. the nonzero eigenvalues of AA^T and A^TA coincide. This concludes the proof.

6.1 Case 1

We consider Case 1, i.e. C1. This is the case considered in Theorem 12 with $P = H^{-1}$ and an additional rank assumption on C.

Proposition 24. Assume that $CPC^T \in \mathbb{S}^{(m+p)}_{++}$. Then a matrix D with $(D^TD)^{-1} \in \mathcal{L}$ that minimizes the ratio (37) can be computed by solving the semi-definite program

minimize
$$t$$

subject to $tCPC^T \succeq \mathbf{L}$
 $CPC^T \preceq \mathbf{L}$
 $\mathbf{L} \in \mathcal{L}$ (38)

where $\mathbf{L} = (D^T D)^{-1}$. Further, $\mathbf{L} \succeq CPC^T$.

Proof. Since CPC^T has full rank, (37) is the condition number. Thus, according to (Boyd et al., 1994, Section 3.1), (38) can be solved in order to minimize (37). Further, the second constraint implies that $\mathbf{L} \succeq CPC^T$.

6.2 Case 2

Here, we show how to minimize (37) in the second case, C2. This covers both Theorem 12 (with $P = H^{-1}$) and Proposition 15 (with $P = K_{11}$) with the additional assumption that C is not wide and has full column rank. *Proposition 25.* Assume that $QC^TCQ^T \in \mathbb{S}^q_{++}$, where $P \in \mathbb{S}^n_+$ is factorized as $P = Q^TQ$, where $Q \in \mathbb{R}^{q \times n}$ has rank q. Then a matrix D with $(D^TD)^{-1} \in \mathcal{L}$ that minimizes the ratio (37) can be computed by solving the semi-definite program

$$\begin{array}{ll} \text{minimize} & -t \\ \text{subject to} & QC^T M C Q^T \leq I \\ & QC^T M C Q^T \geq t I \\ & M \in \mathcal{L} \end{array}$$
(39)

where $M = (D^T D)$. Further $\mathbf{L} = (D^T D)^{-1} \succeq CPC^T$.

Proof. Since QC^TMCQ^T has full rank, we get from Lemma 23, we get that minimizing the condition number of QC^TMCQ^T is equivalent to minimizing the ratio between the largest and smallest non-zero eigenvalues of $DCPC^TD^T$, i.e. equivalent to solving (37). From (Boyd et al., 1994, Section 3.1), we get that (39) minimizes the condition number of QC^TMCQ^T i.e. it minimizes (37). Further, the first inequality implies through Lemma 23 that $DCPC^TD^TQ \preceq I$, which is equivalent to that $\mathbf{L} = (D^TD)^{-1} \succeq CPC^T$. This concludes the proof.

6.3 Case 3

Here, we consider Case C3, which covers the cases not included in Cases C1 and C2. This covers, e.g. the situation in Proposition 15 with additional assumptions on the rank of C.

Proposition 26. Assume that rank $(QC^TCQ^T) = r$ with $r < \min(m + p, q)$ and that $P \in \mathbb{S}^n_+$ is factorized as $P = Q^TQ$, where $Q \in \mathbb{R}^{q \times n}$ has rank q. Further, assume that $\Phi \in \mathbb{R}^{q \times r}$ is an orthonormal basis for $\mathcal{R}(QC^T)$. Then a matrix D with $(D^TD)^{-1} \in \mathcal{L}$ that minimizes the ratio (37) can be computed by solving the semi-definite program

$$\begin{array}{ll} \text{minimize} & -t \\ \text{subject to} & QC^T M C Q^T \preceq I \\ & \Phi^T Q C^T M C Q^T \Phi \succeq t I \\ & M \in \mathcal{L} \end{array}$$
(40)

where $M = (D^T D)$. Further, $\mathbf{L} = (D^T D)^{-1} \succeq CPC^T$.

Proof. The first inequality in (40) is by Lemma 23 equivalent to that $DCPC^TD^T \leq I$, i.e. $\lambda_1(DCPC^TD^T) \leq 1$.

To lower bound the smallest nonnegative eigenvalue, we need to search in directions perpendicular to the null-space of QC^TMCQ^T . We have that

$$\mathcal{N}(QC^T M C Q^T) = \mathcal{N}(DCQ^T) = \mathcal{N}(CQ^T) \perp \mathcal{R}(QC^T)$$

where the second equality holds since D is assumed invertible. This implies that we need to search in directions that span $\mathcal{R}(QC^T)$. Now, we have that $t \leq \lambda_r(QC^TMCQ^T)$ if and only if $0 \leq x^T(QC^TMCQ^T-tI)x$ for all $x \in \mathcal{R}(QC^T)$. This, in turn, is equivalent to that

$$\Phi^T (QC^T M C Q^T - tI) \Phi \in \mathbb{S}_+^r \tag{41}$$

where $\Phi \in \mathbb{R}^{q \times r}$ is an orthonormal basis to $\mathcal{R}(QC^T)$. Further, since Φ is an orthonormal basis, i.e. $\Phi^T \Phi = I$, (41) is equivalent to $\Phi^T Q C^T M C Q^T \Phi \succeq tI$. This chain of equivalences shows that the second inequality in (40) is equivalent to that $\lambda_r(QC^T M C Q^T) \geq t$. Thus, by maximizing t (or equivalently minimizing -t) the ratio

$$\lambda_1(QC^T M C Q^T) / \lambda_r(QC^T M C Q^T) \leq 1/t$$

is minimized. From Lemma 23 and the reasoning to the proof of Case C2, we conclude that (40) solves (37).

Further, the first inequality implies through Lemma 23 that $\mathbf{L} = (D^T D)^{-1} \succeq CPC^T$. This concludes the proof.

Remark 27. Note that if rank $(QC^TCQ^T) = q$, then $\Phi = I$ is an orthonormal basis to $\mathcal{R}(QC^T)$ and (40) reduces to (39). Thus, (40) is a generalization of (39) to cover also the positive semi-definite case. A similar generalization that reduces to (38) in the positive definite case would rely on searching in directions perpendicular to $\mathcal{N}(DCPC^TD^T) = \mathcal{N}(QC^TD^T) \perp \mathcal{R}(DCQ^T)$ to lower bound the smallest non-zero eigenvalue. This implies that the search directions depend on the decision variables D, which makes such a generalization more elaborate.

7. MODEL PREDICTIVE CONTROL

In this section, we pose some standard model predictive control problems and show how they can be solved using the methods presented in this paper. The resulting algorithms will have simple arithmetic operations only which allows for easier implementation in embedded systems. We also show how to choose the **L**-matrix in each case.

Example 28. We consider MPC optimization problems of the form

minimize
$$\sum_{t=0}^{N-1} \frac{1}{2} \left(x_t^T Q x_t + u_t^T R u_t \right) + \frac{1}{2} x_N^T Q_f x_N$$
subject to
$$x_{t+1} = \Phi x_t + \Gamma u_t, \quad t = 0, \dots, N-1$$
$$x_{\min} \le x_t \le x_{\max}, \quad t = 0, \dots, N$$
$$u_{\min} \le u_t \le u_{\max}, \quad t = 0, \dots, N-1$$
$$x_0 = \bar{x}$$

where $\bar{x}, x_t \in \mathbb{R}^{n_x}, u_t \in \mathbb{R}^{n_u}, \Phi \in \mathbb{R}^{n_x \times n_x}, \Gamma \in \mathbb{R}^{n_x \times n_u}$ and $Q \in \mathbb{S}^{n_x}_{++}, R \in \mathbb{S}^{n_u}_{++}, Q_f \in \mathbb{S}^{n_x}_{++}$ are all diagonal. Letting $y = (x_0, \ldots, x_N, u_0, \ldots, u_{N-1})$, this can be cast as

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}y^T H y\\ \text{subject to} & Ay = b\bar{x}\\ & y_{\min} \leq y \leq y_{\max} \end{array}$$

where H, A, b, y_{\min} , and y_{\max} are structured according to y. We choose $f(y) = \frac{1}{2}y^T Hy$, g = 0, and $h = I_{\mathcal{Y}}$ where $I_{\mathcal{Y}}$ is the indicator function to

$$\mathcal{Y} = \{ y \in \mathbb{R}^{(N+1)n_x + Nn_u} \mid y_{\min} \le y \le y_{\max} \}.$$

This implicitly implies that we introduce dual variables λ for the equality constraints $Ay = b\bar{x}$. The algorithm becomes:

$$y^{k} = \arg\min_{y} \left\{ \frac{1}{2} y^{T} H y + I_{\mathcal{Y}}(y) + z^{T} A x \right\}$$
(42)

$$\lambda^k = z^k + \mathbf{L}_{\lambda}^{-1} (Ay^k - b\bar{x}) \tag{43}$$

$$t^{k+1} = \frac{1+\sqrt{1+4(t^k)^2}}{2} \tag{44}$$

$$z^{k+1} = \lambda^k + \left(\frac{t^k - 1}{t^{k+1}}\right) \left(\lambda^k - \lambda^{k-1}\right) \tag{45}$$

where the first step (42) can be implemented as

$$y^{k} = \max\left(\min\left(-H^{-1}A^{T}z^{k}, y_{\max}\right), y_{\min}\right)$$
(46)

due to the structure of the problem. The preceding section suggests that $\mathbf{L}_{\lambda} = (D^T D)^{-1} \succeq A H^{-1} A^T$ should be chosen such that $I \approx D A H^{-1} A^T D^T$. Since A is sparse and H^{-1} is diagonal due to the MPC problem formulation, Dcan be chosen to get equality in $I \approx D A H^{-1} A^T D^T$, i.e. we can choose $\mathbf{L}_{\lambda} = (D^T D)^{-1} = A H^{-1} A^T$. The algorithm requires the computation of $\mathbf{L}_{\lambda}^{-1} z$, where $z = A y^k - b \bar{x}$, in each iteration. Since $\mathbf{L}_{\lambda} = A H^{-1} A$ is sparse, this can efficiently be implemented by offline storing the sparse Cholesky factorization $R^T R = S^T \mathbf{L}_{\lambda} S$, where R is sparse and upper triangular, and S is a permutation matrix. The online computation of $\mathbf{L}_{\lambda}^{-1}z$ then reduces to one forward and one backward solve, which can be very efficiently implemented.

The algorithm in this example is a generalization of the algorithm in Richter et al. (2013), where the matrix **L** is chosen as $\mathbf{L} = \|AH^{-1}A^T\|_2 I$. In the numerical section we will see that this generalization can significantly improve the convergence rate.

Next, we present an algorithm that works for arbitrary positive definite cost matrices, and arbitrary linear constraints.

 $Example\ 29.$ We consider MPC optimization problems of the form

minimize
$$\sum_{t=0}^{N-1} \frac{1}{2} \left(x_t^T Q x_t + u_t^T R u_t \right) + \frac{1}{2} x_N^T Q_f x_N$$

subject to
$$x_{t+1} = \Phi x_t + \Gamma u_t, \qquad t = 0, \dots, N-1$$
$$\underline{d}_x \leq B_x x_t \leq \overline{d}_x, \qquad t = 0, \dots, N-1$$
$$\underline{d}_u \leq B_u u_t \leq \overline{d}_u, \qquad t = 0, \dots, N-1$$
$$x_0 = \overline{x}, \underline{d}_N \leq B_N x_N \leq \overline{d}_N$$

where $\bar{x}, x_t \in \mathbb{R}^{n_x}, u_t \in \mathbb{R}^{n_u}, \Phi \in \mathbb{R}^{n_x \times n_x}, \Gamma \in \mathbb{R}^{n_x \times n_u}, B_x \in \mathbb{R}^{p_x \times n_x}, B_u \in \mathbb{R}^{p_u \times n_x}, B_N \in \mathbb{R}^{p_N \times n_x}, \underline{d}_x, \bar{d}_x \in \mathbb{R}^{p_x}, \underline{d}_u, \bar{d}_u \in \mathbb{R}^{p_u}, \underline{d}_N, \bar{d}_N \in \mathbb{R}^{p_N}, Q \in \mathbb{S}^{n_x}_{++}, R \in \mathbb{S}^{n_u}_{++}, and Q_f \in \mathbb{S}^{n_x}_{++}$. We let $y = (x_0, \dots, x_N, u_0, \dots, u_{N-1})$ and define B = blkdiag $(\bar{B}_x, B_N, \bar{B}_u)$ where \bar{B}_x = blkdiag (B_x, \dots, B_x) and \bar{B}_u = blkdiag (B_u, \dots, B_u) . We also introduce $\underline{d} = (\underline{d}_x, \dots, \underline{d}_x, \underline{d}_N, \underline{d}_u, \dots, \underline{d}_u)$ and $\overline{d} = (\bar{d}_x, \dots, \bar{d}_x, \bar{d}_N, \bar{d}_u, \dots, d_u)$. This implies that all inequality constraints are described by $\underline{d} \leq By \leq \overline{d}$. Using this notation, the optimization problem can be rewritten as

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}y^T H y\\ \text{subject to} & Ay = b\bar{x}\\ & By = v\\ & \underline{d} \leq v \leq \bar{d} \end{array}$$

We let $f(y) = \frac{1}{2}y^T Hy$, $h = I_{Ay=b\bar{x}}$, and $g = I_{\mathcal{Y}}$ where $\mathcal{Y} = \{y \in \mathbb{R}^{(N+1)n_x+Nn_u} \mid \underline{d} \leq y \leq \overline{d}\}$. Since h is the indicator function for the equality constraints $Ay = b\bar{x}$, we do not need to introduce dual variables for those constraints. However, we introduce dual variables μ for By = v. Letting $H_A = AH^{-1}A^T$, the algorithm becomes $u^k = H^{-1}(A^T H_{-1}^{-1}(AH^{-1}B^T v^k + b\bar{x}) - B^T v^k)$ (47)

$$\mu^{k} = \operatorname{prox}_{a^{\star}}^{\mathbf{L}_{\mu}}(v^{k} + \mathbf{L}_{\mu}^{-1}By^{k})$$
(48)

$$t^{k+1} = \frac{1+\sqrt{1+4(t^k)^2}}{2} \tag{49}$$

$$v^{k+1} = \mu^k + \left(\frac{t^{k-1}}{t^{k+1}}\right) (\mu^k - \mu^{k-1})$$
(50)

where the y^k iterate follows from solving $\min_x \{f(x) + I_{Ax=b\bar{x}}(x) + (v^k)^T Bx\}$. In an implementation, the y^k -update can be implemented as in (47). Then, for efficiency, the matrix multiplications should be computed offline and stored for online use. Depending on the sparsity of H, A, and B, it might be more efficient to use the KKT-system from which (47) is deduced, namely

$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} y^k \\ \xi \end{bmatrix} = \begin{bmatrix} -B^T v^k \\ b\bar{x} \end{bmatrix}$$

Then, a sparse LDL-factorization of the KKT-matrix $\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix}$ is computed offline for online use. The online

computational burden to compute the y^k -update then becomes one forward and one backward solve. Whichever method that has the lower number of flops should be chosen.

By restricting \mathbf{L}_{μ} to be diagonal, the second step, i.e. (48), can be implemented as

$$\mu^{k} = \min(v^{k} + \mathbf{L}_{\mu}^{-1}(By^{k} - \underline{d}), \max(v^{k} + \mathbf{L}_{\mu}^{-1}(By^{k} - \overline{d}), 0))$$

To get fast convergence, the diagonal \mathbf{L}_{μ} should be computed as in Section 6. Note that, in this example, the matrix P used in Section 6 can be either $P = K_{11}$, where K_{11} is implicitly defined in (29), or $P = H^{-1}$. Since $K_{11} \leq H^{-1}$, the latter choice is expected to give a somewhat slower convergence.

The splitting method used here is the same as the one used in Patrinos and Bemporad (2014). However, this is more general since we allow for \mathbf{L}_{μ} -matrices that are not a multiple of the identity matrix. Also, the same splitting is used in O'Donoghue et al. (2013); Jerez et al. (2013), where ADMM (see Boyd et al. (2011)) is used to solve the optimization problem.

8. NUMERICAL EXAMPLE

The proposed algorithms are evaluated by applying them to the AFTI-16 aircraft model in Kapasouris et al. (1990); Bemporad et al. (1997). This problem is also a tutorial example in the MPC toolbox in MATLAB. As in Bemporad et al. (1997) and the MPC toolbox tutorial, the continuous time model from Kapasouris et al. (1990) is sampled using zero-order hold every 0.05 s. The system has four states $x = (x_1, x_2, x_3, x_4)$, two outputs $y = (y_1, y_2)$, two inputs $u = (u_1, u_2)$, and obeys the following dynamics

$$\begin{aligned} x^+ &= \begin{bmatrix} 0.999 & -3.008 & -0.113 & -1.608\\ -0.000 & 0.986 & 0.048 & 0.000\\ 0.000 & 2.083 & 1.009 & -0.000\\ 0.000 & 0.053 & 0.050 & 1.000 \end{bmatrix} x + \begin{bmatrix} -0.080 & -0.635\\ -0.029 & -0.014\\ -0.868 & -0.092\\ -0.022 & -0.002 \end{bmatrix} u, \\ y &= \begin{bmatrix} 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix} x \end{aligned}$$

where x^+ denotes the state in the next time step. The dynamics, input, and output matrices are denoted by Φ , Γ , C respectively, i.e. we have $x^+ = \Phi x + \Gamma u$, y = Cx. The system is unstable, the magnitude of the largest eigenvalue of the dynamics matrix is 1.313. The outputs are the attack and pitch angles, while the inputs are the elevator and flaperon angles. The inputs are physically constrained to satisfy $|u_i| \leq 25^\circ$, i = 1, 2. The outputs are soft constrained to satisfy $-s_1 - 0.5 \leq y_1 \leq 0.5 + s_2$ and $-s_3 - 100 \leq y_2 \leq 100 + s_4$ respectively, where $s = (s_1, s_2, s_3, s_4) \geq 0$ are slack variables. The cost in each time step is

$$\ell(x, u, s) = \frac{1}{2} ((x - x_r)^T Q(x - x_r) + u^T R u + s^T S s)$$

where $Q = C^T Q_y C + Q_x$, where $Q_y = 10^2 I$ and $Q_x = \text{diag}(10^{-4}, 0, 10^{-3}, 0)$, x_r is such that $y_r = Cx_r$ where y_r is the output reference that can vary in each step, $R = 10^{-2}I$, and $S = 10^6 I$. This gives condition number 10^{10} of the full cost matrix. Further, the terminal cost is Q, and the control and prediction horizon is N = 10. The numerical data in Tables 1 and 2 is obtained by following a reference trajectory on the output. The objective is to change the pitch angle from 0° to 10° and then back to 0° while the angle of attack satisfies the output constraints

 $-0.5^\circ \leq y_1 \leq 0.5^\circ.$ The constraints on the angle of attack limits the rate on how fast the pitch angle can be changed.

In Table 1, the proposed algorithms are evaluated by comparing them to other first order methods recently proposed in the literature for embedded model predictive control, namely Richter et al. (2013); Patrinos and Bemporad (2014); O'Donoghue et al. (2013); Jerez et al. (2013). In Table 2, the execution time of a C implementation of Algorithm 2 is compared to the execution time of FORCES, Domahidi et al. (2012), which is a C code generator for MPC-problems, and to the commercial solver MOSEK.

All algorithms in the comparison in Table 1 are implemented in MATLAB, while the algorithms in Table 2 are implemented in C. Further, all simulations are performed on a Linux machine using a single core running at 2.9 GHz. To create an easily transferable and fair termination criterion, the optimal solution to each optimization problem y^* is computed to high accuracy using an interior point solver. Where applicable, the optimality condition is $||y^k - y^*||_2/||y^*||_2 \leq 0.005$, where y^k is the primal iterate in the algorithm. This implies that a relative accuracy of 0.5% of the primal solution is required.

First, we discuss the results in Table 1. The algorithms in Example 28, i.e. (42)-(45), and Example 29, i.e. (47)-(50), have been applied to this problem. Due to the slack variables, (46) cannot replace (42) for the y^k update. However, the y^k minimization is separable in the constraints and each of the projections can be solved by a multiparametric program with two regions. This is almost as computationally inexpensive as the y^k update in (46). Further, we use $\mathbf{L}_{\lambda} = AH^{-1}A^T$. Algorithm (42)-(45) is a generalization of Richter et al. (2013) that allows for general matrices \mathbf{L}_{λ} . The algorithm in Richter et al. (2013) is obtained by setting $\mathbf{L}_{\lambda} = \|AH^{-1}A^T\|_2 I$. The numerical evaluation in Table 1 reveals that this generalization improves the execution time with more than three orders of magnitude for this problem. The formulation in Example 29, i.e. (47)-(50), directly covers this MPC formulation with soft constraints. For this algorithm, we compute \mathbf{L}_{μ} as in Section 6 using both $P = K_{11}$ and $P = H^{-1}$. The resulting algorithm is a generalization of the algorithm in Patrinos and Bemporad (2014). The algorithm in Patrinos and Bemporad (2014) is given by setting $\mathbf{L}_{\mu} = \|BH^{-1}B^T\|_2 I$ or $\mathbf{L}_{\mu} = \|BK_{11}B^T\|_2 I$ in the iterations (47)-(50). Table 1 indicates that this generalization improves the algorithm by one to two orders of magnitude compared to Patrinos and Bemporad (2014). Further, (47)-(50) is based on the same splitting as the method in O'Donoghue et al. (2013); Jerez et al. (2013). The difference is that here, the problem is solved with a generalized dual gradient method, while in O'Donoghue et al. (2013); Jerez et al. (2013) it is solved using ADMM. In ADMM, the ρ -parameter need to be chosen. However, no exact guidelines are yet known for this choice, and the performance of the algorithm often relies heavily on this parameter. We compare our algorithm with ADMM using the best ρ that we found, $\rho = 3$, and with one larger and one smaller ρ . Table 1 reports that the execution time for our method is one to two orders of magnitude smaller (or more if the ρ -parameter in O'Donoghue et al. (2013); Jerez et al. (2013) is chosen suboptimally) than the algorithm proposed in O'Donoghue et al. (2013); Jerez et al. (2013).

		exec ti	exec time (ms) nb		iters
Algorithm	Parameters	avg.	max	avg.	max
(42)-(45)	$\mathbf{L}_{\lambda} = AH^{-1}A^{T}$	2.3	12.1	21.7	102
Richter et al. (2013)	$\mathbf{L}_{\lambda} = \ AH^{-1}A^T\ _2 I$	4713.9	28411	50845	308210
(47)-(50)	\mathbf{L}_{μ} comp. as in Sec.6 w. $P = K_{11}$	1.4	7.1	23.5	128
(47)-(50)	\mathbf{L}_{μ} comp. as in Sec.6 w. $P = H^{-1}$	1.2	5.8	20.0	105
Patrinos and Bemporad (2014)	$\mathbf{L}_{\mu} = \ BK_{11}B^T\ _2 I$	98.5	673.0	1835.9	12686
Patrinos and Bemporad (2014)	$\mathbf{L}_{\mu} = \ BH^{-1}B^T\ _2 I$	98.9	679.4	1850.1	12783
O'Donoghue et al. (2013); Jerez et al. (2013)	$\rho = 0.3$	193.9	920.6	3129.5	15037
O'Donoghue et al. (2013); Jerez et al. (2013)	$\rho = 3$	29.7	142.2	457.3	2179
O'Donoghue et al. (2013) ; Jerez et al. (2013)	$\rho = 30$	35.1	264.4	556.7	4194

Table 1. Comparison to other first-order methods, all implemented in MATLAB.

Table2.Comparison to state-of-the-art
solvers, all implemented in C.

		exec time (ms)	
Algorithm	Parameters	avg.	max
(42)-(45)	$\mathbf{L}_{\lambda} = AH^{-1}A^{T}$	0.079	0.232
(47)-(50)	\mathbf{L}_{μ} as in Sec.6 w. $P = H^{-1}$	0.061	0.196
FORCES	-	0.347	0.592
MOSEK	-	4.9	5.4

In Table 2, we compare different solvers implemented in C. For the algorithms presented in this paper, we generate C code that take the reference trajectory and the initial state as inputs. Compared to the corresponding MATLAB implementations in Table 1, the generated C code is more than 20 times faster. These implementations are compared to FORCES and MOSEK. FORCES, see Domahidi et al. (2012), is an optimized interior point C code generator for MPC problems. The structure of the MPC problem is exploited to significantly reduce the computational time when solving the KKT-system in each iteration. The comparison also includes MOSEK, which is a general commercial QP-solver that does not have the advantage of generating code for this specific problem beforehand. The numerical evaluation in Table 2 shows that our algorithms and FORCES, for both of which C code is generated for this specific problem instance, outperform the general purpose commercial C solver MOSEK with more than one order of magnitude. Further, Table 2 reveals that our two algorithms perform similarly and that they are at least two to three times faster than FORCES.

9. CONCLUSIONS

We have proposed a generalization of dual fast gradient methods. This generalization allows the algorithm to, in each iteration, minimize a quadratic upper bound to the negative dual function with different curvature in different directions. This is in contrast to the standard fast dual gradient method where a quadratic upper bound to the negative dual with the same curvature in all directions is minimized in each iteration. This generalization is made possible by the main contribution of this paper that characterizes the set of matrices that can be used to describe a quadratic upper bound to the negative dual function. The numerical evaluation on an ill-conditioned aircraft problem reveals that the proposed algorithms outperform several other MPC problem solvers recently proposed in the literature.

REFERENCES

- A. Beck and M. Teboulle. A fast iterative shrinkagethresholding algorithm for linear inverse problems. SIAM J. Imaging Sciences, 2(1):183–202, October 2009.
- A. Bemporad, A. Casavola, and E. Mosca. Nonlinear control of constrained linear systems via predictive reference management. *IEEE Transactions on Automatic Control*, 42(3):340–349, 1997.
- A. Bemporad, M. Morari, V. Dua, and E.N. Pistikopoulos. The explicit linear quadratic regulator for constrained systems. *Automatica*, 38(1):3–20, January 2002.
- D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, MA, 2nd edition, 1999.
- S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, 2004.
- S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. Linear Matrix Inequalities in System and Control Theory, volume 15 of Studies in Applied Mathematics. SIAM, Philadelphia, PA, June 1994.
- S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations* and *Trends in Machine Learning*, 3(1):1–122, 2011.
- A. Domahidi, A. Zgraggen, M.N. Zeilinger, M. Morari, and C.N. Jones. Efficient interior point methods for multistage problems arising in receding horizon control. In *IEEE Conference on Decision and Control (CDC)*, pages 668–674, Maui, HI, USA, December 2012.
- P. Giselsson. Improving fast dual ascent for MPC Part I: The distributed case. *Automatica*, 2014a. Submitted. Available http://arxiv.org/abs/1312.3012.
- P. Giselsson. Improved fast dual gradient methods for embedded model predictive control. In *In Proceedings* of 2014 *IFAC World Congress*, 2014b. Accepted for publication. Available https://www.control.lth.se/ Staff/PontusGiselsson.html.
- J. L. Jerez, P. J. Goulart, S. Richter, G. A. Constantinides, E. C. Kerrigan, and M. Morari. Embedded online optimization for model predictive control at megahertz rates. *IEEE Transactions on Automatic Control*, 2013. Submitted.
- P. Kapasouris, M. Athans, and G. Stein. Design of feedback control systems for unstable plants with saturating actuators. In *Proceedings of the IFAC Symposium on Nonlinear Control System Design*, pages 302–307. Pergamon Press, 1990.
- Y. Nesterov. A method of solving a convex programming problem with convergence rate O (1/k²). Soviet Mathematics Doklady, 27(2):372–376, 1983.

- Y. Nesterov. Introductory Lectures on Convex Optimization: A Basic Course. Springer Netherlands, 1st edition, 2003. ISBN 1402075537.
- Y. Nesterov. Smooth minimization of non-smooth functions. Math. Program., 103(1):127–152, May 2005.
- B. O'Donoghue, G. Stathopoulos, and S. Boyd. A splitting method for optimal control. *IEEE Transactions on Control Systems Technology*, 21(6):2432–2442, 2013.
- P. Patrinos and A. Bemporad. An accelerated dual gradient-projection algorithm for embedded linear model predictive control. *IEEE Transactions on Automatic Control*, 59(1):18–33, 2014.
- S. Richter, C. N. Jones, and M. Morari. Certification aspects of the fast gradient method for solving the dual of parametric convex programs. *Mathematical Methods* of Operations Research, 77(3):305–321, 2013.
- K. T. Rockafellar. *Convex Analysis*, volume 28. Princeton University Press, Princeton, NJ, 1970.
- P. Tseng. On accelerated proximal gradient methods for convex-concave optimization. Technical report. Available: http://www.csie.ntu.edu.tw/~b97058/tseng/ papers/apgm.pdf, May 2008.
- W. Zuo and Z. Lin. A generalized accelerated proximal gradient approach for total-variation-based image restoration. *IEEE Transactions on Image Processing*, 20(10):2748–2759, October 2011.