

Diagonal Scaling in Douglas-Rachford Splitting and ADMM

Pontus Giselsson* and Stephen Boyd*

Abstract—The performance of Douglas-Rachford splitting and the alternating direction method of multipliers (ADMM) (i.e. Douglas-Rachford splitting on the dual problem) are sensitive to conditioning of the problem data. For a restricted class of problems that enjoy a linear rate of convergence, we show in this paper how to precondition the optimization data to optimize a bound on that rate. We also generalize the preconditioning methods to problems that do not satisfy all assumptions needed to guarantee a linear convergence. The efficiency of the proposed preconditioning is confirmed in a numerical example, where improvements of more than one order of magnitude are observed compared to when no preconditioning is used.

I. INTRODUCTION

Optimization problems of the form

$$\begin{aligned} & \text{minimize} && f(x) + g(y) \\ & \text{subject to} && Ax = y \end{aligned} \quad (1)$$

where $x \in \mathbf{R}^n$ is the variable and f and g are convex, arise in numerous applications ranging from compressed sensing [6] and statistical estimation [17] to model predictive control [24] and image restoration. There exist a variety of algorithms for solving convex problems of the form (1), many of which are treated in the monograph [22]. The methods include primal and dual forward-backward splitting methods [8] and their accelerated variants [3], the Arrow-Hurwicz method [1], Douglas Rachford splitting [10] and Peaceman-Rachford splitting [23], the alternating direction method of multipliers (ADMM) [16], [13], [5] (which is Douglas-Rachford splitting applied to the dual problem [12], [11]), and linearized ADMM [7].

In this paper, we focus on Douglas-Rachford splitting, Peaceman-Rachford splitting and ADMM. These methods are well known to converge sub-linearly under very general assumptions, see e.g. [11]. However, the convergence time can vary greatly depending on conditioning of the problem data and on the algorithm parameters. Yet, very little is known on how to precondition the data and how to select algorithm parameters to achieve a well performing algorithm in the general case. In the context of finding a zero of two maximal monotone operators, it is in [19] shown that when one of the operators is Lipschitz continuous and strongly monotone, the Douglas-Rachford algorithm converges with a linear rate. Also, it is shown how to select the algorithm parameter to optimize the bound on this linear rate. In the case of applying Douglas-Rachford splitting to the dual of (1) (i.e. applying ADMM to (1)) with f or g strongly convex and

smooth, the bound on the linear rate from [19] is improved in [9], and another parameter selection is provided.

For a more restricted class of problems, another improvement of the linear convergence rate bound for ADMM is provided in [14]. They consider problems where f is quadratic and strongly convex, i.e. $f(x) = \frac{1}{2}x^T Hx + \xi^T x$ with H positive definite, and g is the indicator function for the constraint set $\mathcal{Y} = \{y \in \mathbf{R}^m \mid y \leq b\}$. They show how to choose the algorithm parameter and how to individually scale the equality constraints $Ax = y$ such that the linear convergence rate factor in the bound is optimized.

In this paper, we further improve on the linear convergence rate bound for the Douglas-Rachford splitting from [19], and for Peaceman-Rachford splitting. We also show that Peaceman-Rachford has a faster theoretical linear rate than Douglas-Rachford, in the case of f being strongly convex and smooth. These results are translated to the ADMM case by applying Douglas-Rachford splitting and Peaceman-Rachford splitting to the dual problem. The tight strong convexity and smoothness characterizations of the dual problem in [15], are in this paper utilized to improve the bound on the convergence rate of ADMM compared to [9]. Our convergence results also generalize the results in [14] to handle any smooth and strongly convex f and any proper, closed, and convex g .

The obtained bounds on the linear rates for Douglas-Rachford splitting, Peaceman-Rachford splitting and ADMM depend on the problem scaling and on algorithm parameters. We show how to select these to optimize the respective bounds. We also propose scaling and parameter selection procedures for problems where some of the assumptions needed to achieve a linear convergence rate are not met. These extensions significantly enlarge the class of problems for which the scaling and parameter selection procedures can be successfully applied. A numerical example is provided that shows the efficiency of the proposed scaling. For the considered problem, the execution time is decreased with about one order of magnitude compared to when no scaling is used.

II. PRELIMINARIES AND NOTATION

A. Notation

We denote by \mathbf{R} the set of real numbers, \mathbf{R}^n the set of real column-vectors of length n , and $\mathbf{R}^{m \times n}$ the set of real matrices with m rows and n columns. Further $\bar{\mathbf{R}} := \mathbf{R} \cup \{\infty\}$ denotes the extended real line. We use $\langle \cdot, \cdot \rangle$ as the inner product in the Euclidean space, i.e. $\langle x, y \rangle = x^T y$. Further $\|\cdot\|$ denotes the standard Euclidean norm, while $\|x\|_M =$

* Electrical Engineering Department, Stanford University. Email: {pontusg,boyd}@stanford.edu

$\sqrt{x^T M x}$ denotes the $\|\cdot\|_M$ -norm when $M \in \mathbf{R}^{n \times n}$ is positive definite, but is also used for the $\|\cdot\|_M$ -semi norm when $M \in \mathbf{R}^{n \times n}$ is merely positive semi-definite. We denote by $I_{\mathcal{X}}$ and $I_{g(\cdot) \leq c}$ the indicator functions for the sets \mathcal{X} and $\mathcal{Y} = \{y \in \text{dom}(g) \mid g(y) \leq c\}$ respectively. The conjugate function to f is denoted by $f^*(y) \triangleq \sup_x \{\langle y, x \rangle - f(x)\}$. Finally, the class of closed, proper, and convex functions $f : \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ is denoted by $\Gamma_0(\mathbf{R}^n)$.

B. Preliminaries

In this section, we introduce some definitions and preliminary results to be referenced later. We will introduce these concepts on the Euclidean space \mathbf{R}^n , but most definitions and results hold for general Hilbert spaces. The definitions and results stated below are standard and can be found, e.g. in [26], [2], [20]. Nonstandard results are given short proofs.

Definition 1: A *set-valued operator* (or *operator*) $A : \mathbf{R}^n \rightrightarrows \mathbf{R}^n$ maps each element in \mathbf{R}^n to a set in \mathbf{R}^n .

Definition 2: A *single-valued operator* (or *mapping*) is an operator A that is *single-valued* everywhere on \mathbf{R}^n , i.e. $A(x)$ is a singleton for all $x \in \mathbf{R}^n$. This is denoted by $A : \mathbf{R}^n \rightarrow \mathbf{R}^n$.

Definition 3: The *graph* of a set-valued operator $A : \mathbf{R}^n \rightrightarrows \mathbf{R}^n$ is defined as

$$\text{gph}(A) := \{(x, u) \mid u \in A(x)\}.$$

Any set-valued operator is (uniquely) described by its graph.

Definition 4: An operator $A : \mathbf{R}^n \rightrightarrows \mathbf{R}^n$ is *monotone* if

$$\langle u - v, x - y \rangle \geq 0$$

for all $(x, u) \in \text{gph}(A)$ and $(y, v) \in \text{gph}(A)$.

Definition 5: An operator $A : \mathbf{R}^n \rightrightarrows \mathbf{R}^n$ is σ -*strongly monotone* if

$$\langle u - v, x - y \rangle \geq \sigma \|x - y\|^2$$

for all $(x, u) \in \text{gph}(A)$ and $(y, v) \in \text{gph}(A)$.

Definition 6: A monotone operator $A : \mathbf{R}^n \rightrightarrows \mathbf{R}^n$ is *maximal monotone* if $\text{gph}(A)$ is not a proper subset of the graph of any other monotone operator $\hat{A} : \mathbf{R}^n \rightrightarrows \mathbf{R}^n$.

Example 1: The subdifferential ∂f of a function $f \in \Gamma_0(\mathbf{R}^n)$ is maximally monotone. (The reverse statement is, however, not true.)

Definition 7: A (single-valued) mapping $A : \mathbf{R}^n \rightarrow \mathbf{R}^n$ is β -*Lipschitz continuous* if

$$\|A(x) - A(y)\| \leq \beta \|x - y\|.$$

If $\beta = 1$ then A is *nonexpansive* and if $\beta \in (0, 1)$ then A is *contractive*.

Definition 8: A function $f \in \Gamma_0(\mathbf{R}^n)$ is β -*strongly convex* if

$$f(x) \geq f(y) + \langle u, x - y \rangle + \frac{\beta}{2} \|x - y\|^2$$

hold for all $x, y \in \mathbf{R}^n$ and all $u \in \partial f(y)$.

Definition 9: A general (nonconvex), closed, function $f : \mathbf{R}^n \rightarrow \mathbf{R}$ is β -*smooth* if it is differentiable and

$$|f(x) - f(y) - \langle \nabla f(y), x - y \rangle| \leq \frac{\beta}{2} \|x - y\|^2 \quad (2)$$

holds for all $x, y \in \mathbf{R}^n$.

Remark 1: If in addition f is convex, i.e. $f \in \Gamma_0(\mathbf{R}^n)$, β -smoothness is defined as that

$$f(x) \leq f(y) + \langle \nabla f(y), x - y \rangle + \frac{\beta}{2} \|x - y\|^2$$

holds for all $x, y \in \mathbf{R}^n$.

In the following proposition, which is proven in [26, Example 12.59, Proposition 12.60], some dual properties are stated.

Proposition 1: Suppose that $f \in \Gamma_0(\mathbf{R}^n)$. Then the following are equivalent:

- (i) f is β -strongly convex
- (ii) ∂f is β -strongly monotone
- (iii) f^* is $\frac{1}{\beta}$ -smooth
- (iv) ∇f^* is $\frac{1}{\beta}$ -Lipschitz continuous

Corollary 1: The converse statement (i.e., with f and f^* interchanged) also holds for $f \in \Gamma_0(\mathbf{R}^n)$ since $f = (f^*)^*$, see [25, Theorem 12.2].

Remark 2: The equivalence between (iii) and (iv) in Proposition 1 holds also for general (not necessarily convex) smooth functions $f : \mathbf{R}^n \rightarrow \mathbf{R}$.

Definition 10: The *resolvent* of a maximal monotone operator $A : \mathbf{R}^n \rightrightarrows \mathbf{R}^n$ is defined as

$$R_A := (I + A)^{-1}.$$

Definition 11: The *proximal operator* of a function $f \in \Gamma_0(\mathbf{R}^n)$ is given by

$$\text{prox}_{\gamma f}(y) := \underset{x}{\text{argmin}} \left\{ f(x) + \frac{1}{2\gamma} \|x - y\|^2 \right\}.$$

The proximal operator is a special case of the resolvent. Specifically, if $A = \gamma \partial f$ for some $f \in \Gamma_0(\mathbf{R}^n)$, we have $R_{\gamma \partial f} = \text{prox}_{\gamma f}$.

Definition 12: The function f_γ is defined as $f_\gamma := \gamma f + \frac{1}{2} \|\cdot\|^2$, where $\gamma > 0$.

Proposition 2: Assume that $f \in \Gamma_0(\mathbf{R}^n)$, then $\text{prox}_{\gamma f}(y) = \nabla f_\gamma^*(y)$, where f_γ is defined in Definition 12.

Proof. We have

$$\begin{aligned} \text{prox}_{\gamma f} &= \underset{x}{\text{argmin}} \left\{ f(x) + \frac{1}{2\gamma} \|x - y\|^2 \right\} \\ &= \underset{x}{\text{argmax}} \left\{ \langle y, x \rangle - \gamma f(x) - \frac{1}{2} \|x\|^2 \right\} \\ &= \underset{x}{\text{argmax}} \left\{ \langle y, x \rangle - f_\gamma(x) \right\} = \partial f_\gamma^*(y) \end{aligned}$$

where the last step is due to [25, Theorem 23.5]. Further, since f_γ is 1-strongly convex, Proposition 1 implies that f_γ^* is smooth, hence differentiable. \square

Proposition 3: Assume that $f \in \Gamma_0(\mathbf{R}^n)$ is β -strongly convex. Then $\text{prox}_{\gamma f} : \mathbf{R}^n \rightarrow \mathbf{R}^n$ is $\frac{1}{1+\gamma\beta}$ -contractive.

Proof. Since f is β -strongly convex, f_γ is $(1+\gamma\beta)$ -strongly convex. Apply Propositions 2 and 1 to get the result. \square

Proposition 4: Assume that $f \in \Gamma_0(\mathbf{R}^n)$ is β -smooth. Then $\text{prox}_{\gamma f} : \mathbf{R}^n \rightarrow \mathbf{R}^n$ is $\frac{1}{1+\gamma\beta}$ -strongly monotone.

Proof. Since f is β -smooth, f_γ is $(1 + \gamma\beta)$ -smooth. Apply Propositions 2 and 1 to get the result. \square

Definition 13: The *reflected proximal operator* (also referred to as *Cayley operator*) to $f \in \Gamma_0(\mathbf{R}^n)$ is defined as

$$C_{\gamma f} := 2\text{prox}_{\gamma f} - I.$$

III. PROBLEM FORMULATION

We consider convex composite optimization problems of the form

$$\text{minimize } f(x) + g(Ax) \quad (3)$$

that satisfy the following assumptions:

Assumption 1:

- (i) The function $f : \mathbf{R}^n \rightarrow \mathbf{R}$ is such that $f - \frac{1}{2}\|\cdot\|_H^2$ is convex and $\frac{1}{2}\|\cdot\|_M^2 - f$ is convex.
- (ii) The function $g \in \Gamma_0(\mathbf{R}^m)$.
- (iii) The matrix $A \in \mathbf{R}^m \rightarrow \mathbf{R}^n$ has full row rank.

Remark 3: That $f - \frac{1}{2}\|\cdot\|_H^2$ is convex is equivalent to that f satisfies

$$f(x) \geq f(y) + \langle u, x - y \rangle + \frac{1}{2}\|x - y\|_H^2 \quad (4)$$

for all $x, y \in \mathbf{R}^n$ and $u \in \partial f(y)$, i.e., that f is 1-strongly convex w.r.t. a space with inner product $\langle x, y \rangle = x^T y$ and norm $\|x\|_H$. This implies that f is $\lambda_{\min}(H)$ -strongly convex w.r.t. the Euclidean space. That $\frac{1}{2}\|\cdot\|_M^2 - f$ is convex implies that f satisfies

$$f(x) \leq f(y) + \langle \nabla f(y), x - y \rangle + \frac{1}{2}\|x - y\|_M^2 \quad (5)$$

for all $x, y \in \mathbf{R}^n$, i.e., that f is 1-smooth w.r.t. a space with inner product $\langle x, y \rangle = x^T y$ and norm $\|x\|_M$. This implies that f is $\lambda_{\max}(M)$ -smooth w.r.t. the Euclidean space. If $f(x) = \frac{1}{2}x^T H x + h^T x$, then $f - \frac{1}{2}\|\cdot\|_H^2$ and $\frac{1}{2}\|\cdot\|_H^2 - f$ are convex, and the upper and lower bounds (4) and (5) coincide.

The (negative) Fenchel dual problem to (3) is given by

$$\text{minimize } d(\mu) + g^*(\mu) \quad (6)$$

where

$$d(\mu) := f^*(-A^T \mu). \quad (7)$$

The problem (3) can be solved either directly, or via its dual (6). In this paper, we will solve (3) by applying generalized Douglas-Rachford splitting to the primal problem (3), and by applying generalized Douglas-Rachford splitting to the dual problem (6) (which is equivalent to applying ADMM to the primal (3)). Under the assumptions stated in Assumption 1, we show linear convergence of the generalized Douglas-Rachford algorithm when applied to the primal problem (3) and to the dual problem (6). We also show how to scale the problem data and select algorithm parameters to optimize the obtained linear convergence rate bound.

IV. GENERALIZED DOUGLAS RACHFORD SPLITTING

The generalized Douglas Rachford algorithm can be applied to solve problems of the form (3). It is most computationally efficient in the case where $A = I$, which is why we restrict ourselves to this case. The generalized Douglas-Rachford algorithm is given by the iteration

$$z^{k+1} = ((1 - \theta)I + \theta C_{\gamma g} C_{\gamma f}) z^k \quad (8)$$

where $C_{\gamma f} = 2\text{prox}_{\gamma f} - I$ is the reflected proximal operator in Definition 13 and $\theta \in (0, 1]$. The iteration (8) is a θ -averaged iteration of $C_{\gamma g} C_{\gamma f}$, i.e., it goes a fraction θ of the distance to the point $C_{\gamma g} C_{\gamma f} z^k$ from the current point z^k . Since $C_{\gamma g}$ and $C_{\gamma f}$ (and hence $C_{\gamma g} C_{\gamma f}$) are nonexpansive in the general case [2, Corollary 23.10], the generalized Douglas-Rachford algorithm converges sub-linearly to a fixed point of $C_{\gamma g} C_{\gamma f}$ for $\theta \in (0, 1)$ (since θ -averaged iterations of nonexpansive operators converge sub-linearly to a fixed-point of the nonexpansive operator [2, Corollary 5.14]). Special cases of (8) are obtained by letting $\theta = \frac{1}{2}$, which gives the standard Douglas-Rachford algorithm [10], and by letting $\theta = 1$, which gives the Peaceman-Rachford algorithm [23]. That is, the Peaceman-Rachford algorithm is a non-averaged iteration of $C_{\gamma g} C_{\gamma f}$ and cannot be guaranteed to converge in the general case.

A more explicit formulation of the generalized Douglas-Rachford algorithm is given by the iterations

$$\begin{aligned} x^k &= \underset{x}{\text{argmin}} \left\{ f(x) + \frac{1}{2\gamma} \|x - z^k\|^2 \right\} \\ y^k &= \underset{y}{\text{argmin}} \left\{ g(y) + \frac{1}{2\gamma} \|y - 2x^k + z^k\|^2 \right\} \\ z^{k+1} &= z^k + 2\theta(y^k - x^k) \end{aligned}$$

where x^k is known to converge to an optimal solution of (3) (if it exists) and z^k (which is the same as z^k in (8)) converges sub-linearly to a fixed-point of $C_{\gamma g} C_{\gamma f}$ (if it exists), see [2, Corollary 27.7]. The convergence of the generalized Douglas-Rachford algorithm is for some problem classes linear. We will show that Assumption 1 defines one such class. To show this, we first show that $C_{\gamma f}$ is contractive under Assumption 1.

Proposition 5: Suppose that Assumption 1 holds. Further, let $L = \lambda_{\max}(M)$, $\sigma = \lambda_{\min}(H)$, and $\kappa = L/\sigma$, and let the algorithm parameter $\gamma = \frac{1}{\sqrt{\sigma L}}$. Then $C_{\gamma f}$ is β -contractive with $\beta = \frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}$.

A proof to this proposition is found in Appendix I.

Based on this contraction result, we establish a linear rate of convergence for generalized Douglas-Rachford splitting in the following proposition. This proposition is proven in Appendix II.

Proposition 6: Suppose that Assumption 1 holds. Then for $\gamma = 1/\sqrt{\lambda_{\max}(M)\lambda_{\min}(H)}$ the generalized Douglas-Rachford algorithm (8) converges linearly with rate $\frac{\sqrt{\kappa}+1-2\theta}{\sqrt{\kappa}+1}$, i.e.

$$\|z^{k+1} - \bar{z}\| \leq \left(\frac{\sqrt{\kappa}+1-2\theta}{\sqrt{\kappa}+1} \right)^k \|z^0 - \bar{z}\|$$

where $\kappa = \frac{\lambda_{\max}(M)}{\lambda_{\min}(H)}$ and \bar{z} is a fixed-point of (8).

The bound on the linear convergence rate for the generalized Douglas-Rachford algorithm that is established in Proposition 6, is a decreasing function of θ . Thus, the larger θ is, the better the bound becomes. This implies that Peaceman-Rachford splitting ((8) with $\theta = 1$) has a better theoretical rate $\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}$ than standard Douglas-Rachford ((8) with $\theta = \frac{1}{2}$), which has rate $\frac{\sqrt{\kappa}}{\sqrt{\kappa}+1}$, when Assumption 1 holds.

From the convergence rate result in Proposition 6, we also conclude that the bound on the linear convergence rate of the generalized Douglas-Rachford algorithm is improved by improving the conditioning of f i.e. by reducing $\kappa = \lambda_{\max}(M)/\lambda_{\min}(H)$. This can be achieved in different ways. One option is to choose a space on which κ is small and to apply the algorithm on that space. For instance if $f(x) = \frac{1}{2}x^T Hx$, then $\kappa = 1$ on a space equipped with inner product $\langle x, y \rangle_H = x^T H y$ and induced norm $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle_H}$, while κ on \mathbf{R}^n is $\frac{\lambda_{\max}(H)}{\lambda_{\min}(H)}$. Another option is to perform the generalized Douglas-Rachford iterations in the Euclidean space, and to perform a linear change of variables that minimizes κ for the new problem. In this paper we have chosen to focus on the latter approach.

A. Preconditioning

To improve the conditioning of the problem to be solved by the generalized Douglas-Rachford algorithm, we perform a linear change of variables $Dq = x$, where $D \in \mathbf{R}^{n \times n}$ is invertible, and define the functions

$$\begin{aligned} f_D(q) &:= f(Dq) \\ g_D(q) &:= g(Dq). \end{aligned}$$

Then, we apply generalized Douglas-Rachford splitting to

$$\text{minimize } f_D(x) + g_D(x)$$

to get

$$\begin{aligned} q^k &= \operatorname{argmin}_q \left\{ f_D(q) + \frac{1}{2\gamma} \|q - v^k\|^2 \right\} \\ r^k &= \operatorname{argmin}_r \left\{ g_D(r) + \frac{1}{2\gamma} \|r - 2q^k + v^k\|^2 \right\} \\ v^{k+1} &= v^k + \theta(r^k - q^k). \end{aligned}$$

The smoothness and strong convexity properties of f_D are given by the following proposition.

Proposition 7: Suppose that Assumption 1 holds. Then f_D is $\lambda_{\min}(DHD)$ -strongly convex and $\lambda_{\max}(DM D)$ -smooth.

Proof. Assumption 1 implies that (4) holds. By letting $Dq = x$ and $Dr = y$, we get $D\partial f(y) = D\partial f(Dr) = \partial f_D(r)$, and from (4) we get

$$\begin{aligned} f(x) &\geq f(y) + \langle u, x - y \rangle + \frac{1}{2} \|x - y\|_H^2 \\ &= f(Dr) + \langle D^{-1}w, D(q - r) \rangle + \frac{1}{2} \|Dr - Dq\|_H^2 \\ &= f_D(r) + \langle w, q - r \rangle + \frac{1}{2} \|r - q\|_{D^T H D}^2. \end{aligned}$$

Since $f(x) = f_D(q)$, this implies that f_D is $\lambda_{\min}(DHD)$ -strongly convex. The smoothness claim is proven similarly

by using the smoothness definition in Remark 1 instead of the strong convexity definition in Definition 8. \square

Ideally, the matrix D should be chosen to minimize the ratio $\lambda_{\max}(D^T M D)/\lambda_{\min}(D^T H D)$. However, the prox-operator for the q^k -update (and similarly for the r^k update) is evaluated as

$$\begin{aligned} q^k &= \operatorname{argmin}_q \left\{ f_D(q) + \frac{1}{2\gamma} \|q - v^k\|^2 \right\} \\ &= D^{-1} \operatorname{argmin}_x \left\{ f(x) + \frac{1}{2\gamma} \|D^{-1}x - v^k\|^2 \right\} \\ &= D^{-1} \operatorname{argmin}_x \left\{ f(x) + \frac{1}{2\gamma} \|x - Dv^k\|_{(DD^T)^{-1}}^2 \right\}. \end{aligned}$$

Often f or g is separable down to the component. For such problems, choosing D non-diagonal would increase the computational burden in each iteration. Thus, the objective is to minimize the ratio $\lambda_{\max}(D^T M D)/\lambda_{\min}(D^T H D)$ using a diagonal preconditioner D . The reader is referred to [15, Section 6], for methods to minimize such a ratio and for heuristics to reduce the same.

V. ADMM

ADMM is Douglas-Rachford splitting applied to the dual problem (6), i.e. generalized Douglas-Rachford splitting with $\theta = \frac{1}{2}$. Under- and over-relaxed ADMM is obtained by applying generalized Douglas-Rachford splitting to the dual, where $\theta < \frac{1}{2}$ gives under-relaxation and $\theta > \frac{1}{2}$ gives over-relaxation, [11]. The linear convergence result of the generalized Douglas-Rachford algorithm in Proposition 6 is based on strong convexity and smoothness assumptions on the primal function f . In the following proposition, we show similar properties for the dual function d , defined in (7), under Assumption 1.

Proposition 8: Suppose that Assumption 1 holds. Then convexity of $f - \|\cdot\|_H^2$ implies that d , defined in (7), satisfies

$$d(\mu) \leq d(\nu) + \langle \nabla d(\nu), \mu - \nu \rangle + \frac{1}{2} \|\mu - \nu\|_{AH^{-1}A^T}^2 \quad (9)$$

for all $\mu, \nu \in \mathbf{R}^m$. Further, convexity of $\|\cdot\|_M^2 - f$ implies that

$$d(\mu) \geq d(\nu) + \langle \nabla d(\nu), \mu - \nu \rangle + \frac{1}{2} \|\mu - \nu\|_{AM^{-1}A^T}^2 \quad (10)$$

holds for all $\mu, \nu \in \mathbf{R}^m$.

Proof. The first claim follows from [15, Proposition 29]. The second claim follows from [15, Proposition 29] by using [15, Remark 19] instead of [15, Proposition 18]. \square

The result implies that d is $\lambda_{\max}(AH^{-1}A^T)$ -smooth and, if in addition A has full row rank (as in Assumption 1), $\lambda_{\min}(AM^{-1}A^T)$ -strongly convex. Thus, when applying generalized Douglas-Rachford splitting to the dual problem (6) (or equivalently applying ADMM to the primal), Proposition 6 shows a linear convergence of the algorithm. Proposition 6 further suggests that the convergence is faster for smaller ratios $\lambda_{\max}(AH^{-1}A^T)/\lambda_{\min}(AM^{-1}A^T)$ and with maximal over-relaxation (i.e. $\theta = 1$). Next, we will show how to precondition the problem data to improve the ratio $\lambda_{\max}(AH^{-1}A^T)/\lambda_{\min}(AM^{-1}A^T)$.

A. Preconditioning

Similarly to in the primal Douglas-Rachford case, we precondition the problem data by performing a linear change of variables $E^T \nu = \mu$, with invertible $E \in \mathbf{R}^{m \times m}$, to get

$$\begin{aligned} d_E(\nu) &:= d(E^T \nu) \\ g_E^*(\nu) &:= g^*(E^T \nu). \end{aligned}$$

Then we apply generalized Douglas-Rachford splitting to

$$\text{minimize } d_E(\nu) + g_E^*(\nu) \quad (11)$$

which is the dual of

$$\begin{aligned} \text{minimize } & f(x) + g(y) \\ \text{subject to } & EAx = Ey. \end{aligned} \quad (12)$$

Proposition 8 implies that d_E is $\lambda_{\max}(EAH^{-1}A^TE^T)$ -smooth and $\lambda_{\min}(EAM^{-1}A^TE^T)$ -strongly convex. Thus, to optimize the linear convergence rate bound, the ratio $\lambda_{\max}(EAH^{-1}A^TE^T)/\lambda_{\min}(EAM^{-1}A^TE^T)$ should be minimized. As in the primal Douglas-Rachford case, this minimization is subject to structural constraints on E . Typically, if f or g is separable, the prox-operations in ADMM are kept simple if E is diagonal. Again, the reader is referred to [15, Section 6] for different methods to find diagonal and non-diagonal E that reduce or minimize the ratio $\lambda_{\max}(EAH^{-1}A^TE^T)/\lambda_{\min}(EAM^{-1}A^TE^T)$.

VI. EXTENSIONS

In this section, we discuss how to compute preconditioners and select the γ -parameter when some of the assumption in Assumption 1 are not met. We discuss loss of smoothness in the objective, and loss of rank condition on the matrix A . Loss of strong convexity in the objective can be treated similarly to the loss of smoothness, but is omitted due to space considerations.

The extensions will be presented for ADMM, since it efficiently solves more general problems than Douglas-Rachford splitting on the primal. Also, ADMM, or Douglas-Rachford splitting on the dual, for the case of $A = I$ is equivalent to Douglas-Rachford splitting on the primal (but the preconditioners relate to each other as $D = E^{-1}$).

A. Loss of smoothness

In this section, we assume that Assumption 1 holds, but that $\|\cdot\|_M^2 - f$ is not convex, i.e. that f is not smooth. This occurs for instance when solving problems of the form

$$\text{minimize } \underbrace{\frac{1}{2}x^T Hx + h^T x + \hat{f}(x)}_{f(x)} + g(Ax)$$

where $H \in \mathbf{R}^{n \times n}$ is positive definite, $h \in \mathbf{R}^n$, $A \in \mathbf{R}^{m \times n}$ with rank m and $\hat{f} \in \Gamma_0(\mathbf{R}^n)$. Proposition 8 states that the lack of smoothness in f implies that d is not strongly convex. Thus, when solving such problems using ADMM, or Douglas-Rachford splitting on the dual, we loose strong convexity in the objective. However, from Proposition 8 we still have

$$d(\mu) \leq d(\nu) + \langle \nabla d(\nu), \mu - \nu \rangle + \frac{1}{2} \|\mu - \nu\|_{EAH^{-1}A^TE^T}. \quad (13)$$

In [15, Proposition 31], it is shown that for many choices of \hat{f} (such as the 1-norm or the indicator function of a constraint set with non-empty interior) (13) is tight in some full-dimensional subset of \mathbf{R}^m . Here, we use the heuristic to make d as round as possible in that subset, and to choose γ optimally according to the shape of d inside the subset. That is, we propose to select an E that minimizes $\lambda_{\max}(EAH^{-1}A^TE^T)/\lambda_{\min}(EAH^{-1}A^TE^T)$, and to set $\gamma = 1/\sqrt{\lambda_{\max}(EAH^{-1}A^TE^T)\lambda_{\min}(EAH^{-1}A^TE^T)}$ (in accordance with Proposition 6).

For $\hat{f}(x) = I_{Bx=b}(x)$, where $B \in \mathbf{R}^{p \times n}$ has full row rank, it is shown in [15, Proposition 33] that the bound in (13) is not tight (note that the interior of the set defined by $Bx = b$ is empty). In this case, it is shown that d satisfies

$$d(\mu) = d(\nu) + \langle \nabla d(\nu), \mu - \nu \rangle + \frac{1}{2} \|\mu - \nu\|_{EAK_{11}A^TE^T},$$

where

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

This suggests that we should select E to minimize the condition number of $EAK_{11}A^TE^T$. However, the matrix $EAK_{11}A^TE^T$ does not have full rank. Thus, we instead propose to select E that minimizes the ratio $\lambda_{\max}(EAK_{11}A^TE^T)/\lambda_{\min>0}(EAK_{11}A^TE^T)$, where $\lambda_{\min>0}$ denotes the smallest non-zero eigenvalue, and (in accordance with Proposition 6) choose $\gamma = 1/\sqrt{\lambda_{\max}(EAK_{11}A^TE^T)\lambda_{\min>0}(EAK_{11}A^TE^T)}$. Minimization of the pseudo condition number $\lambda_{\max}/\lambda_{\min>0}$ can be posed as a convex optimization problem and be solved exactly, see [15, Section 6] which also contains heuristics to reduce the pseudo condition number.

B. Loss of rank-condition in A

In this section, we consider the case where $A \in \mathbf{R}^{m \times n}$ does not have full row rank. This is common, e.g., when solving problems of the form

$$\begin{aligned} \text{minimize } & f(x) \\ \text{subject to } & Ax \leq d \end{aligned}$$

where there are more constraints than variables. Letting $Ax = y$ and $g(y) = I_{y \leq d}(y)$ results in an inexpensive prox-operation for g , which might not be the case if splitting according to $x = y$ and $g(y) = I_{Ay \leq d}(y)$. Even if A does not have full row-rank, the results in Proposition 8 still hold. However, $EAH^{-1}A^TE^T$ and $EAM^{-1}A^TE^T$ does not have full rank, which means that we cannot minimize the ratio $\lambda_{\max}(EAH^{-1}A^TE^T)/\lambda_{\min}(EAM^{-1}A^TE^T)$. Again, we propose to choose E that minimizes the ratio between the largest and smallest nonzero eigenvalues, i.e. that minimize $\lambda_{\max}(EAH^{-1}A^TE^T)/\lambda_{\min>0}(EAM^{-1}A^TE^T)$, where, again, $\lambda_{\min>0}$ denotes the smallest non-zero eigenvalue. Also, similarly to before, we propose to select γ as $\gamma = 1/\sqrt{\lambda_{\max}(EAH^{-1}A^TE^T)\lambda_{\min>0}(EAM^{-1}A^TE^T)}$.

VII. NUMERICAL EXAMPLE

In this section, we evaluate the preconditioning by applying ADMM to the (small-scale) aircraft control problem from [18], [4]. As in [4], the continuous time model from [18] 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

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

$$y = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} x$$

where x^+ denotes the state in the next time step. 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

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

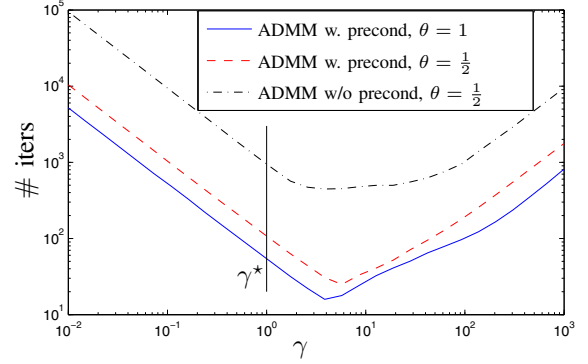
where x_r is a reference, $Q = \text{diag}(10^{-4}, 10^2, 10^{-3}, 10^2)$, $R = 10^{-2}I$, and $S = 10^6I$. This gives a condition number of 10^{10} of the full cost matrix. Further, the terminal cost is Q , and the control and prediction horizon is $N = 10$. The numerical data in Figure 1 is obtained 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. The full optimization problem can be written on the form

$$\begin{aligned} & \text{minimize} && \underbrace{\frac{1}{2}z^T H z + r_t^T z + I_{Bz=bx_t}(z)}_{f(z)} + \underbrace{I_{d \leq y \leq \bar{d}}(z')}_{g(z')} \\ & \text{subject to} && C z = z' \end{aligned}$$

where x_t and r_t may change from one sampling instant to the next.

This is exactly the optimization problem formulation discussed in the end of Section VI-A. In Figure 1, the performance of the ADMM algorithm for different values of γ and for different preconditioning is presented. Since the numerical example treated here is a model predictive control application, we can spend much computational effort offline to compute a preconditioner that will be used in all samples in the online controller. We compute a preconditioner E that minimizes the condition number of $ECH^{-1}C^T E^T$ (minimization of the (pseudo) condition number of $ECK_{11}C^T E^T$ gives about the same performance and is therefore omitted). In Figure 1, the performance of ADMM with and without preconditioning is shown. The figure also compares Douglas-Rachford applied on the dual (i.e. ADMM) and Peaceman-Rachford applied on the dual, (i.e. ADMM with over-relaxation $\theta = 1$), when using preconditioning. In this particular example, improvements of about one order of magnitude

Fig. 1. Average number of iterations for different γ -values, with and without preconditioning, and for different relaxation θ .



are achieved when using preconditioning compared to when no preconditioning is used. Figure 1 also shows that ADMM with over-relaxation performs better than standard ADMM with no relaxation. The empirically best average iteration count for over-relaxed ADMM when using preconditioning is 15.9 iterations, for standard ADMM it is 24.9 iterations, and for standard ADMM without preconditioning, (which is the algorithm proposed in [21]), is 446.1 iterations. The proposed γ -parameter selection is denoted by γ^* in Figure 1 (E or C is scaled to get $\gamma^* = 1$ for all examples). Figure 1 shows that γ^* does not coincide with the empirically found best γ , but still gives a reasonable choice of γ in all cases.

VIII. CONCLUSIONS

We have presented methods to scale the problem data and select algorithm parameters for Douglas-Rachford splitting, Peaceman-Rachford splitting and ADMM. A numerical example is provided in which the scaling improves the performance of the ADMM algorithm with about one order of magnitude compared to when no scaling is used.

REFERENCES

- [1] K. J. Arrow, L. Hurwicz, and H. Uzawa. *Studies in Linear and Nonlinear Programming*. Stanford University Press, 1958.
- [2] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, 2011.
- [3] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sciences*, 2(1):183–202, 2009.
- [4] 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.
- [5] 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.
- [6] E. J. Candes, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, Feb 2006.
- [7] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011.
- [8] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *SIAM journal on Multiscale Modeling and Simulation*, 4(4):1168–1200, 2005.

- [9] W. Deng and W. Yin. On the global and linear convergence of the generalized alternating direction method of multipliers. Technical Report CAAM 12-14, Rice University, 2012.
- [10] J. Douglas and H. H. Rachford. On the numerical solution of heat conduction problems in two and three space variables. *Trans. Amer. Math. Soc.*, 82:421–439, 1956.
- [11] J. Eckstein. *Splitting methods for monotone operators with applications to parallel optimization*. PhD thesis, MIT, 1989.
- [12] D. Gabay. Applications of the method of multipliers to variational inequalities. In M. Fortin and R. Glowinski, editors, *Augmented Lagrangian Methods: Applications to the Solution of Boundary-Value Problems*. North-Holland: Amsterdam, 1983.
- [13] D. Gabay and B. Mercier. A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Computers and Mathematics with Applications*, 2(1):17–40, 1976.
- [14] E. Ghadimi, A. Teixeira, I. Shames, and M. Johansson. Optimal parameter selection for the alternating direction method of multipliers (ADMM): Quadratic problems. 2013. Submitted. Available: <http://arxiv.org/abs/1306.2454>.
- [15] P. Giselsson and S. Boyd. Metric selection in fast dual forward backward splitting. *Automatica*, 2014. Submitted.
- [16] R. Glowinski and A. Marroco. Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de dirichlet non linéaires. *ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique*, 9:41–76, 1975.
- [17] T. Hastie, R. Tibshirani, and J. Friedman. *The elements of statistical learning: data mining, inference and prediction*. Springer, 2nd edition, 2009.
- [18] P. Kamasouris, 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.
- [19] P. L. Lions and B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis*, 16(6):964–979, 1979.
- [20] Y. Nesterov. *Introductory Lectures on Convex Optimization: A Basic Course*. Springer Netherlands, 1st edition, 2003.
- [21] 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.
- [22] N. Parikh and S. Boyd. Proximal algorithms. *Foundations and Trends in Optimization*, 1(3):123–231, 2014.
- [23] D. W. Peaceman and H. H. Rachford. The numerical solution of parabolic and elliptic differential equations. *Journal of the Society for Industrial and Applied Mathematics*, 3(1):28–41, 1955.
- [24] J. B. Rawlings and D. Q. Mayne. *Model Predictive Control: Theory and Design*. Nob Hill Publishing, Madison, WI, 2009.
- [25] R. T. Rockafellar. *Convex Analysis*, volume 28. Princeton University Press, Princeton, NJ, 1970.
- [26] R. T. Rockafellar and R. J-B. Wets. *Variational Analysis*. Springer, Berlin, 1998.

APPENDIX I PROOF OF PROPOSITION 5

Proof. Define $\hat{f} = 2f_\gamma^* - \frac{1}{2} \|\cdot\|^2$ (where f_γ is defined in Definition 12). Through Proposition 2, we get that $\nabla \hat{f} = 2\nabla f_\gamma^* - I = 2\text{prox}_{\gamma f} - I = C_{\gamma f}$. We get

$$\begin{aligned} \langle \nabla \hat{f}(y), x - y \rangle &= \langle 2\nabla f_\gamma^*(y) - y, x - y \rangle \\ &\leq 2(f_\gamma^*(x) - f_\gamma^*(y) - \frac{1}{2(\gamma L + 1)} \|x - y\|^2) \\ &\quad - (\frac{1}{2} \|x\|^2 - \frac{1}{2} \|y\|^2 - \frac{1}{2} \|x - y\|^2) \\ &= \hat{f}(x) - \hat{f}(y) + \frac{\gamma L - 1}{2(\gamma L + 1)} \|x - y\|^2 \end{aligned}$$

where Proposition 4 and Proposition 1 are used in the inequality. We also have

$$\begin{aligned} \langle \nabla \hat{f}(y), x - y \rangle &= \langle 2\nabla f_\gamma^*(y) - y, x - y \rangle \\ &\geq 2(f_\gamma^*(x) - f_\gamma^*(y) - \frac{1}{2(\gamma\sigma + 1)} \|x - y\|^2) \\ &\quad - (\frac{1}{2} \|x\|^2 - \frac{1}{2} \|y\|^2 - \frac{1}{2} \|x - y\|^2) \\ &= \hat{f}(x) - \hat{f}(y) + \frac{\gamma\sigma - 1}{2(\gamma\sigma + 1)} \|x - y\|^2 \end{aligned}$$

where Proposition 3 and Proposition 1 are used in the inequality. This implies that

$$\begin{aligned} \frac{\gamma\sigma - 1}{2(\gamma\sigma + 1)} \|x - y\|^2 &\leq \langle \nabla \hat{f}(y), x - y \rangle + \hat{f}(y) - \hat{f}(x) \\ &\leq \frac{\gamma L - 1}{2(\gamma L + 1)} \|x - y\|^2 \end{aligned}$$

or equivalently (by negating the first inequality)

$$\begin{aligned} |\langle \nabla \hat{f}(y), x - y \rangle + \hat{f}(y) - \hat{f}(x)| \\ \leq \frac{1}{2} \max\left(\frac{\gamma L - 1}{\gamma L + 1}, \frac{1 - \gamma\sigma}{\gamma\sigma + 1}\right) \|x - y\|^2. \end{aligned} \quad (14)$$

Since $\nabla \hat{f} = C_{\gamma f}$, Remark 2 and (14) imply that $C_{\gamma f}$ is β -Lipschitz continuous, with

$$\beta = \max\left(\frac{\gamma L - 1}{1 + \gamma L}, \frac{1 - \gamma\sigma}{1 + \gamma\sigma}\right).$$

To minimize β , γ should be chosen such that the arguments have equal magnitude, i.e. $\frac{\gamma L - 1}{1 + \gamma L} = \frac{1 - \gamma\sigma}{1 + \gamma\sigma}$, which is obtained by letting $\gamma = \frac{1}{\sqrt{\sigma L}}$. This choice gives $\beta = \frac{\gamma L - 1}{1 + \gamma L} = \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} < 1$, which concludes the proof. \square

APPENDIX II PROOF OF PROPOSITION 6

Proof. By [2, Corollary 23.10] $C_{\gamma g}$ is nonexpansive and by Proposition 5 $C_{\gamma f}$ is $\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}$ -contractive. Thus the composition $C_{\gamma g} C_{\gamma f}$ is $\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}$ -contractive since

$$\begin{aligned} \|C_{\gamma g} C_{\gamma f} z_1 - C_{\gamma g} C_{\gamma f} z_2\| &\leq \|C_{\gamma f} z_1 - C_{\gamma f} z_2\| \\ &\leq \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \|z_1 - z_2\|. \end{aligned} \quad (15)$$

Now, let $T = (1 - \theta)I + \theta C_{\gamma g} C_{\gamma f}$ be the generalized Douglas-Rachford operator in (8). Since $\bar{z} = T\bar{z}$, we get

$$\begin{aligned} \|z^{k+1} - \bar{z}\| &= \|Tz^k - T\bar{z}\| \\ &= \|(1 - \theta)(z^k - \bar{z}) + \theta(C_{\gamma g} C_{\gamma f} z^k - C_{\gamma g} C_{\gamma f} \bar{z})\| \\ &\leq \theta \|(C_{\gamma g} C_{\gamma f} z^k - C_{\gamma g} C_{\gamma f} \bar{z})\| + (1 - \theta) \|z^k - \bar{z}\| \\ &\leq \left(\theta \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} + (1 - \theta)\right) \|z^k - \bar{z}\| \\ &= \frac{\sqrt{\kappa} + 1 - 2\theta}{\sqrt{\kappa} + 1} \|z^k - \bar{z}\| \end{aligned}$$

where (15) is used in the second inequality. This concludes the proof. \square