Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 53-2 (2020) 7350–7354

Global exponential stability of the Douglas-Rachford splitting dynamics ⋆

Sepideh Hassan-Moghaddam and Mihailo R. Jovanović

Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA 90089; (e-mails: hassanmo@usc.edu, mihailo@usc.edu)

Abstract: Many modern optimization problems admit a composite form in which the objective function is given by the sum of a smooth term and a nonsmooth regularizer. Such problems can be solved via proximal methods and their variants, including the Douglas-Rachford (DR) splitting algorithm. In this paper, we view the DR splitting flow as a dynamical system and leverage techniques from control theory to study its global stability properties. In particular, for problems with strongly convex objective functions, we utilize the theory of integral quadratic constraints to prove global exponential stability of the ordinary differential equation that governs the evolution of the DR splitting flow. In our analysis, we use the fact that this algorithm can be interpreted as a variable-metric gradient method on the DR envelope and exploit structural properties of nonlinear terms that arise from composition of reflected proximal operators.

Keywords: Control for optimization, Douglas-Rachford splitting, global exponential stability, integral quadratic constraints, nonlinear dynamics, nonsmooth optimization, proximal algorithms.

1. INTRODUCTION

Examining optimization algorithms as continuous-time dynamical systems has been an active topic since the seminal work of Arrow, Hurwicz, and Uzawa (Arrow et al., 1958). This viewpoint can provide important insight into performance of optimization algorithms and, during the last decade, it has been advanced and extended to a broad class of problems including convergence analysis of primal-dual (Feijer and Paganini, 2010; Wang and Elia, 2011; Cherukuri et al., 2016, 2018; Dhingra et al., 2019; Qu and Li, 2018) and accelerated (Su et al., 2016; Wibisono et al., 2016; França et al., 2018; Shi et al., 2018; Muehlebach and Jordan, 2019; Poveda and Li, 2019) first-order methods. Furthermore, establishing the connection between theory of ordinary differential equations (ODEs) and numerical optimization algorithms has been a topic of many studies, including Schropp and Singer (2000); for recent efforts, see Wibisono et al. (2016); Zhang et al. (2018).

Most algorithms can be viewed as a feedback interconnection of linear dynamical systems with nonlinearities that posses certain structural properties. This system-theoretic interpretation was exploited in Lessard et al. (2016) and further advanced in a number or recent papers (Dhingra et al., 2019; Hu et al., 2017; Hu and Lessard, 2017; Fazlyab et al., 2018; Hatanaka et al., 2018; Hassan-Moghaddam and Jovanović, 2018a,b; Seidman et al., 2019). The key idea is to exploit structural features of linear and nonlinear terms and utilize theory and techniques from stability analysis of nonlinear dynamical systems to study properties of optimization algorithms. This approach provides new methods for studying not only convergence rate but also robustness of optimization routines (Mohammadi et al., 2018, 2019a,b; Michalowsky et al., 2019) and can lead to new classes of algorithms that strike a desired tradeoff between the speed and robustness.

In this paper, we utilize techniques from control theory to establish global stability properties of an ordinary differential equation that describes the Douglas-Rachford (DR) splitting flows. This algorithm provide an effective tool for solving nonsmooth convex optimization problems in which the objective function is given by a sum of a differentiable term and a nondifferentiable regularizer. For strongly convex problems, we exploit the fact that the DR splitting algorithm (Douglas and Rachford, 1956) can be interpreted as a variable-metric gradient method on DR envelope (Patrinos et al., 2014) and prove global exponential stability by utilizing the theory of integral quadratic constrains (IQCs).

The paper is structured as follows. In Section 2, we formulate the nonsmooth composite optimization problem and provide background material. In Section 3, we introduce a continuous-time gradient flow dynamics based on the celebrated DR splitting algorithm and utilize the theory of IQCs to prove global exponential stability for strongly convex problems. We conclude the paper in Section 4.

2. PROBLEM FORMULATION AND BACKGROUND

We consider a composite optimization problem,

$$\underset{x}{\text{minimize}}\;\; f(x) \,+\, g(x) \tag{1}$$

where $x \in \mathbb{R}^n$ is the optimization variable, $f\colon \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function with a Lipschitz continuous gradient, and $g\colon \mathbb{R}^m \to \mathbb{R}$ is a nondifferentiable convex function. Such optimization problems arise in a

⋆ Supported in part by the National Science Foundation under awards ECCS-1708906 and ECCS-1809833.

10.1016/j.ifacol.2020.12.1254

number of different applications and different first- and second-order algorithms can be employed to solve them. We are interested in studying global stability properties of the method based on Douglas-Rachford splitting dynamics. In the rest of this section, we provide background material that we utilize in the paper.

2.1 Proximal operators and the associated envelopes

The proximal operator of a proper, closed, and convex function g is defined as

$$\mathbf{prox}_{\mu g}(v) \;:=\; \underset{z}{\operatorname{argmin}} \left(g(z) \;+\; \frac{1}{2\mu}\,\|z \,-\, v\|_2^2 \right).$$

The value function of this optimization problem determines the associated Moreau envelope,

$$M_{\mu g}(v) \;:=\; g(\mathbf{prox}_{\mu g}(v)) \,+\, \frac{1}{2\mu}\,\|\mathbf{prox}_{\mu g}(v) \,-\, v\|_2^2$$

which is a continuously differentiable function even when g is not (Parikh and Boyd, 2013),

$$\nabla M_{\mu g}(v) \;=\; \frac{1}{\mu}\,(v \,-\, \mathbf{prox}_{\mu g}(v)).$$

Combining the last two expressions yields,

$$M_{\mu g}(v) \;=\; g(\mathbf{prox}_{\mu g}(v)) \,+\, \frac{\mu}{2}\,\|\nabla M_{\mu g}(v)\|_2^2.$$

The Moreau envelope of g can be used to introduce the forward-backward (FB) envelope (Patrinos et al., 2014; Stella et al., 2017; Themelis et al., 2018) of the composite function

$$F(x) \;:=\; f(x) \,+\, g(x).$$

The FB envelope is determined by the value function of the problem

$$\underset{v}{\text{minimize}} \;\; J(x,v) \tag{2a}$$

where J approximates F via a quadratic expansion of the function f around x,

$$\begin{aligned} J(x,v) \;:=&\; f(x) \,+\, \langle \nabla f(x), v - x\rangle \,+\, \tfrac{1}{2\mu}\,\|v \,-\, x\|_2^2 \,+\, g(v) \\ =&\; g(v) \,+\, \tfrac{1}{2\mu}\,\|v \,-\, (x \,-\, \mu\nabla f(x))\|_2^2 \,+ \\ &\; f(x) \,-\, \tfrac{\mu}{2}\,\|\nabla f(x)\|_2^2. \end{aligned} \tag{2b}$$

The optimal solution of (2) is determined by

$$v_\mu^\star(x) \;=\; \mathbf{prox}_{\mu g}(x \,-\, \mu\nabla f(x))$$

and it can be used to obtain the FB envelope of the function F,

$$\begin{aligned} F_\mu(x) \;:=&\; J(x, v_\mu^\star(x)) \;=\; J(x, \mathbf{prox}_{\mu g}(x \,-\, \mu\nabla f(x))) \\ =&\; f(x) \,-\, \mu\,\langle \nabla f(x), G_\mu(x)\rangle \,+\, \frac{\mu}{2}\,\|G_\mu(x)\|_2^2 \,+ \\ &\; g(\mathbf{prox}_{\mu g}(x \,-\, \mu\,\nabla f(x))) \end{aligned} \tag{3}$$

where G_μ is the generalized gradient map,

$$G_\mu(x) \;:=\; \frac{1}{\mu}\,(x \,-\, \mathbf{prox}_{\mu g}(x \,-\, \mu\nabla f(x))). \tag{4}$$

Alternatively, the FB envelope F_μ can be also expressed as

$$F_\mu(x) = f(x) \,+\, M_{\mu g}(x \,-\, \mu\,\nabla f(x)) \,-\, \frac{\mu}{2}\,\|\nabla f(x)\|_2^2. \tag{5}$$

Moreover, when f is twice continuously differentiable, F_μ is continuously differentiable and its gradient is determined by (Patrinos et al., 2014),

$$\nabla F_\mu(x) \;=\; (I \,-\, \mu\nabla^2 f(x))\,G_\mu(x). \tag{6}$$

The Douglas-Rachford (DR) envelope is another useful object that is obtained by evaluating the FB envelope at $\mathbf{prox}_{\mu f}(x)$ (Themelis and Patrinos, 2017),

$$F_\mu^{DR}(x) \;:=\; F_\mu(\mathbf{prox}_{\mu f}(x)). \tag{7}$$

Alternatively, the DR envelope can be expressed as

$$\begin{aligned} F_\mu^{DR}(x) \;=&\; M_{\mu g}(x - 2\,\mu\nabla M_{\mu f}(x)) \,+\, M_{\mu f}(x) \,- \\ &\; \mu\|\nabla M_{\mu f}(x)\|_2^2. \end{aligned} \tag{8}$$

From the definition of the proximal operator of the continuously differentiable function f, we have

$$\mu\nabla f(\mathbf{prox}_{\mu f}(x)) \,+\, \mathbf{prox}_{\mu f}(x) \,-\, x \;=\; 0 \tag{9}$$

and, thus,

$$\nabla M_{\mu f}(x) \;=\; \nabla f(\mathbf{prox}_{\mu f}(x)). \tag{10}$$

Equality (7) follows from substituting the expression for $\nabla M_{\mu f}(x)$ into (8), using equation (9), and leveraging the properties of the Moreau envelope,

$$\begin{aligned} F_\mu^{DR}(x) \;=&\; M_{\mu g}(x - 2\,\mu\nabla f(\mathbf{prox}_{\mu f}(x)) + f(\mathbf{prox}_{\mu f}(x)) \,+ \\ &\; \frac{1}{2\mu}\,\|\mathbf{prox}_{\mu f}(x) \,-\, x\|_2^2 - \mu\,\|\nabla f(\mathbf{prox}_{\mu f}(x))\|_2^2 \\ =&\; M_{\mu g}(\mathbf{prox}_{\mu f}(x) -\, \mu\nabla f(\mathbf{prox}_{\mu f}(x)) \,+ \\ &\; f(\mathbf{prox}_{\mu f}(x)) - \frac{\mu}{2}\,\|\nabla f(\mathbf{prox}_{\mu f}(x))\|_2^2 \\ =&\; F_\mu(\mathbf{prox}_{\mu f}(x)). \end{aligned}$$

If f is twice continuously differentiable with $\nabla^2 f(x) \preceq L_f I$ for all x, the DR envelope is continuously differentiable and its gradient is given by (Themelis and Patrinos, 2017)

$$\nabla F_\mu^{DR}(x) \;=\; \frac{1}{\mu}\,(2\nabla\mathbf{prox}_{\mu f}(x) \,-\, I)\,G_\mu^{DR}(x) \tag{11}$$

where

$$\nabla\mathbf{prox}_{\mu f}(x) \;=\; \big(I \,+\, \mu\,\nabla^2 f(\mathbf{prox}_{\mu f}(x))\big)^{-1}$$

and

$$G_\mu^{DR}(x) \;:=\; \mathbf{prox}_{\mu f}(x) - \mathbf{prox}_{\mu g}(2\mathbf{prox}_{\mu f}(x) - x). \tag{12}$$

2.2 Strong convexity and Lipschitz continuity

The function f is strongly convex with parameter m_f if for any x and $\hat{x}$,

$$f(\hat{x}) \;\geq\; f(x) \,+\, \langle \nabla f(x), \hat{x} \,-\, x\rangle \,+\, \frac{m_f}{2}\,\|\hat{x} \,-\, x\|_2^2$$

and equivalently,

$$\|\nabla f(x) - \nabla f(\hat{x})\|_2 \;\geq\; m_f\|x \,-\, \hat{x}\|_2. \tag{13}$$

The gradient of a continuously-differentiable function f is Lipschitz continuous with parameter L_f if for any x and $\hat{x}$,

$$f(\hat{x}) \;\leq\; f(x) \,+\, \langle \nabla f(x), \hat{x} \,-\, x\rangle \,+\, \frac{L_f}{2}\,\|\hat{x} \,-\, x\|_2^2$$

and equivalently,

$$\|\nabla f(x) - \nabla f(\hat{x})\|_2 \;\leq\; L_f\|x \,-\, \hat{x}\|_2. \tag{14}$$

Moreover, if an m_f-strongly convex function f has an L_f-Lipschitz continuous gradient, the following inequality holds for any x and $\hat{x}$,

$$\langle \nabla f(x) - \nabla f(\hat{x}), x - \hat{x} \rangle \geq \frac{m_f L_f}{m_f + L_f} \|x - \hat{x}\|_2^2 + \frac{1}{m_f + L_f} \|\nabla f(x) - \nabla f(\hat{x})\|_2^2. \tag{15}$$

Furthermore, the subgradient ∂g of a nondifferentiable function g is defined as the set of points $z \in \partial g(x)$ that for any x and $\hat{x}$ satisfy,

$$g(\hat{x}) \geq g(x) + z^T(\hat{x} - x). \tag{16}$$

3. GLOBAL EXPONENTIAL STABILITY

We next introduce a continuous-time gradient flow dynamics based on the celebrated Douglas-Rachford splitting algorithm (Douglas and Rachford, 1956) and establish global exponential stability for strongly convex f.

3.1 Douglas-Rachford splitting dynamics

The optimality condition for non-smooth composite optimization problem (1) is given by

$$0 \in \nabla f(x) + \partial g(x).$$

Multiplication by μ and addition/subtraction of x yields,

$$0 \in [I + \mu \nabla f](x) + \mu \partial g(x) - x.$$

Since the proximal operator associated with μf is determined by the resolvent operator of $\mu \nabla f$, we have

$$x = (I + \mu \nabla f)^{-1}(x - \mu \partial g(x)) = \mathbf{prox}_{\mu f}(x - \mu \partial g(x)).$$

Introducing a new variable $z := x - \mu \partial g(x)$ allows us to bring the optimality condition into the following form

$$x = \mathbf{prox}_{\mu f}(z)$$

or, equivalently,

$$\mu \partial g(x) = \mathbf{prox}_{\mu f}(z) - z.$$

Now, adding x to both sides of this equation yields

$$\begin{aligned} [I + \mu \partial g](x) &= x + \mathbf{prox}_{\mu f}(z) - z \\ &= 2\mathbf{prox}_{\mu f}(z) - z \end{aligned}$$

which leads to,

$$x^\star = \mathbf{prox}_{\mu g}(2\,\mathbf{prox}_{\mu f}(z^\star) - z^\star) = \mathbf{prox}_{\mu f}(z^\star). \tag{17a}$$

Furthermore, the reflected proximal operators (Giselsson and Boyd, 2017) of the functions f and g,

$$\begin{aligned} R_{\mu f}(z) &:= [2\,\mathbf{prox}_{\mu f} - I](z) \\ R_{\mu g}(z) &:= [2\,\mathbf{prox}_{\mu g} - I](z) \end{aligned}$$

can be used to write optimality condition (17a) as

$$[R_{\mu g} R_{\mu f}](z^\star) = z^\star. \tag{17b}$$

This follows from (17a) and

$$[R_{\mu g} R_{\mu f}](z) = z + 2(\mathbf{prox}_{\mu g}(2\,\mathbf{prox}_{\mu f}(z) - z) - \mathbf{prox}_{\mu f}(z)).$$

Building on the optimality conditions, the DR splitting algorithm consists of the following iterative steps,

$$\begin{aligned} x^k &= \mathbf{prox}_{\mu f}(z^k) \\ y^k &= \mathbf{prox}_{\mu g}(2x^k - z^k) \\ z^{k+1} &= z^k + 2\alpha(y^k - x^k). \end{aligned} \tag{18}$$

Under standard convexity assumptions (Eckstein and Bertsekas, 1992), the DR splitting algorithm converges for $\alpha \in (0, 1)$. Combining all the steps in (18) yields the first-order recurrence,

$$\begin{aligned} z^{k+1} &= z^k + 2\alpha \left(\mathbf{prox}_{\mu g}(2x^k - z^k) - x^k\right) \\ &= [(1 - \alpha)I + \alpha R_{\mu g} R_{\mu f}](z^k) \end{aligned} \tag{19}$$

where z^k converges to the fixed point of the operator $R_{\mu f} R_{\mu g}$ and x^k converges to the optimal solution of (1).

Optimality conditions (17) can be used to obtain the continuous-time gradient flow dynamics to compute $z^\star$,

$$\begin{aligned} \dot{z} &= -z + [R_{\mu g} R_{\mu f}](z) \\ &= \mathbf{prox}_{\mu g}(2\,\mathbf{prox}_{\mu f}(z) - z) - \mathbf{prox}_{\mu f}(z) \\ &= -G_\mu^{DR}(z) \end{aligned} \tag{20}$$

where $G_\mu^{DR}(z)$ is given by (12). We note that the discrete-time system (19) results from the explicit forward Euler discretization of (20) with the stepsize α.

Remark 1. Using the definition of $\nabla F_\mu^{DR}(x)$ in (11), the continuous-time system (20) can be written as

$$\dot{z} = -\mu \left(2\nabla \mathbf{prox}_{\mu f}(z) - I\right)^{-1} \nabla F_\mu^{DR}(z)$$

where the inverse is well-defined for $\mu \in (0, 1/L_f)$. Thus, the DR splitting algorithm can be interpreted as a variable-metric gradient method on the DR envelope F_μ^{DR} (Patrinos et al., 2014).

3.2 Global exponential stability via theory of IQCs

The continuous-time dynamics (20) can be also seen as a feedback interconnection of an LTI system

$$\begin{aligned} \dot{z} &= A\,z + B\,u \\ \xi &= C\,z \end{aligned} \tag{21a}$$

with a nonlinear term,

$$u(z) := [R_{\mu g} R_{\mu f}](z). \tag{21b}$$

Herein, the matrices in (21a) are given by

$$A = -I,\; B = I,\; C = I \tag{21c}$$

and the corresponding transfer function is

$$H(s) = C(sI - A)^{-1}B = \frac{1}{s + 1}\,I. \tag{21d}$$

We first characterize properties of nonlinearity u in (21b) and then utilize the theory of integral quadratic constraints to establish the conditions for global exponential stability of (21) under the following assumption.

Assumption 1. Let the differentiable part f of the objective function in (1) be strongly convex with parameter m_f, let ∇f be Lipschitz continuous with parameter L_f, and let the regularization function g be proper, closed, and convex.

Lemma 1. Let Assumption 1 hold and let $\mu \in (0, 2/L_f)$. Then, the operator $R_{\mu f}$ is σ-contractive,

$$\|R_{\mu f}(x) - R_{\mu f}(y)\|_2 \leq \sigma \|x - y\|_2$$

where σ is given by

$$\sigma = \max\{|1 - \mu m_f|, |1 - \mu L_f|\} < 1. \tag{22}$$

Proof. Given $z_x := \mathbf{prox}_{\mu f}(x)$ and $z_y := \mathbf{prox}_{\mu f}(y)$, x and y can be computed as follows

$$x = z_x + \mu \nabla f(z_x),\; y = z_y + \mu \nabla f(z_y).$$

Thus,

$$\|R_{\mu f}(x) - R_{\mu f}(y)\|^2 = \|2(z_x - z_y) - (x-y)\|^2 = \|(z_x - z_y) - \mu(\nabla f(z_x) - \nabla f(z_y))\|^2 = \|z_x - z_y\|^2 + \|\mu(\nabla f(z_x) - \nabla f(z_y))\|^2 - 2\mu\langle \nabla f(z_x) - \nabla f(z_y), z_x - z_y\rangle \le \max\{(1-\mu L_f)^2, (1-\mu m_f)^2\}\|z_x - z_y\|^2 \le \sigma\|x-y\|^2.$$

where the firm non-expansiveness of the proximal operator is used in the last step. From the above expressions, we see that $\sigma < 1$ if and only if,

$$-1 < 1 - \mu L_f < 1 \;\text{ and }\; -1 < 1 - \mu m_f < 1.$$

Since $m_f \le L_f$, these conditions hold for $\mu \in (0, 2/L_f)$, which completes the proof.

Lemma 2. Let Assumption 1 hold and let $\mu \in (0, 2/L_f)$. Then, the operator $R_{\mu g}$ is firmly non-expansive.

Proof.

$$\|R_{\mu g}(x) - R_{\mu g}(y)\|_2^2 = 4\|\mathbf{prox}_{\mu f}(x) - \mathbf{prox}_{\mu f}(y)\|_2^2 + \|x-y\|_2^2 - 4\langle x - y, \mathbf{prox}_{\mu f}(x) - \mathbf{prox}_{\mu f}(y)\rangle \le \|x-y\|_2^2.$$

Remark 2. Since $R_{\mu g}$ is firmly non-expansive and $R_{\mu f}$ is σ-contractive, the composite operator $R_{\mu g}R_{\mu f}$ is also σ-contractive. This yields the following lemma.

Lemma 3. Let Assumption 1 hold. Then, for any $z \in \mathbb{R}^n$, $\hat{z} \in \mathbb{R}^n$, $u := [R_{\mu g}R_{\mu f}](z)$, and $\hat{u} := [R_{\mu g}R_{\mu f}](\hat{z})$, the pointwise quadratic inequality

$$\begin{bmatrix} z - \hat{z} \\ u - \hat{u} \end{bmatrix}^T \underbrace{\begin{bmatrix} \sigma^2 I & 0 \\ 0 & -I \end{bmatrix}}_{\Pi} \begin{bmatrix} z - \hat{z} \\ u - \hat{u} \end{bmatrix} \ge 0 \tag{23}$$

holds, where

$$\sigma = \max\{|1 - \mu m_f|, |1 - \mu L_f|\}. \tag{24}$$

We next employ the KYP lemma in the frequency domain (Rantzer, 1996)

$$\begin{bmatrix} H_\rho(j\omega) \\ I \end{bmatrix}^* \Pi \begin{bmatrix} H_\rho(j\omega) \\ I \end{bmatrix} \prec 0, \;\; \forall\,\omega \in \mathbb{R} \tag{25}$$

where Π is given by (23), ω is the temporal frequency, and

$$H_\rho(j\omega) := C\,(j\omega I - (A + \rho I))^{-1}B = \frac{1}{j\omega + 1 - \rho}\,I \tag{26}$$

to establish the global exponential stability of (21).

Theorem 4. Let Assumption 1 hold. Then, the DR splitting dynamics (21) are globally exponentially stable, i.e., there is $c > 0$ and $\rho \in (0, 1-\sigma)$ such that,

$$\|z(t) - z^\star\| \le c\,\mathrm{e}^{-\rho t}\|z(0) - z^\star\|, \;\; \forall\, t \ge 0.$$

Proof. The KYP lemma implies the global exponential stability of (21) if there exists $\rho \in (0, 1)$ such that,

$$\sigma^2 H_\rho(j\omega)^* H_\rho(j\omega) - I \prec 0, \;\; \forall\,\omega \in \mathbb{R}, \tag{27}$$

where $H_\rho(j\omega)$ is given by (26). Inequality (27) is satisfied if,

$$\sigma^2 - (1-\rho)^2 - \omega^2 < 0, \;\; \forall\,\omega \in \mathbb{R}$$

which proves $\rho < 1 - \sigma$.

3.3 Douglas-Rachford splitting on the dual problem

The DR splitting algorithm cannot be used to directly solve a problem with a more general linear constraint,

$$\begin{array}{ll} \underset{x,\,z}{\text{minimize}} & f(x) + g(z) \\ \text{subject to} & Tx + Sz = r \end{array} \tag{28}$$

where $T \in \mathbb{R}^{m\times n}$, $S \in \mathbb{R}^{m\times n}$, and $r \in \mathbb{R}^m$ are problem data. However, it can be utilized to solve the dual problem,

$$\underset{\zeta}{\text{minimize}} \quad f_1(\zeta) + g_1(\zeta) \tag{29}$$

where $f_1(\zeta) = f^\star(-T^T\zeta) + r^T\zeta$, $g_1(\zeta) = g^\star(-S^T\zeta)$, and $h^\star(\zeta) := \sup_x(\zeta^T x - h(x))$ is the conjugate of the function h. It is a standard fact (Eckstein and Bertsekas, 1992; Gabay, 1983) that solving the dual problem (29) via the DR splitting algorithm is equivalent to using ADMM for the original problem (28). Next, we introduce a gradient flow dynamics based on the DR splitting algorithm for solving (29) and demonstrate global exponentially stability under the following assumption.

Assumption 2. Let the differentiable part f of the objective function in (28) be strongly convex with parameter m_f, let ∇f be Lipschitz continuous with parameter L_f, let the function g be proper, lower semicontinuous, and convex, and let the matrix T be full row rank.

The continuous-time DR splitting algorithm

$$\dot{\zeta} = \mathbf{prox}_{\mu g_1}(2\,\mathbf{prox}_{\mu f_1}(\zeta) - \zeta) - \mathbf{prox}_{\mu f_1}(\zeta) \tag{30}$$

can be used to compute the optimal solution $\zeta^\star$ to (29). It is well-known that the conjugate functions are convex Boyd and Vandenberghe (2004). Since (30) is identical to (20), if f_1 satisfies the conditions in Assumption 1 global exponential stability of (30) follows from Theorem 4. For a full row rank T, f_1 is strongly convex and ∇f_1 is Lipschitz continuous with parameters (Giselsson and Boyd, 2017, Proposition 4),

$$L_{f_1} = \|T^T\|^2/m_f, \;\; m_{f_1} = \theta^2/L_f$$

where θ is a positive parameter that always exists and satisfies $\|T^T\nu\| \ge \theta\|\nu\|$ for all $\nu \in \mathbb{R}^m$. Thus, both f_1 and g_1 satisfy Assumption 2 and global exponential stability of (30) follows from Theorem 4.

4. CONCLUDING REMARKS

We study a class of nonsmooth optimization problems in which it is desired to minimize the sum of a continuously differentiable strongly convex function with a Lipschitz continuous gradient and a nondifferentiable proper, closed, and convex function. We employ the theory of IQCs to prove global exponential stability of Douglas-Rachford splitting dynamics. Our ongoing work focuses on establishing stability properties in the absence of strong convexity.

REFERENCES

Arrow, K.J., Hurwicz, L., and Uzawa, H. (1958). Studies in linear and non-linear programming.

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization.* Cambridge University Press.

Cherukuri, A., Mallada, E., and Cortes, J. (2016). Asymptotic convergence of constrained primal–dual dynamics. *Syst. Control Lett.*, 87, 10–15.

Cherukuri, A., Mallada, E., Low, S., and Cortes, J. (2018). The role of convexity on saddle-point dynamics: Lyapunov function and robustness. *IEEE Trans. Automat. Control*, 63(8), 2449–2464.

Dhingra, N.K., Khong, S.Z., and Jovanović, M.R. (2019). The proximal augmented Lagrangian method for nonsmooth composite optimization. *IEEE Trans. Automat. Control*, 64(7), 2861–2868.

Douglas, J. and Rachford, H. (1956). On the numerical solution of heat conduction problems in two and three space variables. *Trans. Amer. Math. Soc.*, 82(2), 421–439.

Eckstein, J. and Bertsekas, D.P. (1992). On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Program.*, 55(1-3), 293–318.

Fazlyab, M., Ribeiro, A., Morari, M., and Preciado, V.M. (2018). Analysis of optimization algorithms via integral quadratic constraints: Nonstrongly convex problems. *SIAM J. Optim.*, 28(3), 2654–2689.

Feijer, D. and Paganini, F. (2010). Stability of primal–dual gradient dynamics and applications to network optimization. *Automatica*, 46(12), 1974–1981.

França, G., Robinson, D., and Vidal, R. (2018). ADMM and accelerated ADMM as continuous dynamical systems. ArXiv:1805.06579.

Gabay, D. (1983). Applications of the method of multipliers to variational inequalities. In *Studies in Mathematics and its Applications*, volume 15, 299–331.

Giselsson, P. and Boyd, S. (2017). Linear convergence and metric selection for Douglas-Rachford splitting and ADMM. *IEEE Trans. Automat. Control*, 62(2), 532–544.

Hassan-Moghaddam, S. and Jovanović, M.R. (2018a). Distributed proximal augmented Lagrangian method for nonsmooth composite optimization. In *Proceedings of the 2018 American Control Conference*, 2047–2052.

Hassan-Moghaddam, S. and Jovanović, M.R. (2018b). On the exponential convergence rate of proximal gradient flow algorithms. In *Proceedings of the 57th IEEE Conference on Decision and Control*, 4246–4251.

Hatanaka, T., Chopra, N., Ishizaki, T., and Li, N. (2018). Passivity-based distributed optimization with communication delays using PI consensus algorithm. *IEEE Trans. Automat. Control*, 63(12), 4421–4428.

Hu, B. and Lessard, L. (2017). Dissipativity theory for Nesterov's accelerated method. In *Proceedings of the 34th International Conference on Machine Learning*, 1549–1557.

Hu, B., Seiler, P., and Rantzer, A. (2017). A unified analysis of stochastic optimization methods using jump system theory and quadratic constraints. In *Proceedings of the 2017 Conference on Learning Theory*, 1157–1189.

Lessard, L., Recht, B., and Packard, A. (2016). Analysis and design of optimization algorithms via integral quadratic constraints. *SIAM J. Optim.*, 26(1), 57–95.

Michalowsky, S., Scherer, C., and Ebenbauer, C. (2019). Robust and structure exploiting optimization algorithms: An integral quadratic constraint approach. ArXiv:1905.00279.

Mohammadi, H., Razaviyayn, M., and Jovanović, M.R. (2018). Variance amplification of accelerated first-order algorithms for strongly convex quadratic optimization problems. In *Proceedings of the 57th IEEE Conference on Decision and Control*, 5753–5758. Miami, FL.

Mohammadi, H., Razaviyayn, M., and Jovanović, M.R. (2019a). Performance of noisy Nesterov's accelerated method for strongly convex optimization problems. In *Proceedings of the 2019 American Control Conference*, 3426–3431. Philadelphia, PA.

Mohammadi, H., Razaviyayn, M., and Jovanović, M.R. (2019b). Robustness of accelerated first-order algorithms for strongly convex optimization problems. *IEEE Trans. Automat. Control*. Submitted; also arXiv:1905.11011.

Muehlebach, M. and Jordan, M.I. (2019). A dynamical systems perspective on Nesterov acceleration. ArXiv:1905.07436.

Parikh, N. and Boyd, S. (2013). Proximal algorithms. *Found. Trends Opt.*, 1(3), 123–231.

Patrinos, P., Stella, L., and Bemporad, A. (2014). Forward-backward truncated Newton methods for convex composite optimization. ArXiv:1402.6655.

Poveda, J.I. and Li, N. (2019). Inducing uniform asymptotic stability in time-varying accelerated optimization dynamics via hybrid regularization. ArXiv:1905.12110.

Qu, G. and Li, N. (2018). On the exponential stability of primal-dual gradient dynamics. *IEEE Control Syst. Lett.*, 3(1), 43–48.

Rantzer, A. (1996). On the Kalman-Yakubovich-Popov lemma. *Syst. Control Lett.*, 28(1), 7–10.

Schropp, J. and Singer, I. (2000). A dynamical systems approach to constrained minimization. *Numer. Func. Anal. Opt.*, 21(3-4), 537–551.

Seidman, J., Fazlyab, M., Preciado, V., and Pappas, G. (2019). A control-theoretic approach to analysis and parameter selection of Douglas-Rachford splitting. ArXiv:1903.11525.

Shi, B., Du, S., Jordan, M.I., and Su, W. (2018). Understanding the acceleration phenomenon via high-resolution differential equations. ArXiv:1810.08907.

Stella, L., Themelis, A., and Patrinos, P. (2017). Forward–backward quasi-Newton methods for nonsmooth optimization problems. *Comput. Optim. Appl.*, 67(3), 443–487.

Su, W., Boyd, S., and Candes, E. (2016). A differential equation for modeling Nesterov's accelerated gradient method: Theory and insights. *J. Mach. Learn. Res.*, 17, 1–43.

Themelis, A. and Patrinos, P. (2017). Douglas-Rachford splitting and ADMM for nonconvex optimization: tight convergence results. ArXiv:1709.05747.

Themelis, A., Stella, L., and Patrinos, P. (2018). Forward-backward envelope for the sum of two nonconvex functions: Further properties and nonmonotone line-search algorithms. *SIAM J. Optim.*, 28(3), 2274–2303.

Wang, J. and Elia, N. (2011). A control perspective for centralized and distributed convex optimization. In *Proceedings of the 50th IEEE Conference on Decision and Control*, 3800–3805.

Wibisono, A., Wilson, A.C., and Jordan, M.I. (2016). A variational perspective on accelerated methods in optimization. *Proc. Natl. Acad. Sci.*, 113(47), E7351–E7358.

Zhang, J., Mokhtari, A., Sra, S., and Jadbabaie, A. (2018). Direct Runge-Kutta discretization achieves acceleration. In *Advances in Neural Information Processing Systems*, 3900–3909.