ESAIM: COCV 25 (2019) 57
https://doi.org/10.1051/cocv/2018063

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

A WEAKLY CONVERGENT FULLY INEXACT DOUGLAS-RACHFORD METHOD WITH RELATIVE ERROR TOLERANCE*

Benar Fux Svaiter**

Abstract. Douglas-Rachford method is a splitting algorithm for finding a zero of the sum of two maximal monotone operators. Each of its iterations requires the sequential solution of two proximal subproblems. The aim of this work is to present a fully inexact version of Douglas-Rachford method wherein both proximal subproblems are solved approximately within a relative error tolerance. We also present a semi-inexact variant in which the first subproblem is solved exactly and the second one inexactly. We prove that both methods generate sequences weakly convergent to the solution of the underlying inclusion problem, if any.

Mathematics Subject Classification. 49M27, 47H05, 65G99, 65K05, 49J45.

Received September 8, 2018. Accepted November 5, 2018.

1. Introduction

Douglas-Rachford method [6], originally proposed for solving discretized heat equations, was extended by Lions and Mercier [17] for finding a zero of the sum of two maximal monotone operators. This method is, presently, the subject of intense research due to its efficiency, its use for solving PDEs, large-scale optimization problems (even some non-convex ones), imaging problems, and its connection with the alternating direction method (see [2–4, 11–16] and the references therein).

Eckstein and Bertseka [7] proved that Douglas-Rachford method can be regarded as an instance of the proximal point method applied to an implicitly defined operator. Recently, Eckstein and Yao [9] cleverly used this result to propose an inexact version of Douglas-Rachford method with relative error tolerance based on Solodov and Svaiter hybrid proximal-extragradient method [20, 21, 23, 24]. Each iteration of Douglas-Rachford method requires the sequential solution of two proximal subproblems. Eckstein-Yao algorithm is a semi-inexact version of this method in the sense that it allows for inexactness on the first proximal subproblem and requires the second one to be solved exactly. Complexity of Eckstein-Yao inexact Douglas-Rachford method was derived by Marques Alves and Geremia [1].

*Partially supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico CNPq grant 306247/2015-1 and Fundação Carlos Chagas Filho de Amparo á Pesquisa do Estado do Rio de Janeiro FAPERJ grant Cientistas de Nosso Estado E-26/203.318/2017.

Keywords and phrases: Douglas - Rachford method, relative error, weak convergence, splitting.

IMPA, Estrada Dona Castorina 110, 22460-320 Rio de Janeiro, Brazil.

** Corresponding author: benar@impa.br

Article published by EDP Sciences

The main contribution of this work is the introduction and analysis of a fully inexact version of Douglas-Rachford method wherein both proximal subproblems are solved approximately within a relative error tolerance. To the best of our knowledge, this is the first fully inexact version of Douglas-Rachford method with a relative error tolerance. We also propose a semi-inexact version for the cases where one of the proximal subproblems can be solved exactly. In the semi-inexact case of Douglas-Rachford, our proposal is to solve the first subproblem exactly and the second one inexactly, so that the error criterion to be satisfied is immediately available (during the computation of the approximate solution of the second subproblem).

Consider the monotone inclusion problem

$$0 \in A(x) + B(x),$$

where A and B are maximal monotone operators in a Hilbert space H. Douglas-Rachford method, originally proposed for solving linear monotone problems [6], was extended by Lions and Mercier [17] for solving this inclusion problem by means of the iteration

$$v_{k+1} = J_{\lambda A}(2J_{\lambda B} - I)v_k + (I - J_{\lambda B})v_k, \qquad (k = 1, 2, \dots),$$

where $\lambda > 0$ and $J_{\lambda A} = (I + \lambda A)^{-1}$ and $J_{\lambda B} = (I + \lambda B)^{-1}$ are the *proximal operators* of λA and λB, respectively. It follows from Minty's theorem [18] that these operators are point-to-point and their domain is the whole H. Defining

$$x_{k-1} = J_{\lambda B}(v_k),\ b_{k-1} = \frac{v_k - x_{k-1}}{\lambda},\ y_k = J_{\lambda A}(x_{k-1} - \lambda b_{k-1}),\ a_k = \frac{x_{k-1} - \lambda b_{k-1} - y_k}{\lambda},$$

it is easy to verify that Douglas-Rachford method can be written as: choose $x_0 \in H$, $b_0 \in B(x_0)$ and for $k = 1, 2, \dots$

$$\begin{aligned} &\text{compute } y_k,\ a_k \in H \ \text{ such that } \quad a_k \in A(y_k), \quad \lambda a_k + y_k = x_{k-1} - \lambda b_{k-1},\\ &\text{compute } x_k,\ b_k \in H \ \text{ such that } \quad b_k \in B(x_k), \quad \lambda b_k + x_k = y_k + \lambda b_{k-1}. \end{aligned}$$

Since $x_k = x_{k-1} - \lambda(a_k + b_k)$ and $b_k = b_{k-1} - \lambda^{-1}(x_k - y_k)$, we propose the following (fully) inexact version of Douglas-Rachford method: choose $z_0, w_0 \in H$ and for $k = 1, 2, \dots$

$$\begin{aligned} &\text{compute } y_k,\ a_k \in H \text{ and } x_k,\ b_k \in H \text{ such that:}\\ &y_k, a_k \ \text{ is an approximate solution of } \quad a \in A(y), \quad \lambda a + y = z_{k-1} - \lambda w_{k-1},\\ &x_k, b_k \ \text{ is an approximate solution of } \quad b \in B(x), \quad \lambda b + x = y_k + \lambda w_{k-1},\\ &\text{set } z_k = z_{k-1} - (1 - t_k)\lambda(a_k + b_k),\ w_k = w_{k-1} - (1 - t_k)\lambda^{-1}(x_k - y_k), \end{aligned}$$

where t_k is an under-relaxation parameter which is zero whenever it is "safe" to do so (more of which latter). If both subproblems are exactly solved and $t_k = 1$, at the kth iteration of this generic method, then $z_k = x_k$ and $w_k = b_k$. Following [20, 24], we will allow errors in the inclusions *and* in the equations in the computation of y_k, a_k and x_k, b_k. Errors in the inclusion will be considered using the ε-enlargement [5] of maximal monotone operators.

Since our semi-inexact method is related to [9], it is worth discussing their differences. Eckstein-Yao semi-inexact version of Douglas-Rachford method [9] does not require the introduction of a relaxation factor, as ours do, so that their inexact version is formally closer than ours to the exact method. Their version allows for errors in the first subproblem, while our semi-inexact version allows for errors in the second one. Weak convergence, in infinite dimensional spaces, of their version is an interesting open question, while we prove here weak convergence of our version. Their version has a very good practical performance ([9], Sect. 7) and nice complexity properties [1], while the practical performance and complexity properties of our version are, as of now, open questions. Weak convergence on Douglas-Rachford method was established by the author in [26] for the inexact version with the summable error tolerance. Here, we use the techniques and ideas of that work to prove weak convergence under a relative error tolerance.

This work is organized as follows. In Section 2, we establish the notation and prove some basic results. In Section 3, we present a fully inexact Douglas-Rachford method and analyze its convergence properties. In Section 4, we present a semi-inexact Douglas-Rachford method and analyze its convergence properties. In Appendices A and B, we prove some technical results, complementary proofs are provided in Appendix C.

2. Basic definitions and results

From now on H is a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\|x\| = \sqrt{\langle x, x\rangle}$. In $H \times H$ we consider the canonical inner product and associated norm of Hilbert spaces products.

We are concerned with the inclusion problem

$$0 \in A(x) + B(x), \tag{2.1}$$

where $A : H \rightrightarrows H$ and $B : H \rightrightarrows H$ are maximal monotone operators. The *extended solution set* [8] of this problem is

$$S_e(A, B) = \{(z, w) \ : \ w \in B(z),\ -w \in A(z)\}. \tag{2.2}$$

It is trivial to verify that

$$0 \in A(z) + B(z) \iff (z, w) \in S_e(A, B) \text{ for some } w \in H.$$

The ε-enlargement [5] of a maximal monotone operator $T : H \rightrightarrows H$ is defined as

$$T^{[\varepsilon]}(z) = \{v \in H \ : \ \langle z - z', v - v'\rangle \geq -\varepsilon \ \ \forall z' \in H,\ v' \in T(z')\} \qquad (\varepsilon \geq 0, \ \ z \in H). \tag{2.3}$$

The properties of the ε-enlargement below follows trivially from the above definition.

Proposition 2.1. *If $T : H \rightrightarrows H$ is maximal monotone, then*

1. *$T^{[\varepsilon=0]} = T$;*
2. *$T^{[\varepsilon_1]}(x) \subseteq T^{[\varepsilon_2]}(x)$ for any $0 \leq \varepsilon_1 \leq \varepsilon_2$ and $x \in H$.*

In each iteration of Douglas-Rachford method, one shall compute $(I + \lambda T)^{-1}\zeta$, first with $T = A$ and $\zeta = x_{k-1} - \lambda b_{k-1}$ and then with $T = B$ and $\zeta = y_k + \lambda b_{k-1}$ (or vice versa, with the suitable change of indexes). Let T be a maximal monotone operator in H. Computation of $z = (I + \lambda T)^{-1}\zeta$ can be decoupled in an inclusion and an equation, which we call the *proximal inclusion-equation system*:

$$v \in T(z), \qquad \lambda v + z - \zeta = 0. \tag{2.4}$$

If we allow errors in the inclusion, by means of the ε-enlargement of T, and errors in the equation we get

$$v \in T^{[\varepsilon]}(z), \qquad \lambda v + z - \zeta = r,$$

where ε is the *error in the inclusion* and r is the *residual in the equation* at (2.4). In some sense, $\|r\|^2 + 2\lambda\varepsilon$ quantifies the overall error in the inexact solution of (2.4), due to the next result proved in [22].

Proposition 2.2 ([22], Cor. 1)**.** *Suppose $T : H \rightrightarrows H$ is maximal monotone, $\lambda > 0$, and $\zeta \in H$. If*

$$v^* \in T(z^*), \quad \lambda v^* + z^* - \zeta = 0, \qquad v \in T^{[\varepsilon]}(z), \quad \lambda v + z - \zeta = r$$

then $\|\lambda(v^* - v)\|^2 + \|z^* - z\|^2 \leq \|r\|^2 + 2\lambda\varepsilon$.

The next lemma will be instrumental in the convergence analysis of the inexact methods we propose in this work.

Lemma 2.3. *Let $z, w \in X$ and $\lambda > 0$. Suppose*

$$\begin{aligned} \lambda a + y &= z - \lambda w + r, \quad a \in A^{[\varepsilon]}(y), \\ \lambda b + x &= y + \lambda w + s, \quad b \in B^{[\mu]}(x), \end{aligned}$$

and define $z' = z - \lambda(a + b)$, $w' = w - \lambda^{-1}(x - y)$. For any $(z^*, w^*) \in S_e(A, B)$,

$$\|(z^*, \lambda w^*) - (z, \lambda w)\|^2 \geq \|(z^*, \lambda w^*) - (z', \lambda w')\|^2 + \|\lambda(a + w)\|^2 - [\|s\|^2 + 2\lambda(\langle r, a + b\rangle + \varepsilon + \mu)].$$

Proof. Fix $(z^*, w^*) \in S_e(A, B)$ and let

$$\pi = \langle z^* - z', z' - z\rangle + \lambda^2\langle w^* - w', w' - w\rangle.$$

In view of definition (2.2), $w^* \in B(z^*)$ and $-w^* \in A(z^*)$. It follows from these inclusions, the inclusions $a \in A^{[\varepsilon]}(y)$ and $b \in B^{[\mu]}(x)$, and definition (2.3), that

$$\begin{aligned} \langle z^* - z', z' - z\rangle &= \lambda\langle z^* - z', w^* - b\rangle + \lambda\langle z^* - z', -w^* - a\rangle \\ &= \lambda[\langle z^* - x, w^* - b\rangle + \langle x - z', w^* - b\rangle + \langle z^* - y, -w^* - a\rangle + \langle y - z', -w^* - a\rangle] \\ &\geq \lambda[\langle x - z', w^* - b\rangle + \langle y - z', -w^* - a\rangle - \varepsilon - \mu]\,. \end{aligned}$$

Direct combination of this inequality with the definitions of π and w', and direct algebraic manipulations yields

$$\begin{aligned} \pi &\geq \lambda[\langle x - z', w^* - b\rangle + \langle y - z', -w^* - a\rangle - \varepsilon - \mu + \langle w^* - w', y - x\rangle] \\ &= \lambda[\langle x - z', w^* - b\rangle + \langle y - z', (b - w^*) - b - a\rangle - \varepsilon - \mu + \langle (w^* - b) + b - w', y - x\rangle] \\ &= \lambda[\langle y - z', -b - a\rangle + \langle b - w', y - x\rangle - \varepsilon - \mu] \\ &= \lambda\langle y - x + r + s, -a - b\rangle + \langle s, y - x\rangle - \lambda(\varepsilon + \mu) \\ &= \lambda\langle x - y, a + b\rangle - [\lambda\langle r, a + b\rangle + \langle s, \lambda(a + b) + x - y\rangle + \lambda(\varepsilon + \mu)]. \end{aligned}$$

Therefore,

$$\begin{aligned} \|(z^*, \lambda w^*) - (z, \lambda w)\|^2 &= \|(z^*, \lambda w^*) - (z', \lambda w')\|^2 + \|(z', \lambda w') - (z, \lambda w)\|^2 + 2\pi \\ &\geq \|(z^*, \lambda w^*) - (z', \lambda w')\|^2 + \lambda^2\|a + b\|^2 + \|x - y\|^2 \\ &\quad + 2\lambda\langle x - y, a + b\rangle - 2[\lambda\langle r, a + b\rangle + \langle s, \lambda(a + b) + x - y\rangle + \lambda(\varepsilon + \mu)] \\ &= \|(z^*, w^*) - (z', w')\|^2 + \|\lambda(a + b) + x - y\|^2 \\ &\quad - 2[\lambda\langle r, a + b\rangle + \langle s, \lambda(a + b) + y - x\rangle + \lambda(\varepsilon + \mu)]. \end{aligned}$$

To end the proof, observe that

$$\|\lambda(a + b) + x - y\|^2 - 2\langle s, \lambda(a + b) + x - y\rangle = \|\lambda(a + b) + x - y - s\|^2 - \|s\|^2 = \|\lambda(a + w)\|^2 - \|s\|^2,$$

and combine the two above equations. $\square$

3. A fully inexact Douglas-Rachford method with relative error tolerance

In this section we present an inexact Douglas-Rachford method wherein both proximal subproblems are to be solved within a relative error tolerance. We prove that sequences generated by this method converge weakly to a point in $S_e(A,B)$, whenever the solution set of (2.1) is non-empty.

Algorithm I: Inexact DR method with relative error tolerance.

(0) Let $z_0, w_0 \in H$, $\lambda > 0$, $0 < \sigma < \nu < 1$. For $k = 1, 2, 3, \ldots,$
(1) Find $a_k, y_k, \varepsilon_k, b_k, x_k, \mu_k$ such that

$$
\begin{aligned}
&a_k \in A^{[\varepsilon_k]}(y_k), \quad b_k \in B^{[\mu_k]}(x_k),\\
&\|\lambda a_k + y_k - (z_{k-1} - \lambda w_{k-1})\|^2 + \|\lambda b_k + x_k - (y_k + \lambda w_{k-1})\|^2\\
&\;+ 2\lambda(\varepsilon_k + \mu_k) \le \frac{\sigma^2}{4}(\lambda^2\|a_k + b_k\|^2 + \|x_k - y_k\|^2);
\end{aligned}
\tag{3.1}
$$

(2) set

$$
\begin{aligned}
&\delta_k = \|\lambda a_k + y_k - (z_{k-1} - \lambda w_{k-1})\|^2 + \|\lambda b_k + x_k - (y_k + \lambda w_{k-1})\|^2 + 2\lambda(\varepsilon_k + \mu_k),\\
&\rho_k = \|\lambda(a_k + b_k)\|^2 + \|x_k - y_k\|^2,\\
&t_k = \begin{cases} 0, & \text{if } a_k + b_k = x_k - y_k = 0\\ \nu \max\left\{0, \sqrt{\dfrac{4\delta_k}{\sigma^2\rho_k}} - \dfrac{\|\lambda(a_k + w_{k-1})\|^2}{\rho_k}\right\}, & \text{otherwise,} \end{cases}\\
&z_k = z_{k-1} - (1 - t_k)\lambda(a_k + b_k), \quad w_k = w_{k-1} - (1 - t_k)\lambda^{-1}(x_k - y_k).
\end{aligned}
\tag{3.2}
$$

We did not specify how to compute a_k, y_k, ε_k, b_k, x_k, and μ_k, which adds generality to the method. In Appendix A we show that under some mild conditions (on A and B) step (1) is computable. Whenever $t_k = 0$, $z_k = z_{k-1} - \lambda(a_k + b_k)$ and $w_k = w_{k-1} - \lambda^{-1}(x_k - y_k)$, so that formally we retrieve a Douglas-Rachford iteration. In those iterations where $t_k \ne 0$, we will have an under-relaxed Douglas-Rachford iteration.

To simplify the convergence analysis, let r_k and s_k denote the residuals in the equations of the proximal inclusion-equations systems to be solved for A and B at the kth iteration, that is,

$$
r_k = \lambda a_k + y_k - (z_{k-1} - \lambda w_{k-1}), \quad s_k = \lambda b_k + x_k - (y_k + \lambda w_{k-1}), \qquad (k = 1, 2, \ldots). \tag{3.3}
$$

With this notation, (3.1) and the first line of (3.2) writes

$$
\begin{aligned}
&a_k \in A^{[\varepsilon_k]}(y_k), \quad \lambda a_k + y_k = z_{k-1} - \lambda w_{k-1} + r_k,\\
&b_k \in B^{[\mu_k]}(x_k), \quad \lambda b_k + x_k = y_k + \lambda w_{k-1} + s_k,\\
&\delta_k = \|r_k\|^2 + \|s_k\|^2 + 2\lambda(\varepsilon_k + \mu_k), \qquad \delta_k \le \frac{\sigma^2}{4}\rho_k.
\end{aligned}
\tag{3.4}
$$

It follows from the definition of t_k at (3.2) that for all k,

$$
0 \le t_k \le \nu < 1, \qquad \frac{\nu}{\sigma}\sqrt{4\delta_k\rho_k} - \nu\|\lambda(a_k + w_{k-1})\|^2 \le t_k\rho_k, \qquad t_k^2\rho_k \le \frac{4\nu^2}{\sigma^2}\delta_k. \tag{3.5}
$$

From now on in this section,

$$p_k = (z_k, \lambda w_k), \qquad (k = 0, 1, 2, \dots). \tag{3.6}$$

First we will prove that the sequence $\{p_k\}$ converges Féjer to the set of points $(z, \lambda w)$ where $(z, w) \in S_e(A, B)$.

Lemma 3.1. *For any $(z, w) \in S_e(A, B)$ and all k*

$$\|(z, \lambda w) - p_{k-1}\|^2 \geq \|(z, \lambda w) - p_k\|^2 + (1 - t_k)\left\{\frac{\nu - \sigma}{\sigma}\sqrt{4\delta_k\rho_k} + (1 - \nu)\|\lambda(a_k + w_{k-1})\|^2\right\},$$

$$\|(z, \lambda w) - p_0\|^2 \geq \|(z, \lambda w) - p_k\|^2 + \sum_{i=1}^{k}(1 - t_i)\left\{\frac{\nu - \sigma}{\sigma}\sqrt{4\delta_i\rho_i} + (1 - \nu)\|\lambda(a_i + w_{i-1})\|^2\right\}.$$

Proof. Fix $(z, w) \in S_e(A, B)$ and let $p^* = (z, \lambda w)$. Define, for $k = 1, 2, \dots$,

$$z'_k = z_{k-1} - \lambda(a_k + b_k), \quad w'_k = w_{k-1} - \lambda^{-1}(x_k - y_k), \quad p'_k = (z'_k, \lambda w'_k). \tag{3.7}$$

It follows from (3.6), (3.4) and Lemma 2.3 that

$$\|p^* - p_{k-1}\|^2 \geq \|p^* - p'_k\|^2 + \|\lambda(a_k + w_{k-1})\|^2 - [\|s_k\|^2 + 2\lambda(\langle r_k, a_k + b_k\rangle + \varepsilon_k + \mu_k)].$$

Using Cauchy-Schwarz inequality, the definition of ρ_k in (3.2), and the third line on (3.4) we conclude that

$$\|s_k\|^2 + 2\lambda(\langle r_k, a_k + b_k\rangle + \varepsilon_k + \mu_k) \leq \|s_k\|^2 + 2\lambda(\varepsilon_k + \mu_k) + 2\|r_k\|\sqrt{\rho_k} = \delta_k - \|r_k\|^2 + 2\|r_k\|\sqrt{\rho_k}$$

and $\|r_k\| \leq \sqrt{\delta_k} \leq \sigma\sqrt{\rho_k}/2$. Since the expression on the right-hand side of the above equality is increasing for $\|r_k\| \leq \sqrt{\rho_k}$, its maximum value on $[0, \sqrt{\delta_k}]$ is attained at $\|r_k\| = \sqrt{\delta_k}$. Combining these observations with the two above inequalities we conclude that

$$\|p^* - p_{k-1}\|^2 \geq \|p^* - p'_k\|^2 + \|\lambda(a_k + w_{k-1})\|^2 - 2\sqrt{\delta_k\rho_k}.$$

Since $p_k = t_k p_{k-1} + (1 - t_k)p'_k$ and $\|p_{k-1} - p'_k\|^2 = \rho_k$,

$$\begin{aligned}
\|p^* - p_{k-1}\|^2 &= t_k\|p^* - p_{k-1}\|^2 + (1 - t_k)\|p^* - p_{k-1}\|^2 \\
&\geq t_k\|p^* - p_{k-1}\|^2 + (1 - t_k)\left[\|p^* - p'_k\|^2 + \|\lambda(a_k + w_{k-1})\|^2 - 2\sqrt{\delta_k\rho_k}\right] \\
&= \|t_k(p^* - p_{k-1}) + (1 - t_k)(p^* - p'_k)\|^2 + (1 - t_k)t_k\|p_{k-1} - p'_k\|^2 \\
&\quad + (1 - t_k)\left[\|\lambda(a_k + w_{k-1})\|^2 - 2\sqrt{\delta_k\rho_k}\right] \\
&= \|p^* - p_k\|^2 + (1 - t_k)\left[t_k\rho_k + \|\lambda(a_k + w_{k-1})\|^2 - 2\sqrt{\delta_k\rho_k}\right].
\end{aligned}$$

To end the proof of the first inequality, use the above inequality and the next to last inequality in (3.5). The second inequality of the lemma follows trivially from the first one. ☐

Lemma 3.2. *If $S_e(A, B)$ is nonempty, then*

1. *$\{z_k\}$ and $\{w_k\}$ are bounded;*
2. *$r_k \to 0$, $s_k \to 0$, $\varepsilon_k \to 0$ and $\mu_k \to 0$ as $k \to \infty$;*
3. *$a_k + w_{k-1} \to 0$ and $y_k - z_{k-1} \to 0$ as $k \to \infty$;*
4. *$b_k - w_k \to 0$ and $x_k - z_k \to 0$ as $k \to \infty$.*

Proof. Take (z, w) in $S_e(A, B)$. It follows from Lemma 3.1 that $\{(z_k, \lambda w_k)\}$ is bounded, which proves item 1, and that

$$\sum_{k=1}^{\infty} \sqrt{\delta_k \rho_k} < \infty, \qquad \sum_{k=1}^{\infty} \|a_k + w_{k-1}\|^2 < \infty.$$

Since $\delta_k \leq \sigma^2 \rho_k / 4$, $\sqrt{\delta_k \rho_k} \geq 2\delta_k/\sigma$ and it follows from the first above inequality that $\delta_k \to 0$ as $k \to \infty$. This result, together with the third line of (3.4) proves item 2. The first limit in item 3 follows trivially from the second above inequality while the second limit follows from the first one, the first limit in item 2 and the definition of r_k in (3.3).

It follows from the update formula for z_k and w_k, and from (3.4) that

$$z_k - x_k + r_k + s_k = t_k \lambda (a_k + b_k), \qquad \lambda(w_k - b_k) + s_k = t_k(x_k - y_k).$$

Squaring both sides of each one of the above equations and adding them we conclude that

$$\|z_k - x_k + r_k + s_k\|^2 + \|\lambda(w_k - b_k) + s_k\|^2 = t_k^2 \rho_k.$$

Since $\delta_k \to 0$ as $k \to \infty$, it follows from the above equations and the last inequality in (3.5) that

$$t_k^2 \rho_k \to 0, \quad z_k - x_k + r_k + s_k \to 0, \quad \lambda(w_k - b_k) + s_k \to 0 \qquad \text{as } k \to \infty.$$

Item 4 follows from the above result and from item 2. □

Theorem 3.3. *If $S_e(A, B) \neq \emptyset$, then $\{(x_k, b_k)\}$, $\{(y_k, -a_k)\}$ and $\{(z_k, w_k)\}$ converge weakly to a point in this set.*

Proof. Suppose a subsequence $\{(z_{k_n}, w_{k_n})\}$ converges weakly to some (z, w). It follows from Lemma 3.2 that

$$\begin{aligned}
&x_{k_n} \xrightarrow{w} z, \quad y_{k_n+1} \xrightarrow{w} z, \qquad x_{k_n} - y_{k_n+1} \to 0,\\
&b_{k_n} \xrightarrow{w} w, \quad a_{k_n+1} \xrightarrow{w} -w, \quad a_{k_n+1} + b_{k_n} \to 0,\\
&\mu_{k_n} \to 0, \quad \varepsilon_{k_n+1} \to 0, \qquad\qquad \text{as } k \to \infty.
\end{aligned}$$

Since $b_{k_n} \in B^{[\mu_{k_n}]}(x_{k_n})$, $a_{k_n+1} \in A^{[\varepsilon_{k_n+1}]}(y_{k_n+1})$ for $n = 1, 2, \ldots$, it follows from Lemma B.2 that $w \in B(z)$ and $-w \in A(z)$. Therefore, $(z, w) \in S_e(A, B)$.

We have proved that all weak limit points of $\{(z_k, w_k)\}$ are in $S_e(A, B)$. Therefore, all weak limit points of $\{p_k = (z_k, \lambda w_k)\}$ are in Ω,

$$\Omega = \{(z, \lambda w) \ : \ (z, w) \in S_e(A, B)\}.$$

In Lemma 3.1, we proved that $\{p_k\}$ is Fejér convergent to Ω. Since $\Omega \neq \emptyset$, it follows from these results and from Opial's Lemma [19] that the bounded sequence $\{p_k\}$ has a unique weak limit point and such a point belongs to Ω. Therefore, $\{p_k\}$ converges weakly to a point in $(z, \lambda w)$ where $(z, w) \in S_e(A, B)$. To end the proof, use items 3 and 4 of Lemma 3.2. □

4. A semi-inexact Douglas-Rachford method

In this section we present an inexact Douglas-Rachford method wherein, in each iteration, the first proximal subproblem is to be exactly solved and the second proximal subproblem is to be solved within a relative error tolerance.

A possible advantage of solving the second subproblem approximately, instead of the first one, is that the error criterion to be satisfied is readily available during the computation of the second step, thereby obviating the necessity of computing more than one approximate solution per iteration. Since one of the subproblems is to be solved exactly, following [9], we call this method semi-inexact.

The semi-inexact Douglas-Rachford method to be proposed in this section is quite similar to Algorithm I. We opted to derive a new method, Algorithm II, instead of using Algorithm I with $\varepsilon_k = 0$ and $r_k = 0$ (see Eq. (3.4)) because with a specific analysis of the semi-inexact case we obtain an algorithm with the following features, as compared with Algorithm I:

1. It uses a better (greater) relative error tolerance, in fact by a factor of 4 (*cf.* the error tolerances at (3.1) and (4.1a)).
2. It uses smaller under-relaxation factors, which may be in principle quadratically smaller that the ones used in Algorithm I (*cf.* the definitions of t_k in (3.2) and (4.2)).

Algorithm II: A semi-inexact Douglas-Rachford method with relative error tolerance.

(0) Let $z_0, w_0 \in H$, $\lambda > 0$, $0 < \sigma < \nu < 1$. For $k = 1, 2, 3, \ldots,$

(1) compute a_k, y_k such that

$$a_k \in A(y_k), \quad \lambda a_k + y_k = z_{k-1} - \lambda w_{k-1}, \tag{4.1a}$$

compute b_k, x_k, μ_k such that

$$b_k \in B^{[\mu_k]}(x_k),\ \|\lambda b_k + x_k - (y_k + \lambda w_{k-1})\|^2 + 2\lambda\mu_k \le \sigma^2(\lambda^2\|a_k + b_k\|^2 + \|x_k - y_k\|^2); \tag{4.1b}$$

(2) set

$$\begin{aligned}
&\delta_k = \|\lambda b_k + x_k - (y_k + \lambda w_{k-1})\|^2 + 2\lambda\mu_k,\\
&\rho_k = \|\lambda(a_k + b_k)\|^2 + \|x_k - y_k\|^2,\\
&t_k = \begin{cases} 0 & \text{if } a_k + b_k = x_k - y_k = 0 \\ \nu^2 \max\left\{0, \dfrac{\delta_k}{\sigma^2\rho_k} - \dfrac{\|\lambda(a_k + w_{k-1})\|^2}{\rho_k}\right\} & \text{otherwise,} \end{cases}\\
&z_k = z_{k-1} - (1 - t_k)\lambda(a_k + b_k), \quad w_k = w_{k-1} - (1 - t_k)\lambda^{-1}(x_k - y_k).
\end{aligned} \tag{4.2}$$

From now on in this section, $\{z_k\}$, $\{w_k\}$, $\{a_k\}$ etc. are sequences generated by Algorithm II. To simplify the convergence analysis, let s_k denote the residuals in the equation of the inclusion-equation system to be solved for B at the kth iteration, that is,

$$s_k = \lambda b_k + x_k - (y_k + \lambda w_{k-1}), \qquad (k = 1, 2, \ldots). \tag{4.3}$$

With this notation, (4.1a), (4.1b) and the first line of (4.2) writes

$$\begin{aligned}
&a_k \in A(y_k), && \lambda a_k + y_k = z_{k-1} - \lambda w_{k-1},\\
&b_k \in B^{[\mu_k]}(x_k), && \lambda b_k + x_k = y_k + \lambda w_{k-1} + s_k,\\
&\delta_k = \|s_k\|^2 + 2\lambda\mu_k, && \delta_k \le \sigma^2\rho_k.
\end{aligned} \tag{4.4}$$

It follows from the definition of t_k at (4.2) that for all k,

$$0 \le t_k \le \nu^2 < 1, \qquad \frac{\nu^2}{\sigma^2}\delta_k - \nu^2\|\lambda(a_k + w_{k-1})\|^2 \le t_k\rho_k, \qquad t_k^2\rho_k \le \frac{\nu^4}{\sigma^2}\delta_k. \tag{4.5}$$

From now on in this section,

$$p_k = (z_k, \lambda w_k), \qquad (k = 0, 1, 2, \dots). \tag{4.6}$$

We will prove that the sequence $\{p_k\}$ converges to a point $(z, \lambda w)$ where $(z, w) \in S_e(A, B)$.

Lemma 4.1. *For any $(z^*, w^*) \in S_e(A, B)$ and all k*

$$\|(z^*, \lambda w^*) - p_{k-1}\|^2 \ge \|(z^*, \lambda w^*) - p_k\|^2 + (1 - t_k)\left\{\frac{\nu^2 - \sigma^2}{\sigma^2}\delta_k + (1 - \nu^2)\|\lambda(a_k + w_{k-1})\|^2\right\},$$

$$\|(z^*, \lambda w^*) - p_0\|^2 \ge \|(z^*, \lambda w^*) - p_k\|^2 + \sum_{i=1}^{k}(1 - t_i)\left\{\frac{\nu^2 - \sigma^2}{\sigma^2}\delta_i + (1 - \nu^2)\|\lambda(a_i + w_{i-1})\|^2\right\}.$$

Proof. Fix $(z^*, w^*) \in S_e(A, B)$ and let $p^* = (z^*, \lambda w^*)$. Define, for $k = 1, 2, \dots$,

$$z'_k = z_{k-1} - \lambda(a_k + b_k), \quad w'_k = w_{k-1} - \lambda^{-1}(x_k - y_k), \quad p'_k = (z'_k, \lambda w'_k). \tag{4.7}$$

It follows from (4.6), (4.4) and Lemma 2.3 that

$$\begin{aligned}\|p^* - p_{k-1}\|^2 &\ge \|p^* - p'_k\|^2 + \|\lambda(a_k + w_{k-1})\|^2 - [\|s_k\|^2 + 2\lambda\mu_k] \\ &\ge \|p^* - p'_k\|^2 + \|\lambda(a_k + w_{k-1})\|^2 - \sigma^2\rho_k.\end{aligned}$$

Since $p_k = t_k p_{k-1} + (1 - t_k)p'_k$ and $\|p_{k-1} - p'_k\|^2 = \rho_k$,

$$\begin{aligned}\|p^* - p_{k-1}\|^2 &= t_k\|p^* - p_{k-1}\|^2 + (1 - t_k)\|p^* - p_{k-1}\|^2 \\ &\ge t_k\|p^* - p_{k-1}\|^2 + (1 - t_k)\left[\|p^* - p'_k\|^2 + \|\lambda(a_k + w_{k-1})\|^2 - \delta_k\right] \\ &= \|t_k(p^* - p_{k-1}) + (1 - t_k)(p^* - p'_k)\|^2 + (1 - t_k)t_k\|p_{k-1} - p'_k\|^2 \\ &\quad + (1 - t_k)\left[\|\lambda(a_k + w_{k-1})\|^2 - \delta_k\right] \\ &= \|p^* - p_k\|^2 + (1 - t_k)\left[t_k\rho_k + \|\lambda(a_k + w_{k-1})\|^2 - \delta_k\right].\end{aligned}$$

To end the proof of the first inequality, use the above inequality and the next to last inequality in (4.5).

The second inequality of the lemma follows trivially from the first one. □

The proofs of next lemma and theorem are similar to the ones of Lemma 3.2 and Theorem 3.3. For this reason, these proofs are supplied in Appendix C.

Lemma 4.2. *If $S_e(A, B)$ is nonempty, then*

1. *$\{z_k\}$ and $\{w_k\}$ are bounded;*
2. *$s_k \to 0$ and $\mu_k \to 0$ as $k \to \infty$;*
3. *$a_k + w_{k-1} \to 0$ and $y_k - z_{k-1} \to 0$ as $k \to \infty$;*
4. *$b_k - w_k \to 0$ and $x_k - z_k \to 0$ as $k \to \infty$.*

Theorem 4.3. *If $S_e(A, B) \ne \emptyset$, then $\{(x_k, b_k)\}$, $\{(y_k, -a_k)\}$ and $\{(z_k, w_k)\}$ converge weakly to a point in this set.*

Appendix A. Computability of step (1)

Proposition A.1. *Suppose that $T = A$ and $T = B$ have the following properties*

1. *for any $v, z \in H$, one can verify whether $v \in T(z)$ or $v \notin T(z)$;*
2. *for any $c \in H$, one can generate sequences v_i, z_i, η_i such that $v_i \in T^{[\eta_i]}(z_i)$ for all i and*

$$\eta_i \to 0, \ \|\lambda v_i + z_i - c\| \to 0 \qquad \textit{as } i \to \infty. \tag{A.1}$$

Then, step (1) of Algorithms I and II are computable.

Proof. It suffices to consider one iteration of Algorithm I. Assume that we are at iteration k of Algorithm I. If $-w_{k-1} \in A(z_{k-1})$ and $w_{k-1} \in B(z_{k-1})$ then

$$(a_k, y_k, \varepsilon_k) = (-w_{k-1}, z_{k-1}, 0), \qquad (b_k, x_k, \mu_k) = (w_{k-1}, z_{k-1}, 0)$$

trivially satisfies criterion (3.1).

Suppose $-w_{k-1} \notin A(z_{k-1})$ or $w_{k-1} \notin B(z_{k-1})$ and let

$$\hat{y} = (I + \lambda A)^{-1}(z_{k-1} - \lambda w_{k-1}), \quad \hat{a} = \lambda^{-1}(z_{k-1} - \lambda w_{k-1} - \hat{y}),$$
$$\hat{x} = (I + \lambda B)^{-1}(\hat{y} + \lambda w_{k-1}), \qquad \hat{b} = \lambda^{-1}(\hat{y} + \lambda w_{k-1} - \hat{x}).$$

It follows from these definitions that $\hat{a} \in A(\hat{y})$, $\hat{b} \in B(\hat{x})$, and

$$\lambda \hat{a} + \hat{y} = z_{k-1} - \lambda w_{k-1}, \qquad \lambda \hat{b} + \hat{x} = \hat{y} + \lambda w_{k-1}.$$

If $\hat{a} + \hat{b} = \hat{x} - \hat{y} = 0$, then it follows from the above equalities that $w_{k-1} = \hat{b} = -\hat{a}$, $z_{k-1} = \hat{x} = \hat{y}$ and, consequently, $-w_{k-1} \in A(z_{k-1})$ and $w_{k-1} \in B(z_{k-1})$, in contradiction with our assumption. Therefore, $\hat{a} + \hat{b} \neq 0$ or $\hat{x} - \hat{y} \neq 0$.

In view of the assumption of the proposition, one can generate sequences $\{(a_{k,j}, y_{k,j}, \varepsilon_{k,j})\}_{j \in \mathbb{N}}$ and $\{(b_{k,j}, x_{k,j}, \mu_{k,j})\}$ such that

$$a_{k,j} \in A^{[\varepsilon_{k,j}]}(y_{k,j}), \quad \varepsilon_{k,j} \leq \frac{1}{2\lambda j^2}, \quad \|\lambda a_{k,j} + y_{k,j} - (z_{k-1} + \lambda w_{k-1})\| \leq \frac{1}{j},$$
$$b_{k,j} \in B^{[\mu_{k,j}]}(x_{k,j}), \quad \mu_{k,j} \leq \frac{1}{2\lambda j^2}, \quad \|\lambda b_{k,j} + x_{k,j} - (y_{k-1} + \lambda w_{k-1})\| \leq \frac{1}{j}\,.$$

It follows from the above relations and from Proposition 2.2 that

$$\|\lambda(a_{k,j} - \hat{a})\|^2 + \|y_{k,j} - \hat{y}\|^2 \leq \frac{2}{j^2}.$$

In particular $\|y_{k,j} - \hat{y}\| \leq 2/j$ and

$$\|\lambda b_{k,j} + x_{k,j} - (\hat{y} + \lambda w_{k-1})\| \leq \frac{3}{j}.$$

Using again Proposition 2.2 we conclude that

$$\|b_{k,j} - \hat{b}\|^2 + \|x_{k,j} - \hat{x}\| \leq \frac{10}{j^2}.$$

Therefore,

$$\begin{aligned} &y_{k,j} \to \hat{y}, \quad a_{k,j} \to \hat{a}, \\ &x_{k,j} \to \hat{x}, \quad b_{k,j} \to \hat{b} \qquad \text{as } j \to \infty. \end{aligned}$$

and $\|\lambda(a_{k,j} - b_{k,j})\|^2 + \|x_{k,j} - y_{k,j}\|^2 \to \|\lambda(\hat{a} + \hat{b})\|^2 + \|\hat{x} - \hat{y}\|^2 > 0$ as $j \to \infty$. Since

$$\|\lambda a_{k,j} + y_{k,j} - (z_{k-1} + \lambda w_{k-1})\|^2 + \|\lambda b_{k,j} + x_{k,j} - (y_{k-1} + \lambda w_{k-1})\|^2 + 2\lambda(\varepsilon_{k,j} + \mu_{k,j}) \leq \frac{4}{j^2},$$

for j large enough criterion (3.1) will be satisfied. □

Appendix B. A technical Lemma

Let X be a real Banach space with topological dual X^* and let $\langle x, x^* \rangle$ stand for the duality product $x^*(x)$ for $x \in X$ and $x^* \in X^*$. A point-to-set operator $T : X \rightrightarrows X^*$ is monotone if

$$\langle x - y, x^* - y^* \rangle \geq 0 \qquad \forall x, y \in X,\ x^* \in T(x),\ y^* \in T(y).$$

A monotone operator T is maximal monotone if it is monotone and its graph is maximal in the family of the graphs of monotone operators.

The ε-*enlargement* of a maximal monotone operator T in X is defined as

$$T^{[\varepsilon]}(x) = \{x^* \in X^* \ : \ \langle x - y, x^* - y^* \rangle \geq -\varepsilon\} \qquad\qquad \varepsilon \geq 0,\ x \in X.$$

Fitzpatrick [10] function φ associated with a maximal monotone operator $T : X \rightrightarrows X^*$ is defined as

$$\varphi : X \times X^* \to \overline{\mathbb{R}}, \quad \varphi(x, x^*) = \sup_{y^* \in T(y)} \langle x, y^* \rangle + \langle y, x^* \rangle - \langle y, y^* \rangle. \tag{B.1}$$

Lemma B.1. *Suppose T is a maximal monotone operator in X and let φ be its Fitzpatrick function. Then*

1. *φ is convex and lower semicontinuous in the weak×weak-∗ topology of $X \times X^*$;*
2. *$\langle x, v \rangle \leq \varphi(x, v)$;*
3. *$v \in T(x) \iff \langle x, v \rangle = \varphi(x, v)$;*
4. *$v \in T^{[\varepsilon]}(x) \iff \varphi(x, v) \leq \langle x, v \rangle + \varepsilon$;*

Proof. Items 1, 2, and 3 where proved in [10].

Item 4 was proved in [25]. For the sake of completeness we present a proof. If follows from (B.1) that

$$\varphi(x, x^*) - \langle x, x^* \rangle = \sup_{y^* \in T(y)} \langle x - y, y^* - x^* \rangle.$$

To end the proof, use (2.3) to write

$$x^* \in T^{[\varepsilon]}(x) \iff \varepsilon \geq \langle x - y, y^* - x^* \rangle \ \forall y \in X,\ y^* \in T(y)$$

and combine the two above equations. □

The next technical lemma will be used in the convergence analysis of the inexact Douglas-Rachford methods proposed in this work.

Lemma B.2. *Let X be a real Banach space. If $T_1,\dots,T_m : X \rightrightarrows X^*$ are maximal monotone operators, $v_{i,k} \in T_i^{[\varepsilon_{i,k}]}(x_{i,k})$ for $i=1,\dots,m$ and $k=1,2,\dots,$ and*

$$x_{i,k}-x_{j,k}\to 0 \ \textit{(for } i,j=1,\dots,m\textit{)}, \quad \sum_{i=1}^m v_{i,k}\to v,$$
$$x_{i,k}\xrightarrow{w} x, \qquad v_{i,k}\xrightarrow{w*} v_i, \qquad \varepsilon_{i,k}\to 0 \ \textit{(for } i=1,\dots,m\textit{)},$$

as $k\to\infty$, then $v_i\in T_i(x)$ for $i=1\dots,m$.

Proof. Let φ_i be Fitzpatrick's function of T_i for $i=1\dots,m$. Since $v_{i,k}\in T_i^{[\varepsilon_{i,k}]}(x_{i,k})$,

$$\varphi_i(x_{i,k},v_{i,k})\le\langle x_{i,k},v_{i,k}\rangle+\varepsilon_{i,k} \qquad i=1,\dots,m.$$

Adding these inequalities for $i=1,\dots,m$ we obtain, after trivial algebraic manipulations the inequality

$$\sum_{i=1}^m\varphi_i(x_{i,k},v_{i,k})\le\langle x_{1,k},v\rangle+\left\langle x_{1,k},\sum_{i=1}^m v_{i,k}-v\right\rangle+\sum_{i=1}^m\langle x_{i,k}-x_{1,k},v_{i,k}\rangle+\varepsilon_{i,k}\,.$$

Each φ_i is lower semicontinuous in the weak×weak-∗ topology of $X\times X^*$. Moreover, it follows trivially from the assumptions of the lemma that the sequences $\{x_{i,k}\}$ and $\{v_{i,k}\}$ are bounded for $i=1,\dots,m$. Therefore, taking the $\liminf$ as $k\to\infty$ at both sides of the above inequality and using the lower semicontinuity of φ we conclude that

$$\sum_{i=1}^m\varphi_i(x,v_i)\le\langle x,v\rangle.$$

It also follows trivially from the assumptions of the lemma that $\sum_{i=1}^m v_i=v$. So, we can write this above inequality as

$$0\ge\sum_{i=1}^m\varphi_i(x,v_i)-\langle x,v_i\rangle.$$

Since all terms of this sum are non-negative, each one is equal to 0 and the conclusion follows from item 3 of Lemma B.1. ☐

Appendix C. Proofs of Lemma 4.2 and Theorem 4.3

Proof of Lemma 4.2. Take (z^*,w^*) in $S_e(A,B)$. It follows from Lemma 4.1 that $\{(z_k,\lambda w_k)\}$ is bounded, which proves item 1, and that

$$\sum_{k=1}^\infty\delta_k<\infty, \qquad \sum_{k=1}^\infty\|a_k+w_{k-1}\|^2<\infty.$$

It follows from the first above inequality that $\delta_k\to 0$ as $k\to\infty$. This result, together with the third line of (4.4) proves item 2. Item 3 follows trivially from the second above inequality and the equality in (4.1a).

It follows from the update formulas for z_k and w_k, and from (4.4) that

$$z_k-x_k+s_k=t_k\lambda(a_k+b_k), \qquad \lambda(w_k-b_k)+s_k=t_k(x_k-y_k).$$

Squaring both sides of each one of these equalities and adding them we conclude that

$$\|z_k - x_k + s_k\|^2 + \|\lambda(w_k - b_k) + s_k\|^2 = t_k^2 \rho_k.$$

Since $\delta_k \to 0$ as $k \to \infty$, it follows from the above equations and the last inequality in (4.5) that

$$t_k^2 \rho_k \to 0, \quad z_k - x_k + s_k \to 0, \quad \lambda(w_k - b_k) + s_k \to 0 \qquad \text{as } k \to \infty.$$

Item 4 follows from the above result and from item 2. □

Proof of Theorem 4.3. Suppose a subsequence $\{(z_{k_n}, w_{k_n})\}$ converges weakly to some (z, w). It follows from Lemma 4.2 that

$$\begin{aligned} &x_{k_n} \xrightarrow{w} z, \quad y_{k_n+1} \xrightarrow{w} z, \qquad x_{k_n} - y_{k_n+1} \to 0, \\ &b_{k_n} \xrightarrow{w} w, \quad a_{k_n+1} \xrightarrow{w} -w, \quad a_{k_n+1} + b_{k_n} \to 0, \qquad \text{as } k \to \infty. \\ &\mu_{k_n} \to 0, \quad \varepsilon_{k_n+1} \to 0 \end{aligned}$$

Since $b_{k_n} \in B^{[\mu_{k_n}]}(x_{k_n})$, $a_{k_n+1} \in A(y_{k_n+1})$ for $n = 1, 2, \dots$, it follows from Lemma B.2 that $w \in B(z)$ and $-w \in A(z)$. Therefore, $(z, w) \in S_e(A, B)$.

Define, again,

$$\Omega = \{(z, \lambda w) \;:\; (z, w) \in S_e(A, B)\}.$$

We have proved that all weak limit points of $\{(z_k, w_k)\}$ are in $S_e(A, B)$. Therefore, all weak limit points of $\{p_k = (z_k, \lambda w_k)\}$ are in Ω. In Lemma 4.1, we proved that $\{p_k\}$ is Fejér convergent to Ω. Since $\Omega \neq \emptyset$, it follows from these results and from Opial's Lemma [19] that the bounded sequence $\{p_k\}$ has a unique weak limit point and such a point belongs to Ω. Therefore, $\{p_k\}$ converges weakly to a point in $(z, \lambda w)$ where $(z, w) \in S_e(A, B)$. To end the proof, use items 3 and 4 of Lemma 4.2. □

Acknowledgements. We thank the anonymous referees for corrections, suggestion, and constructive criticisms which improved the original version of this work.

References

[1] M.M. Alves and M. Geremia, Iteration complexity of an inexact douglas-rachford method and of a Douglas-Rachford-Tseng's F-B four-operator splitting method for solving monotone inclusions. *Numer. Algorithms* **82** (2019) 263–295.

[2] F.J.A. Artacho, J.M. Borwein and M.K. Tam, Recent results on Douglas-Rachford methods for combinatorial optimization problems. *J. Optim. Theory Appl.* **163** (2014) 1–30.

[3] T. Aspelmeier, C. Charitha and D.R. Luke, Local linear convergence of the ADMM/Douglas-Rachford algorithms without strong convexity and application to statistical imaging. *SIAM J. Imaging Sci.* **9** (2016) 842–868.

[4] J.M. Borwein and B. Sims, The Douglas-Rachford algorithm in the absence of convexity, in Fixed-point algorithms for inverse problems in science and engineering, Vol. 49 of *Springer Optimization and its Applications*. Springer, New York (2011) 93–109.

[5] R.S. Burachik, A.N. Iusem, and B.F. Svaiter, Enlargement of monotone operators with applications to variational inequalities. *Set-Valued Anal.* **5** (1997) 159–180.

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

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

[8] J. Eckstein and B.F. Svaiter, A family of projective splitting methods for the sum of two maximal monotone operators. *Math. Program.* **111** (2008) 173–199.

[9] J. Eckstein and W. Yao, Relative-error approximate versions of Douglas–Rachford splitting and special cases of the ADMM. *Math. Program.* **170** (2017) 417–444.

[10] S. Fitzpatrick, Representing monotone operators by convex functions, in Workshop/Miniconference on Functional Analysis and Optimization. Vol. 20 of *Proceedings of the Centre for Mathematics and its Analysis*. Australian National University, Canberra (1988) 59–65.

[11] D. Gabay. Application of the method of multipliers to variational inequalities. *Stud. Math. Appl.* **15** (1983) 299–331.

[12] R. Glowinski and S.J. Osher, Splitting Methods in Communication, Imaging, Science, and Engineering. Scientific Computation. Springer, Cham (2016).

[13] R. Hesse and D.R. Luke, Nonconvex notions of regularity and convergence of fundamental algorithms for feasibility problems. *SIAM J. Optim.* **23** (2013) 2397–2419.

[14] R. Hesse, D.R. Luke and P. Neumann, Alternating projections and Douglas-Rachford for sparse affine feasibility. *IEEE Trans. Signal Process.* **62** (2014) 4868–4881.

[15] G. Li and T.K. Pong, Douglas-Rachford splitting for nonconvex optimization with application to nonconvex feasibility problems. *Math. Program.* **159** (2016) 371–401.

[16] J. Liang, J. Fadili, G. Peyré and R. Lukes, Activity identification and local linear convergence of Douglas-Rachford/ADMM under partial smoothness, in Scale space and variational methods in computer vision, Vol. 9087 of *Lecture Notes in Computer Science*. Springer, Cham (2015) 642–653.

[17] P.-L. Lions and B. Mercier, Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.* **16** (1979) 964–979.

[18] G.J. Minty, Monotone (nonlinear) operators in Hilbert space. *Duke Math. J.* **29** (1962) 341–346.

[19] Z. Opial, Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bull. Am. Math. Soc.* **73** (1967) 591–597.

[20] M.V. Solodov and B.F. Svaiter, A hybrid approximate extragradient-proximal point algorithm using the enlargement of a maximal monotone operator. *Set-Valued Anal.* **7** (1999) 323–345.

[21] M.V. Solodov and B.F. Svaiter. A hybrid projection-proximal point algorithm. *J. Convex Anal.* **6** (1990) 59–70.

[22] M.V. Solodov and B.F. Svaiter, Error bounds for proximal point subproblems and associated inexact proximal point algorithms. *Math. Program.* **88** (2000) 371–389.

[23] M.V. Solodov and B.F. Svaiter, An inexact hybrid generalized proximal point algorithm and some new results on the theory of Bregman functions. *Math. Oper. Res.* **25** (2000) 214–230.

[24] M.V. Solodov and B.F. Svaiter, A unified framework for some inexact proximal point algorithms. *Numer. Funct. Anal. Optim.* **22** (2001) 1013–1035.

[25] B.F. Svaiter, A family of enlargements of maximal monotone operators. *Set-Valued Anal.* **8** (2000) 311–328.

[26] B.F. Svaiter, On weak convergence of the Douglas-Rachford method. *SIAM J. Control Optim.* **49** (2011) 280–287.