The Cyclic Douglas-Rachford Algorithm with r-sets-Douglas-Rachford Operators

Francisco J. Aragón Artacho[1], Yair Censor[2] and Aviv Gibali[3,4]

[1]Department of Mathematics, University of Alicante,
03690 San Vicente del Raspeig, Alicante, Spain
{francisco.aragon@ua.es}

[2]Department of Mathematics, University of Haifa
Mt. Carmel, Haifa 3498838, Israel. {yair@math.haifa.ac.il}

[3]Department of Mathematics, ORT Braude College
Karmiel 2161002, Israel.
[4]The Center for Mathematics and Scientific Computation
University of Haifa, Mt. Carmel
Haifa 3498838, Israel. {avivg@braude.ac.il}

Abstract

The Douglas-Rachford (DR) algorithm is an iterative procedure that uses sequential reflections onto convex sets and which has become popular for convex feasibility problems. In this paper we propose a structural generalization that allows to use r-sets-DR operators in a cyclic fashion. We prove convergence and present numerical illustrations of the potential advantage of such operators with $r > 2$ over the classical 2-sets-DR operators in a cyclic algorithm.

1

1   Introduction

We consider the convex feasibility problem (CFP) in a real Hilbert space $\mathcal{H}$. For $i = 0, 1, \cdots, m - 1$, let $C_i \subseteq \mathcal{H}$ be nonempty, closed and convex sets. The CFP is to

$$\text{find a point } x^* \in C := \cap_{i=0}^{m-1} C_i. \tag{1.1}$$

The literature about projection methods for solving this problem is vast, see, e.g., [5], [17], [19, Chapter 5] or the recent [6]. The Douglas–Rachford (DR) algorithm whose origins are in [23] is a recent addition to this class of methods. We are unable to compete with the excellent coverage of the literature on this algorithm furnished in the recent 2017 paper by Bauschke and Moursi [8] and direct the reader there. The DR algorithm has witnessed a surge of interest and publications investigating it in all directions, such as, e.g., for the non-convex and inconsistent case [7, 10, 2]. A particular research direction consists of creating and studying new algorithmic structures that rely on the principles of the original DR algorithm.

This work belongs to this direction. We present and study a new algorithmic structure for the DR algorithm that cyclically uses r-sets-DR operators. In order to explain this recall the original 2-sets-DR algorithm. Given two sets C_0 and C_1 denote by P_{C_i} the orthogonal projection onto C_i and denote the reflection with respect to C_i by $\mathcal{R}_{C_i} = 2P_{C_i} - \text{Id}$, for $i = 0, 1$, where Id is the identity operator on $\mathcal{H}$. With the combined operator $\mathcal{V}_{C_0,C_1} := \mathcal{R}_{C_1}\mathcal{R}_{C_0}$ the original 2-sets-DR operator is defined as

$$\mathcal{T}_{C_0,C_1} := \frac{1}{2} \left(\text{Id} + \mathcal{V}_{C_0,C_1} \right). \tag{1.2}$$

The original DR algorithm, starting from an arbitrary $x_0 \in \mathcal{H}$, employed the sequential iterative process

$$x^{k+1} = \mathcal{T}_{C_0,C_1}(x^k), \quad k \geq 0.$$

It is, thus, restricted to handling only two sets.

2

Borwein and Tam in [11] introduced the cyclic-DR algorithm which is designed to solve CFPs with more than two sets. Their cyclic-DR algorithm applies sequentially the original 2-sets-DR operator (1.2) over subsequent pairs of sets. Censor and Mansour in [18] extended the algorithmic structure to deal with string-averaging and block-iterative structural regimes.

In this work we propose a cyclic DR algorithm that uses r-sets DR operators and prove its convergence. We present numerical illustrations of the potential advantage of r-sets DR operators with $r > 2$ over the original 2-sets-DR operator in this framework. We discovered the insight how to employ r-sets-DR operators which hides in the cyclic DR algorithm of Borwein and Tam [11, Section 3]. The Borwein-Tam cyclic DR algorithm uses 2-sets-DR operators sequentially but for each new pair of sets it uses the last set of the previous pair as the first set in the new pair. Mimicking this recipe enables us to use r-sets-DR operators in a cyclic DR algorithm.

The analysis of convergence of the algorithm presented here is quite standard and relies on tools from fixed point theory and convex analysis. So, the main contribution of the paper is the algorithmic discovery of how to properly employ r-sets DR operators with $r > 2$ in the cyclic DR algorithm. This is a theoretical development that shows that the Borwein and Tam cyclic DR algorithm is a special case of the more general framework proposed here. This opens the door for many future research questions of extending results on the Borwein and Tam cyclic DR algorithm to the new r-sets DR operators with $r > 2$ framework.

The paper is organized as follows. In Section 2 we present definitions and notions needed in the sequel. In Section 3 the r-sets-DR operator and cyclic algorithm are given and the algorithm's convergence is analyzed. Finally, in Section 4 numerical illustrations demonstrate the potential advantage of r-sets DR operators with $r > 2$.

2   Preliminaries

Let $\mathcal{H}$ be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$, and let D be a nonempty, closed and convex subset of $\mathcal{H}$. We write $x^k \rightharpoonup x$ to indicate that the sequence $\left\{ x^k \right\}_{k=0}^{\infty}$ converges weakly to x, and $x^k \to x$ to indicate that the sequence $\left\{ x^k \right\}_{k=0}^{\infty}$ converges strongly to x. We start by recalling the definition and properties of the metric projection operator. For each point $x \in \mathcal{H}$, there exists a unique nearest point in D, denoted by

3

$P_D(x)$. That is,

$$\|x - P_D(x)\| \le \|x - y\|, \text{ for all } y \in D. \tag{2.1}$$

The mapping $P_D : \mathcal{H} \to D$, called *the metric projection of H onto D,* is well-known, see for example [5, Fact 1.5(i)], to be *firmly nonexpansive,* thus, *nonexpansive,* see Definition 2.1 below. The metric projection P_D is characterized [25, Section 3] by the facts that $P_D(x) \in D$ and

$$\langle x - P_D(x), P_D(x) - y \rangle \ge 0, \text{ for all } x \in \mathcal{H}, \ y \in D. \tag{2.2}$$

If D is a hyperplane, or even a closed affine subspace, then (2.2) becomes an equality.

All items in the next definition can be found, e.g., in Cegielski's excellent book [15].

Definition 2.1 *Let $h : \mathcal{H} \to \mathcal{H}$ be an operator and let $D \subset \mathcal{H}$.*

(i) The operator h is called `Lipschitz continuous` *on $D \subset \mathcal{H}$ with constant $L > 0$ if*

$$\|h(x) - h(y)\| \le L\|x - y\|, \text{ for all } x, y \in D. \tag{2.3}$$

(ii) The operator h is called `nonexpansive` *on D if it is 1-Lipschitz continuous.*

(iii) The operator h is called `firmly nonexpansive` *[25] on D if*

$$\langle h(x) - h(y), x - y \rangle \ge \|h(x) - h(y)\|^2, \text{ for all } x, y \in D. \tag{2.4}$$

(iv) The operator h is called `averaged` *[4] if there exists a nonexpansive operator $N : \mathcal{H} \to \mathcal{H}$ and a number $c \in (0, 1)$ such that*

$$h = (1 - c)\,\mathrm{Id} + cN. \tag{2.5}$$

In this case, we say that h is c-av [13].

(v) A nonexpansive operator h satisfies `Condition (W)` *[24] if whenever $\{x^k - y^k\}_{k=1}^{\infty}$ is bounded and $\|x^k - y^k\| - \|h(x^k) - h(y^k)\| \to 0$, it follows that $(x^k - y^k) - (h(x^k) - h(y^k)) \rightharpoonup 0$.*

(vi) The operator h is called `strongly nonexpansive` *[12] if it is nonexpansive and whenever $\{x^k - y^k\}_{k=1}^{\infty}$ is bounded and $\|x^k - y^k\| - \|h(x^k) - h(y^k)\| \to 0$, it follows that $(x^k - y^k) - (h(x^k) - h(y^k)) \to 0$.*

Definition 2.2 *Let $h : \mathcal{H} \to \mathcal{H}$ be an operator with $\mathrm{Fix}(h) := \{x \in \mathcal{H} \mid h(x) = x\} \neq \emptyset$ and let $D \subseteq \mathcal{H}$ be a nonempty, closed and convex set.*

(i) The operator h is called `quasi-nonexpansive` *(QNE) if for all $x \in \mathcal{H}$ and all $z \in \mathrm{Fix}(h)$,*

$$\|h(x) - z\| \leq \|x - z\|. \tag{2.6}$$

(ii) A sequence $\{x_k\}_{k=0}^{\infty} \subset \mathcal{H}$ is said to be `Fejér-monotone` *with respect to D, if for all $k \geq 0$,*

$$\|x^{k+1} - u\| \leq \|x^k - u\|, \ \textit{for any } u \in D. \tag{2.7}$$

Some of the relations between the above classes of operators are collected in the following lemma. For more details and proofs, see Bruck and Reich [12], Baillon et al. [4], Goebel and Reich [25], Byrne [13] and Combettes [20].

Lemma 2.3 *(i) The operator $h : \mathcal{H} \to \mathcal{H}$ is firmly nonexpansive, if and only if it is $1/2$-averaged.*

(ii) If h_1 and h_2 are c_1-av and c_2-av, respectively, then their composition $S = h_1 h_2$ is $(c_1 + c_2 - c_1 c_2)$-av.

(iii) If h_1 and h_2 are averaged and $\mathrm{Fix}(h_1) \cap \mathrm{Fix}(h_2) \neq \emptyset$, then

$$\mathrm{Fix}(h_1) \cap \mathrm{Fix}(h_2) = \mathrm{Fix}(h_1 h_2) = \mathrm{Fix}(w_1 h_1 + w_2 h_2) \tag{2.8}$$

with $w_1 + w_2 = 1$, $w_1, w_2 \in (0,1)$. This result can be generalized for any finite number of averaged operators, see, e.g., [20, Lemma 2.2].

(iv) Every averaged operator is strongly nonexpansive and, therefore, satisfies condition (W).

Another useful property of a sequence of operators is the following, see, e.g., [15, Definition 3.6.1].

Definition 2.4 *Let $\{U_j\}_{j=1}^{\infty}$ be a sequence of operators $U_j : \mathcal{H} \to \mathcal{H}$ and denote $T_\ell = U_\ell U_{\ell-1}, \ldots U_1$. We say that $\{U_j\}_{j=1}^{\infty}$ is* `asymptotically regular` *if*

$$\lim_{\ell \to \infty} \|T_{\ell+1}(x) - T_\ell(x)\| = 0, \ \textit{for all } x \in \mathcal{H}. \tag{2.9}$$

The well-known Opial Theorem [26, Theorm 1], see also [15, Theorem 3.5.1], is presented next.

Theorem 2.5 *Let $\mathcal{H}$ be a real Hilbert space and let $D \subset \mathcal{H}$ be closed and convex. If $h : D \to D$ is an averaged operator with $\mathrm{Fix}(h) \neq \emptyset$ then, for any $x^0 \in D$, the sequence $\left\{x^k\right\}_{k=0}^{\infty}$, generated by $x^{k+1} = h(x^k)$, converges weakly to a point $x^* \in \mathrm{Fix}(h)$.*

3 The r-sets-Douglas-Rachford operator and algorithm

The r-sets-Douglas-Rachford (r-sets-DR) operator was defined in [18] as follows.

Definition 3.1 *[18, Definition 22] Given a sequence of r nonempty closed convex sets, $r \geq 2$, $C_0, C_1, \ldots, C_{r-1} \subseteq \mathcal{H}$, define the composite reflection operator $\mathcal{V}_{C_0,C_1,\ldots,C_{r-1}} : \mathcal{H} \to \mathcal{H}$ by*

$$\mathcal{V}_{C_0,C_1,\ldots,C_{r-1}} := \mathcal{R}_{C_{r-1}} \mathcal{R}_{C_{r-2}} \cdots \mathcal{R}_{C_0}, \tag{3.1}$$

where $\mathcal{R}_{C_i} = 2P_{C_i} - \mathrm{Id}$ is the reflection on the corresponding C_i. The r-sets-DR operator $\mathcal{T}_{C_0,C_1,\ldots,C_{r-1}} : \mathcal{H} \to \mathcal{H}$ is defined by

$$\mathcal{T}_{C_0,C_1,\ldots,C_{r-1}} := \frac{1}{2} \left(\mathrm{Id} + \mathcal{V}_{C_0,C_1,\ldots,C_{r-1}} \right). \tag{3.2}$$

For $r = 2$ the r-sets-DR operator coincides with the original 2-sets-DR operator (1.2) and when it is applied sequentially repeatedly on two sets $m = 2$ the original DR algorithm is recovered. For $r = 3$ the r-sets-DR operator coincides with the 3-sets-DR operator defined in [1, Eq. (2)]. The question whether the 3-sets-DR operator can be applied sequentially repeatedly on three sets $m = 3$ was asked there. However, it is shown, in [1, Example 2.1], that such an iterative process of the form

$$x^{k+1} = \mathcal{T}_{C_0,C_1,C_2}(x^k) \tag{3.3}$$

that uses 3-sets-DR operators sequentially for $m = 3$ need not generate a sequence that converges to a feasible point.

In this paper we discovered the insight how to employ r-sets-DR operators which hides in the cyclic DR algorithm of Borwein and Tam [11, Section 3]. The Borwein-Tam cyclic DR algorithm uses 2-sets-DR operators sequentially but for each new pair of sets it uses the last set of the previous pair as the first set in the new pair. Mimicking this recipe enables us to use r-sets-DR operators in a cyclic DR algorithm.

Given a CFP (1.1) with m sets indexed by $0, 1, \ldots, m-1$, and an integer $r \geq 2$, we compose, for any integer $d \geq 1$, the finite sequence of sets

$$C_{m,r}(d) := C_{((r-1)d-(r-1)) \bmod m}, C_{((r-1)d-(r-2)) \bmod m}, \ldots, C_{((r-1)d) \bmod m}, \quad (3.4)$$

in which the individual sets belong to the family of sets of the given CFP. We further define the operator $S_d : \mathcal{H} \to \mathcal{H}$

$$S_d := \mathcal{T}_{C_{m,r}(d)}, \quad (3.5)$$

performing an r-sets-DR operator on the sets of $C_{m,r}(d)$. We use it to present our r-sets-Douglas-Rachford algorithm.

Algorithm 3.2 _The r-sets-Douglas-Rachford cyclic Algorithm_

Step 0: _Select an arbitrary starting point $x^0 \in \mathcal{H}$ and set $k = 0$._

Step 1: _Given the current iterate x^k, compute_

$$x^{k+1} = S_{k+1}(x^k). \quad (3.6)$$

Step 2: _If $x^k = x^{k+1} = \cdots = x^{k+\lceil m/r \rceil}$ (where $\lceil a \rceil$ stands for the smallest integer greater than or equal to a) then stop. Otherwise, set $k \leftarrow (k+1)$ and return to **Step 1.**_

Example 3.3 _Assume that the CFP contains 5 sets C_0, C_1, C_2, C_3, C_4, and choose $r = 3$. Then_

$$C_{5,3}(1) = C_{((3-1)1-(3-1)) \bmod 5}, C_{((3-1)1-(3-2)) \bmod 5}, \ldots, C_{((3-1)1) \bmod 5} = C_0, C_1, C_2 \quad (3.7)$$

and $S_1 = \mathcal{T}_{C_{5,3}(1)} = \mathcal{T}_{C_0,C_1,C_2}$. Similarly, $S_2 = \mathcal{T}_{C_{5,3}(2)} = \mathcal{T}_{C_2,C_3,C_4}$, $S_3 = \mathcal{T}_{C_{5,3}(3)} = \mathcal{T}_{C_4,C_0,C_1}$, $S_4 = \mathcal{T}_{C_{5,3}(4)} = \mathcal{T}_{C_1,C_2,C_3}$ and $S_5 = \mathcal{T}_{C_{5,3}(5)} = \mathcal{T}_{C_3,C_4,C_0}$, and so on for all integers $d \geq 1$. This realizes the algorithmic structure stated above.

One way to handle the convergence proof of Algorithm 3.2 is to base it on an appropriate generalization of Opial's theorem such as [16, Theorem 9.9], see also [15, Section 3.5]. This approach leads to the next theorem.

Theorem 3.4 *Let $C_i \subseteq \mathcal{H}$, for $i = 0, \ldots, m-1$, be nonempty, closed and convex sets with* $\mathrm{int}\left(\cap_{0=1}^{m-1} C_i\right) \neq \emptyset$. *Let $\{S_k\}_{k=1}^{\infty}$ be the family of operators defined in (3.5). Assume that $S : \mathcal{H} \to \mathcal{H}$ is a nonexpansive operator with* $\mathrm{Fix}(S) \neq \emptyset$ *for which the following assumptions hold:*
(1) $\mathrm{Fix}(S) \subseteq (\cap_{k=1}^{\infty} \mathrm{Fix}(S_k)) \cap \left(\cap_{0=1}^{m-1} C_i\right)$,
(2) $\left\{x^k\right\}_{k=0}^{\infty}$, *generated by Algorithm 3.2, is Fejér-monotone with respect to* $\mathrm{Fix}(S)$,
(3) the inequality $\|S_k(x^k) - x^k\| \geq \beta \|S(x^k) - x^k\|$ *is satisfied for all $k \geq 0$, for some $\beta > 0$.*
Then the sequence $\left\{x^k\right\}_{k=0}^{\infty}$, *generated by Algorithm 3.2, converges weakly to a point $x^* \in \mathrm{Fix}(S)$, and, in particular, $x^* \in \cap_{0=1}^{m-1} C_i$.*

Proof. We first show that the family of operators $\{S_k\}_{k=1}^{\infty}$, defined in (3.5), is quasi-nonexpansive and asymptotically regular (Definitions 2.2 and 2.4 above). Let $d \in \mathbb{N}$, then the composition of reflections operator $\mathcal{V}_{C_{m,r}(d)}$ is nonexpansive and hence $\mathcal{T}_{C_{m,r}(d)}$ is firmly-nonexpansive (1/2-averaged). Thus, this operator is also asymptotically regular, see, e.g., the discussion following Theorem 9.7 in [16]. The asymptotic regularity of $\{S_k\}_{k=1}^{\infty}$ and the assumptions of the theorem enable the use of [14, Theorem 1] (see also [16, Theorem 9.9] and [15, Subsection 3.6]) to obtain the desired result. ∎

Since the conditions of Theorem 3.4 are not easy to verify in practice, we present an alternative convergence result for Algorithm 3.2. Given m nonempty, closed and convex sets C_i, for $i = 0, 1, \ldots, m-1$ and $1 < r \leq m-1$, we look at the string of $(r-1)m$ sets that is composed of $r-1$ copies of $\{C_0, C_1, \ldots, C_{m-1}\}$, i.e.,

$$\underbrace{C_0, C_1, \ldots, C_{m-1}}_{1}, \underbrace{C_0, C_1, \ldots, C_{m-1}}_{2}, \ldots, \underbrace{C_0, C_1, \ldots, C_{m-1}}_{r-1}, \quad (3.8)$$

and define with (3.5) the composite operator Q:

$$Q := S_m \cdots S_2 S_1. \quad (3.9)$$

8

Example 3.5 *To continue Example 3.3, here (3.8) takes the form:*

$$\underbrace{C_0, C_1, \ldots, C_4}_{1}, \underbrace{C_0, C_1, \ldots, C_4}_{2} \tag{3.10}$$

and the operator Q of (3.9) is:

$$Q := S_5 S_4 S_3 S_2 S_1 = \mathcal{T}_{C_3,C_4,C_0} \mathcal{T}_{C_1,C_2,C_3} \mathcal{T}_{C_4,C_0,C_1} \mathcal{T}_{C_2,C_3,C_4} \mathcal{T}_{C_0,C_1,C_2}. \tag{3.11}$$

This kind of algorithmic operator guarantees that the last set that is handled is C_0.

We will prove the convergence of Algorithm 3.2 with S_k in (3.6) replaced by Q, for all $k \geq 1$. We need the following lemma which is based on [18, Corollary 23].

Lemma 3.6 *Let $C_i \subseteq \mathcal{H}$, for $i = 0, 1, \ldots, m-1$, be nonempty, closed and convex sets with $\mathrm{int}\left(\cap_{i=0}^{m-1} C_i\right) \neq \emptyset$. For fixed $r \in \{2, 3, \ldots, m-1\}$, we have*

$$\cap_{i=0}^{r-1} C_i = \mathrm{Fix}(\mathcal{T}_{C_0,C_1,\ldots,C_{r-1}}). \tag{3.12}$$

Proof. Obviously,

$$\emptyset \neq \mathrm{int}\left(\cap_{i=0}^{m-1} C_i\right) \subseteq \mathrm{int}\left(\cap_{i=0}^{r-1} C_i\right). \tag{3.13}$$

Since

$$\mathrm{Fix}(\mathcal{T}_{C_0,C_1,\ldots,C_{r-1}}) = \mathrm{Fix}(\mathcal{V}_{C_0,C_1,\cdots,C_{r-1}}) = \cap_{i=0}^{r-1} \mathrm{Fix}(\mathcal{R}_{C_i}) = \cap_{i=0}^{r-1} C_i, \tag{3.14}$$

combining the above, we get (3.12) as desired. ∎

The alternative convergence result of Algorithm 3.2 follows.

Theorem 3.7 *Let $C_i \subseteq \mathcal{H}$, for $i = 0, \ldots, m-1$, be nonempty, closed and convex sets. If $\mathrm{int}\left(\cap_{0=1}^{m-1} C_i\right) \neq \emptyset$ then any sequence $\left\{x^k\right\}_{k=0}^{\infty}$, generated by Algorithm 3.2 with S_k replaced by Q as in (3.9), converges weakly to a point x^* which solves the convex feasibility problem (1.1).*

9

Proof. Let $r \in \{2, 3, \ldots, m-1\}$. Since the operator $\mathcal{V}_{C_1, C_2, \cdots, C_r}$ is nonexpansive, $\mathcal{T}_{C_1, C_2, \ldots, C_r}$ is firmly-nonexpansive, i.e., $1/2$-averaged. Since composition of averaged operators is averaged, we get that any operator S_d (3.5) is averaged and so is also Q of (3.9).

Next, we study Fix(Q). Since $\cap_{i=0}^{m-1} C_i \neq \emptyset$, Lemma 2.3(iii) and (3.12) yield

$$\text{Fix}(Q) = \text{Fix}(S_m \cdots S_2 S_1) = \cap_{d=1}^m \text{Fix } S_d = \cap_{i=0}^{m-1} C_i. \qquad (3.15)$$

The rest of the proof follows directly from the Opial theorem (Theorem 2.5 above) and the proof is complete. ∎

Remark 3.8 (i) *In the finite-dimensional case, Theorem 3.7 implies also convergence of Algorithm 3.2 with $\{S_k\}_{k=1}^\infty$.*

(ii) *In the special case when $\frac{m}{r-1} = n \in \mathbb{N}$ one can define the operator $\widetilde{Q} := S_n \cdots S_2 S_1$, which means that $\widetilde{Q}$ preforms one full "sweep" over the sets $C_{m-1}, \ldots, C_0$, and use it instead of Q.*

Definition 3.9 *Let $\mathbb{N}$ be the set of natural numbers, $\{h_1, h_2, \ldots\}$ be a sequence of operators, and $r : \mathbb{N} \to \mathbb{N}$. An unrestricted (or random) product of these operators is the sequence $\{S_n\}_{n \in \mathbb{N}}$ defined by $S_n := h_{r(n)} h_{r(n-1)} \cdots h_{r(1)}$.*

We recall the following result by Dye and Reich.

Theorem 3.10 *[24, Theorem 1] Let $T_1 : \mathcal{H} \to \mathcal{H}$ and $T_2 : \mathcal{H} \to \mathcal{H}$ be two (W) nonexpansive mappings on a Hilbert space $\mathcal{H}$, whose fixed point sets have a nonempty intersection. Then any random product $\{S_n\}_{n \in \mathbb{N}}$, from T_1 and T_2 converges weakly (to a common fixed point).*

With the aid of this theorem we can prove that products of projection operators may be interlaced between the r-sets-DR operators in Algorithm 3.2.

Theorem 3.11 *Let $C_i \subseteq \mathcal{H}$, for $i = 0, 1, \ldots, m-1$, be nonempty, closed and convex sets with int $\left(\cap_{0=1}^{m-1} C_i\right) \neq \emptyset$. Given the operators $T_1 = Q$ (where Q is defined in (3.9)) and $T_2 = P_{C_0} P_{C_1} \cdots P_{C_{m-1}}$, any sequence $\{x^k\}_{k=0}^\infty$, generated by any random product from T_1 and T_2, converges weakly to a point x^* which solves the CFP (1.1).*

10

Proof. By (3.15)

$$\text{Fix}(T_1) = \text{Fix}(Q) = \cap_{i=0}^{m-1} C_i \tag{3.16}$$

and clearly also

$$\text{Fix}(T_2) = \text{Fix}(P_{C_0} P_{C_1} \cdots P_{C_{m-1}}) = \cap_{i=0}^{m-1} C_i, \tag{3.17}$$

yielding $\text{Fix}(T_1) \cap \text{Fix}(T_2) = \cap_{i=0}^{m-1} C_i \neq \emptyset$. Since (see the proof of Theorem 3.7) the operator Q is $1/2$-averaged we use Lemma 2.3(iv), to know that it satisfies condition (W). Since T_2 is also averaged, it also satisfies condition (W). Applying Theorem 3.10 the desired result is obtained. ∎

Remark 3.12 *Theorem 3.11 is established with $T_2 = P_{C_0} P_{C_1} \cdots P_{C_{m-1}}$, but as a matter of fact, any (W) nonexpansive operator can be chosen as long as $\text{Fix}(T_2) = \cap_{i=0}^{m-1} C_i$, for example $\mathcal{T}_{C_0,C_1,\cdots,C_{m-1}}$ ((3.2) with $r = m$).*

Remark 3.13 *In [3] a generalized DR operator, called the averaged alternating modified reflections (AAMR) operator, is introduced. It allows to choose any parameters $\alpha \in (0,1)$ and $\beta \in (0,1)$ in the operator $\mathcal{T}_{A,B,\alpha,\beta} : \mathcal{H} \to \mathcal{H}$ given by*

$$\mathcal{T}_{A,B,\alpha,\beta} := (1-\alpha)\,\text{Id} + \alpha(2\beta P_B - \text{Id})(2\beta P_A - \text{Id}) \tag{3.18}$$

where A and B are nonempty, closed and convex sets. We conjecture that our analysis given here can be properly expanded to include r-sets-AAMR operators but we leave it for future work. In this respect, it is worthwhile to note that the condition $\text{int}\left(\cap_{0=1}^{m-1} C_i\right) \neq \emptyset$ seems to be too restrictive. Probably additional convergence properties can be derived by relaxing it. For instance, in finite-dimensions, under a less restrictive condition, linear convergence results are proved in [21] for the cyclic 2-sets-DR algorithm with a generalized DR operator.

4 Numerical demonstrations

We set out to investigate and verify whether r-sets-DR operators with $r > 2$ in a cyclic DR algorithm applied to a CFP are advantageous in any way over the cyclic DR algorithm with $r = 2$ proposed in [11, Section 3]. Our numerical illustrations demonstrate the potential advantage of r-sets DR operators with $r > 2$, especially when the number of sets is large.

Additionally, we included in our numerical experiments also the "Product Space Douglas–Rachford" algorithm, which is based on Pierra's product space formulation [27]. The original 2-sets DR algorithm is applied sequentially to the product set

$$\mathbf{C} := \prod_{i=0}^{m-1} C_i \qquad (4.1)$$

and to the diagonal set

$$\mathbf{D} := \{(x, x, \ldots, x) \in \mathcal{H}^m \mid x \in \mathcal{H}\}. \qquad (4.2)$$

The iterative process obtained in this way has the form

$$x^{k+1} = \mathcal{T}_{\mathbf{C},\mathbf{D}}(x^k), \qquad (4.3)$$

where $\mathcal{T}_{\mathbf{C},\mathbf{D}}$ is the 2-sets-DR operator as in (1.2), see, e.g., [1, Section 3].

We consider two types of CFPs, with linear and quadratic constraints. For each of these type of problems and each problem size, 10 random problems were generated and solved independently. Algorithm 3.2 was run until the stopping criterion

$$\frac{\left\| x^{k+j} - x^{k+j-1} \right\|}{\left\| x^{k+j-1} \right\|} \le 10^{-12}, \quad \text{for all } j = 1, 2, \ldots, \lceil m/r \rceil, \qquad (4.4)$$

was met. All the experiments were run in $\mathbb{R}^n$ (the n-dimensional Euclidean space) with $n = 1000$. Initialization vectors x^0 were generated by randomly uniformly picking their coordinates from the range $[-10, 10]$. All codes were written in Python 2.7 and the tests were run on an Intel Core i7-4770 CPU 3.40GHz with 32GB RAM, under Windows 10 (64-bit).

Example 4.1 (Linear CFPs) *In this example we consider solving a system of linear equations $Ax = 0_m$, where $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$ and $0_m \in \mathbb{R}^m$. Since in real-life, experiments and measurements often come with "noise", we investigate the performances of our algorithm for solving the perturbed system of linear inequalities $-b_i \le \langle a^i, x \rangle \le b_i$, $i = 1, 2, \ldots, m$. The coordinates of a^i were randomly uniformly generated in $[-1, 1]$ and then the vectors were normalized, and b_i was randomly uniformly chosen in $[0, 0.1]$.*

Example 4.2 (Quadratic CFPs) *In this example we followed the experimental setup in [11, Section 5] and generated CFPs consisting of balls of*

various sizes. Each ball was created by picking a ball center a^i with coordinates randomly uniformly generated in the range $[-5, 5]$. Then a radius $b_i := \|a^i\| + \alpha_i$ was defined by adding to the center's distance from the origin $\|a^i\|$ a random number uniformly picked from the range $[0, 0.1]$ guaranteeing that the ball includes the origin, thus, yielding a consistent CFP.

In our first experiment we compare the product space DR algorithm (4.3) with our cyclic r-sets-DR Algorithm 3.2 with different values of r. In Figure 1 we show the running times of the different methods when the number of sets of the CFP varied from 50 to 1000. The stopping criterion (4.4) was also used for the product space DR algorithm, but this time only for $j = 1$. Note that a logarithmic scale was employed for the y-axis. We observe that for 1000 constraints, a number which is relatively small, the product space DR algorithm was nearly 100 times slower than each of the r-sets-DR methods. It is not difficult to understand the main reason why this happens: it requires to work in the product space $\mathbb{R}^{m \times n}$ instead of the original space $\mathbb{R}^n$.

In our second experiment, we compare our cyclic r-sets-DR methods for a wide range of constraints between 200 and 50,000. For each problem size, 10 independent random problems were tested. The averaged runtimes are shown in Figure 2. The performance profiles comparing the methods, shown in Figure 3, were obtained as follows, see [22] and [9]. Let $\mathcal{S}$ denote the set of all 6 solvers compared (namely, the original 2-sets-DR scheme, and the cyclic r-sets-DR algorithm with $r = 3, 5, 10, 20, 50$). Let $\mathcal{P} := \{200, 2500, 5,000, \ldots, 50,000\}$ be the set of problems. Let $t_{p,s}$ be the averaged time required to solve problem $p \in \mathcal{P}$ over the 10 random instances tested, by solver s.

For each problem p and solver s, the *performance ratio* is defined by

$$r_{p,s} := \frac{t_{p,s}}{\min\{t_{p,s} \mid s \in \mathcal{S}\}}. \tag{4.5}$$

The *performance profile* of a solver s is a real-valued function $\pi_s : [1, +\infty) \to [0, 1]$ defined by

$$\pi_s(\tau) := \frac{1}{|\mathcal{P}|} \left|\{p \in \mathcal{P} \mid r_{p,s} \leq \tau\}\right|, \tag{4.6}$$

where $|\mathcal{P}|$ is the cardinality of the test set $\mathcal{P}$. This function indicates the probability that a performance ratio $r_{p,s}$ is within a factor τ of the best possible ratio. Thus, $\pi_s(1)$ represents the portion of problems for which solver $s \in \mathcal{S}$ has the best performance among all other solvers.

13
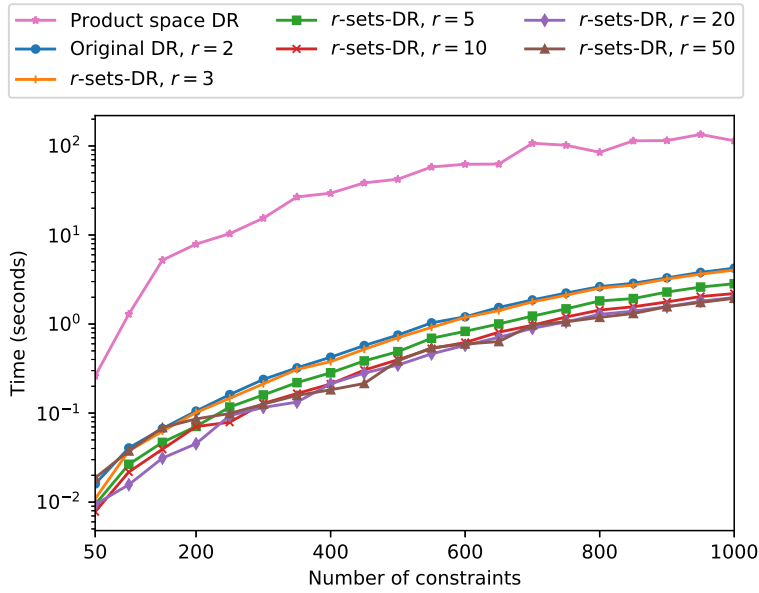
(a) Linear CFPs
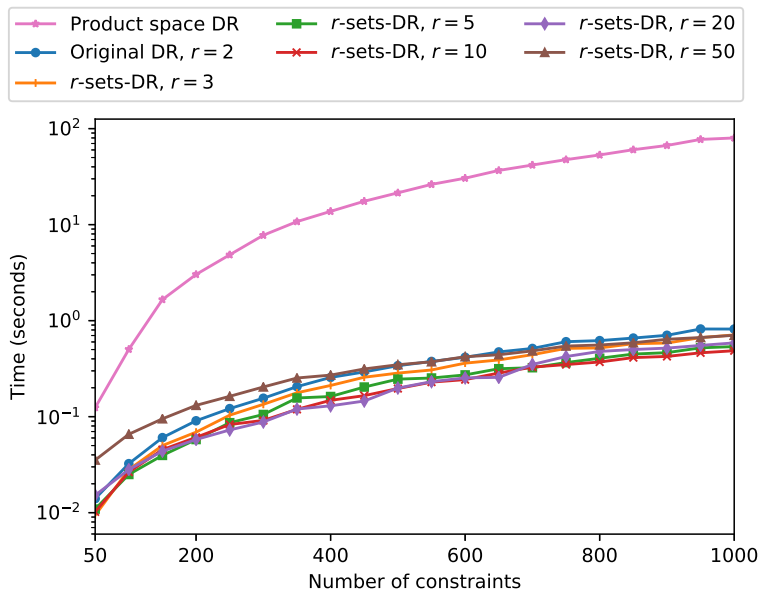
(b) Quadratic CFPs

Figure 1: Runtimes in seconds averaged over 10 independent problems for varying number of constraints. The product space DR algorithm is outperformed by the cyclic r-sets-DR algorithms.

From Figures 2 and 3 we deduce that the cyclic DR algorithm with $r = 2$ is clearly outperformed by the cyclic DR algorithms with the other $r = 3, 5, 10, 20, 50$ r-sets-DR operators. This trend seems even to grow and become more pronounced as the problem sizes grow. The best performance for both linear and quadratic problems that were tested was achieved for $r = 20$, closely followed by $r = 10$. On average, these two algorithms were two times faster than the original cyclic DR algorithm.

In our last experiment, we compare the values of

$$\text{Error}(x^k) := \sum_{i=0}^{m-1} \|P_{C_i}(x^k) - x^k\| \tag{4.7}$$

with respect to the number of iterations and projections employed by each of the methods in one particular random experiment with 10,000 constraints. Of course, the larger r is, the more projections the method uses to compute each iteration. The results, which are presented in Figure 4, clearly show that the original cyclic DR scheme with $r = 2$ uses two times more projections than the r-sets-DR method with $r = 10, 20$ or 50 to achieve the same accuracy.

Extensive numerical study is called for, and indeed planned for future work, to explore further the computational aspects of the of r-sets DR operators with $r > 2$.
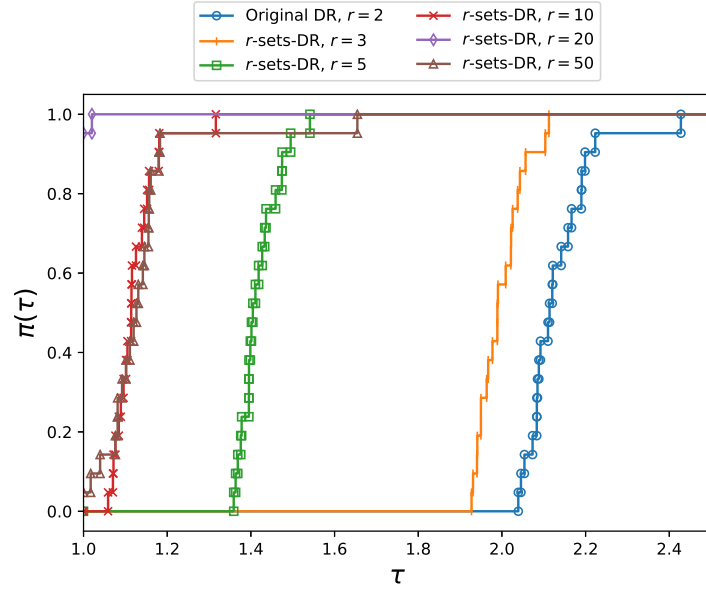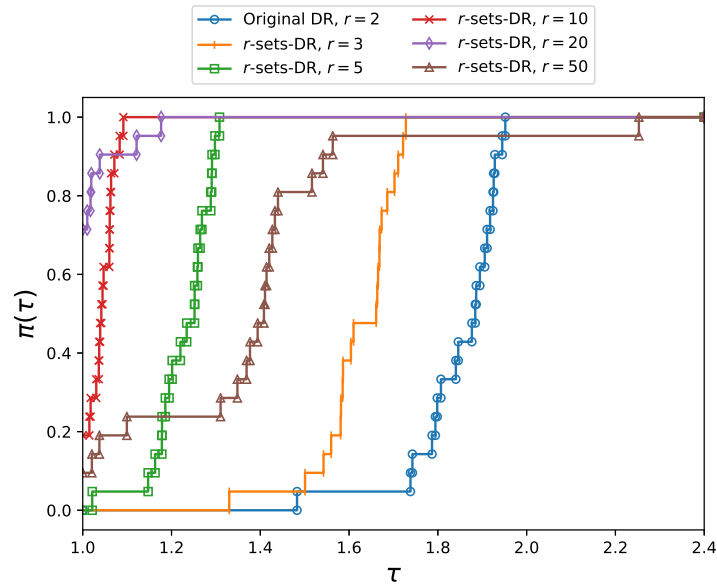
(a) Linear CFPs

(b) Quadratic CFPs

Figure 2: Runtimes in seconds averaged over 10 independent problems for varying number of constraints. The cyclic DR algorithm with $r = 2$ is outperformed by the cyclic DR algorithms with the other $r = 3$, 5, 10, 20, 50 r-sets-DR operators.

(a) Linear CFPs

(b) Quadratic CFPs

Figure 3: Performance profiles over 10 independent problems for varying number of constraints. The cyclic DR algorithm with $r = 2$ is outperformed by the cyclic DR algorithms with the other $r = 3, 5, 10, 20, 50$ r-sets-DR operators.

(a) Linear CFPs (iterations)

(b) Linear CFPs (projections)

(c) Quadratic CFPs (iterations)
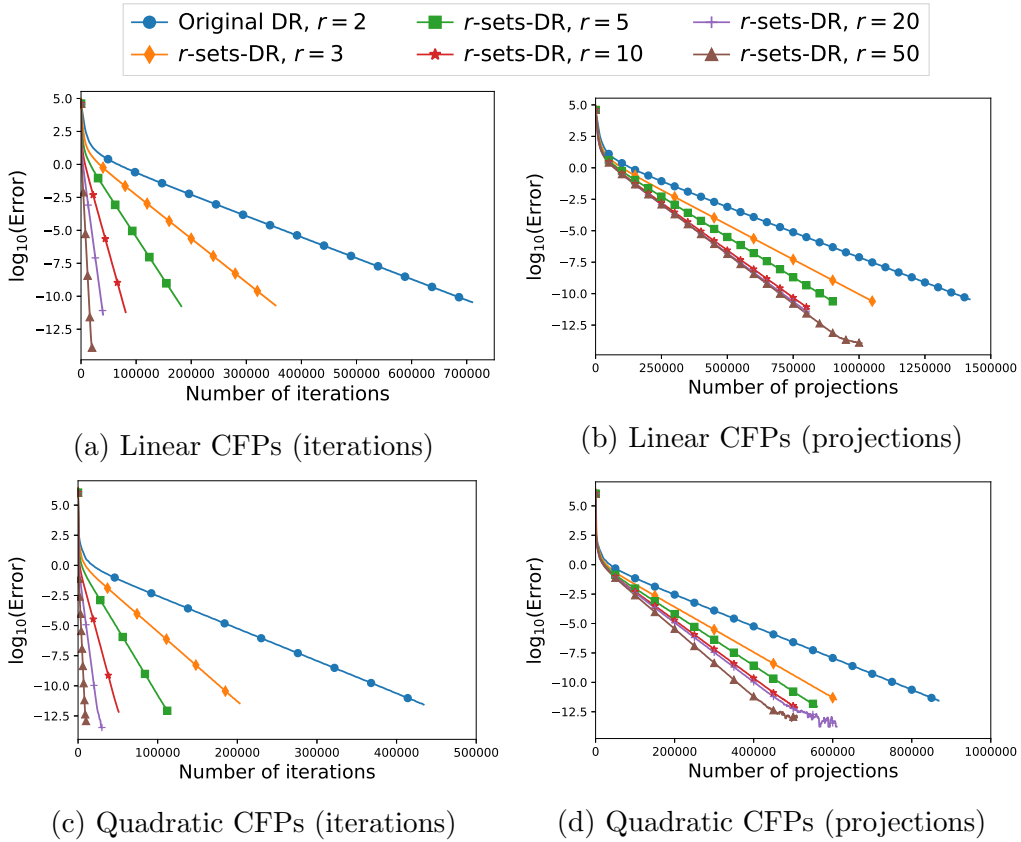
(d) Quadratic CFPs (projections)

Figure 4: Comparison of the value of Error in (4.7) and the number of iterations and projections used for one randomly generated problem with 10,000 constraints. The original cyclic DR algorithm with $r = 2$ needs more projections to achieve the same accuracy than the r-sets DR methods with $r > 2$.

18

References

[1] F.J. Aragón Artacho, J.M. Borwein and M.K. Tam, Recent results on Douglas-Rachford methods for combinatorial optimization problems, *Journal of Optimizaion Theory and Applications* **163** (2014), 1–30.

[2] F.J. Aragón Artacho, J.M. Borwein and M.K. Tam, Global behavior of the Douglas-Rachford method for a nonconvex feasibility problem, *Journal of Global Optimization* **65** (2016), 309–327.

[3] F.J. Aragón Artacho and R. Campoy, A new projection method for finding the closest point in the intersection of convex sets, *Computational Optimization and Applications* **69** (2018), 99–132.

[4] J.-B. Baillon, R.E. Bruck and S. Reich, On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces, *Houston Journal of Mathematics* **4** (1978), 1–9.

[5] H.H. Bauschke and J.M. Borwein, On projection algorithms for solving convex feasibility problems, *SIAM Review* **38** (1996), 367–426.

[6] H.H. Bauschke and P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Second Edition, Springer International Publishing AG, 2017.

[7] H.H. Bauschke, M.N. Dao and S.B. Lindstrom, The Douglas–Rachford algorithm for a hyperplane and a doubleton. *Technical report*, April 2018. Available at: https://arxiv.org/abs/1804.08880.

[8] H.H. Bauschke and W.M. Moursi, On the Douglas-Rachford algorithm, *Mathematical Programming, Series A* **164** (2017), 263–284.

[9] V. Beiranvand, W. Hare and Y. Lucet, Best practices for comparing optimization algorithms, *Optimization and Engineering* **18** (2017), 815–848.

[10] J. Benoist, The Douglas–Rachford algorithm for the case of the sphere and the line, *Journal of Global Optimization* **63** (2015), 363–380.

[11] J.M. Borwein and M.K. Tam, A cyclic Douglas-Rachford iteration scheme, *Journal of Optimizaion Theory and Applications* **160** (2014), 1–29.

[12] R.E. Bruck and S. Reich, Nonexpansive projections and resolvents of accretive operators in Banach spaces, *Houston Journal of Mathematics* **3** (1977), 459–470.

[13] C.L. Byrne, A unified treatment of some iterative algorithms in signal processing and image reconstruction, *Inverse Problems* **20** (1999), 1295–1313.

[14] A. Cegielski, A generalization of the Opial's Theorem, *Control and Cybernetics* **36** (2007), 601–610.

[15] A. Cegielski, *Iterative Methods for Fixed Point Problems in Hilbert Spaces*, Lecture Notes in Mathematics, vol. **2057**, Springer-Verlag, Berlin, Heidelberg, Germany (2012).

[16] A. Cegielski and Y. Censor, Opial-type theorems and the common fixed point problem, in: H.H. Bauschke, R.S. Burachik, P.L. Combettes, V. Elser, D.R. Luke and H. Wolkowicz (Editors), *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, Springer-Verlag, New York, NY, USA, 2011, pp. 155–183.

[17] Y. Censor and A. Cegielski, Projection methods: an annotated bibliography of books and reviews, *Optimization* **64** (2015), 2343–2358.

[18] Y. Censor and R. Mansour, New Douglas-Rachford algorithmic structures and their convergence analyses, *SIAM Journal on Optimization* **26** (2016), 474–487.

[19] Y. Censor and S.A. Zenios, *Parallel Optimization: Theory, Algorithms, and Applications*, Oxford University Press, New York, NY, USA, 1997.

[20] P.L. Combettes, Solving monotone inclusions via compositions of nonexpansive averaged operators, *Optimization* **53** (2004), 475–504.

[21] M.N. Dao and H.M. Phan, Linear convergence of the generalized Douglas-Rachford algorithm for feasibility problems, *Journal of Global Optimization* (2018).
Available on arXiv at: https://arxiv.org/abs/1710.09814.

[22] E.D. Dolan and J.J. Moré, Benchmarking optimization software with performance profiles, *Mathematical Programming* **91** (2002), 201–213.

[23] J. Douglas and H.H. Rachford, On the numerical solution of the heat conduction problem in two and three space variables, *Transactions of the American Mathematical Society* **82** (1956), 421–439.

[24] J.M. Dye and S. Reich, Unrestricted iterations of nonexpansive mappings in Hilbert space, *Nonlinear Analysis* **18** (1992), 199–207.

[25] K. Goebel and S. Reich, *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings*, Marcel Dekker, New York and Basel, 1984.

[26] Z. Opial, Weak convergence of the sequence of successive approximations for nonexpansive mappings, *Bulletin of the American Mathematical Society* **73** (1967), 591–597.

[27] G. Pierra, Decomposition through formalization in a product space, *Mathematical Programming* **28** (1984), 96–115.