Constraint reduction reformulations for projection algorithms with applications to wavelet construction

Minh N. Dao*  Neil D. Dizon†  Jeffrey A. Hogan‡  Matthew K. Tam§

June 10, 2020

Abstract

We introduce a reformulation technique that converts a many-set feasibility problem into an equivalent two-set problem. The reformulation can be viewed as a variant of Pierra’s classical product space decomposition which reduces the dimension of the product space by replacing a pair of constraint sets with their intersection. We refer to this as the constraint reduction reformulation and use it to obtain constraint-reduced variants of well-known projection algorithms such as the Douglas–Rachford algorithm and the method of alternating projections, among others. We prove global convergence of constraint-reduced algorithms in the presence of convexity and local convergence in a nonconvex setting. In order to analyse convergence of the constraint-reduced Douglas–Rachford method, we generalize a classical result which guarantees that the composition of two projectors onto subspaces is a projector onto their intersection. Finally, we apply the constraint-reduced versions of Douglas–Rachford and alternating projections to solve the wavelet feasibility problems, and then compare their performance with their usual product variants.

Keywords: alternating projections · cyclic projections · Douglas–Rachford · fixed point iterations · multiresolution analysis · signal processing · wavelets

Mathematics Subject Classification (MSC 2020): 90C26 · 47H10 · 65K10 · 65T60

1. Introduction

A feasibility problem is the task of finding a point in the intersection of a finite family of sets. Formally, given sets $K_1, K_2, \ldots, K_r$ contained in a Hilbert space, the corresponding feasibility problem is to

$$\text{find } x^* \in K := \bigcap_{j=1}^{r} K_j.$$  \hspace{1cm} (1)

*Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia. E-mail: daoninh@gmail.com
†School of Mathematical and Physical Sciences, University of Newcastle, Callaghan 2308, Australia. E-mail: neil.dizon@newcastle.edu.au
‡School of Mathematical and Physical Sciences, University of Newcastle, Callaghan 2308, Australia. E-mail: jeff.hogan@newcastle.edu.au
§School of Mathematics and Statistics, The University of Melbourne, Parkville 3010, Australia. E-mail: matthew.tam@unimelb.edu.au
In the literature, projection algorithms are often used to solve feasibility problems. The method of alternating projections (MAP) [37] and the Douglas-Rachford (DR) algorithm [21] are well-known examples of projection algorithms which are applicable to two-set feasibility problems. Of these, the DR method has experienced sustained popularity because of its empirical potency in nonconvex settings [1, 2, 3, 9, 12, 17]. Although originally formulated for two-set feasibility problems, it has been extended to many-set feasibility problems by employing the cyclic DR method [10, 11], cyclically anchored DR method [8], cyclic generalized DR method [15, 16], or through Pierra’s product space reformulation [35]. The latter reformulation has the potential drawback of computational inefficiency when the number of constraint sets becomes large. This arises because each additional constraint in the original problem results in an additional product-dimension in the reformulation. A scheme to circumvent this is to replace a pair of constraints by their intersection. We formalize this as a new reformulation technique in Section 3 and use it to introduce variants of well-known projection algorithms. It is further favorable, but not required, if the pair of constraints $K_i$ and $K_j$ satisfy

$$P_{K_i}(K_j) \subseteq K_j$$  \hspace{1cm} (2) 

for some $i, j \in \{1, 2, \ldots, r\}$ with $i \neq j$, where $P_C$ denotes the projector onto a set $C$. As we will show in Section 4, exactly this property appears in the constraint sets arising in the feasibility approach to the wavelet construction.

More precisely, the construction of compactly supported and smooth multidimensional wavelets with orthogonal shifts and multiresolution structure has been recently formulated as a many-set feasibility problem [22, 23, 24] where the DR method, together with other projection algorithms and their many-set extensions, has been successfully employed. In this approach, properties of wavelets which are desirable in signal processing (e.g., compact support, smoothness) are treated as constraints alongside the conditions of multiresolution analysis (MRA) [30, 31], and intersection points yield the coefficients of the corresponding scaling and wavelet functions.

As additional properties such as real-valuedness, symmetry and cardinality [20] are added to the wavelet construction problem, the computational inefficiencies of the product space reformulation outlined above are realized due to the additional constraint sets. Fortunately, but also rather peculiarly to the structure of the wavelet feasibility problem, its constraint sets satisfy the property stated in (2). In particular, we show that the real-valuedness or the symmetry constraint may be combined with constraint sets arising from the conditions of MRA.

The goal of this paper is to present a constraint reduction reformulation for projection algorithms aimed at solving the feasibility problem (1). The main results appear in Section 3 where we formally introduce the reformulation, and use the framework of fixed point theory to study the operators obtained as a result of applying the reformulation to well-known projection algorithms. We give a global convergence analysis for the resulting variant of MAP and DR in the convex setting, and a local convergence analysis in a nonconvex setting. To do so, we extend a classical result regarding commutativity of two projectors on closed subspaces. As we show in Section 4, the reformulation can significantly reduce computational time.

The rest of the paper is organized as follows. Section 2 recalls relevant preliminaries and auxiliary results. Section 3 contains the constraint reduction reformulation together with other new results including the generalization of the classical result on the commutativity of two projectors. And finally in Section 4, we apply the reformulation to wavelet construction cast as a feasibility problem.
2. Preliminaries

Henceforth, we use $\mathcal{H}$ to denote a real Hilbert space endowed with an inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. For $x \in \mathcal{H}$ and $\delta \geq 0$, the closed ball centered at $x$ with radius $\delta$ is $B(x; \delta) := \{ z \in \mathcal{H} : \| z - x \| \leq \delta \}$. We use $I$ to denote the identity mapping on $\mathcal{H}$ which maps any point to itself. Moreover, if $T$ is an operator acting on a set $K$, we write $T(K) = \{ T(x) : x \in K \}$. We also denote the set of fixed points of the operator $T$ by $\text{Fix} T := \{ x \in \mathcal{H} : x = T(x) \}$ when $T$ is single-valued. Further, the product space $\mathcal{H}' = \mathcal{H} \times \mathcal{H} \times \cdots \times \mathcal{H}$ is also a real Hilbert space endowed with the inner product given by

$$\langle x, y \rangle = \sum_{j=1}^{r} \langle x_j, y_j \rangle$$

(3)

for all $x = (x_1, x_2, \ldots, x_n), y = (y_1, y_2, \ldots, y_n) \in \mathcal{H}'$.

2.1. Projectors, reflectors and projection methods

**Definition 2.1.** Let $C$ be a nonempty subset of $\mathcal{H}$. The distance function to $C$ is the function $d_C : \mathcal{H} \to \mathbb{R}$ defined by

$$d_C(x) = \inf_{z \in C} \| x - z \|$$

and the projector onto $C$ is the set-valued operator $P_C : \mathcal{H} \rightrightarrows C$ defined by

$$P_C(x) = \{ c \in C : \| x - c \| = d_C(x) \}.$$

The reflector with respect to $C$ is the set-valued operator $R_C : \mathcal{H} \rightrightarrows \mathcal{H}$ defined by

$$R_C := 2P_C - I.$$

An element of $P_C(x)$ is called a best approximation of $x$ from $C$ or a projection of $x$ onto $C$. Similarly, an element of $R_C(x)$ is called a reflection of $x$ with respect to $C$. If every point in $\mathcal{H}$ has at least one projection onto $C$, then $C$ is said to be proximinal.

Note that the sum in the definition of $R_C$ is understood in the sense of Minkowski set addition. In the case where $P_C$ is single-valued for all $x \in \mathcal{H}$, i.e., $P_C(x) = \{ u \}$ for some $u \in C$, we abuse notation by writing $P_C(x) = u$ and understand $P_C$ as a single-valued operator. It is a direct consequence of the definition that if the projector onto $C$ is single-valued, then the reflector with respect to $C$ is also single-valued. If the set $C$ is closed and convex, then $P_C$ is single-valued [19, Theorem 3.2], and the projections onto $C$ are easily characterized as follows.

**Proposition 2.2.** Let $C$ be a nonempty closed convex subset of $\mathcal{H}$, and $D$ be a nonempty closed affine subspace of $\mathcal{H}$. Then the following statements hold.

(i) $P_C(x) = p$ if and only if $p \in C$ and $\langle x - p, c - p \rangle \leq 0$ for all $c \in C$, 

(ii) $P_D(x) = p$ if and only if $p \in D$ and $\langle x - p, d - p \rangle = 0$ for all $d \in D$.

**Proof.** For (i), see [19, Theorem 4.1] or [6, Theorem 3.16]. For (ii), see [6, Corollary 3.22].

Projectors and reflectors form part of iterative algorithms called projection algorithms for solving feasibility problems. These algorithms exploit the structure of the individual sets which comprise the intersection that is the feasible region. These techniques iterate successively on the individual sets by applying projectors or reflectors, usually in a cyclic fashion.
The earliest formulation of projection methods dates back to the work of von Neumann [37] who showed that the sequence \((x_n)_{n \in \mathbb{N}}\) with \(x_0 \in \mathcal{H}\) and \(x_{n+1} = S_{A,B}(x_n)\), where
\[
s_{A,B} := P_A P_B,
\]
satisfies \(\lim_{n \to \infty} x_n = P_A \cap B(x_0)\) whenever \(A\) and \(B\) are closed subspaces. The operator \(S_{A,B}\) is sometimes called the alternating projection operator, and iterating \(S_{A,B}\) to obtain a projection onto the intersection is referred to as the method of alternating projections. The result was motivated by von Neumann’s return to the question of finding a point on \(A \cap B\) when \(P_A\) and \(P_B\) do not commute. Before then, it was only known that if \(P_A\) and \(P_B\) commute, then \(P_A P_B = P_{A \cap B}\) [37, Chapter XIII]. The following proposition provides several characterizations of this fact.

**Theorem 2.3.** If \(A\) and \(B\) are closed subspaces of \(\mathcal{H}\), then the following are equivalent:

(i) \(P_A P_B = P_B P_A\),

(ii) \(P_A(B) \subseteq B\),

(iii) \(P_B(A) \subseteq A\),

(iv) \(P_A P_B = P_A \cap B\).

**Proof.** See [19, Lemma 9.2].

In the next section, we generalize Theorem 2.3 to the case where \(A\) is a closed affine subspace and \(B\) is a proximinal subset. This generalization is key to our analysis of iterative algorithms.

When applied to two closed convex subsets, and if one of the set is compact or if either of the set is finite dimensional with the distance between them being attained, then strong convergence may be achieved [14]. A natural extension of MAP for many-set feasibility problems is called the method of cyclic projections which iterates by consecutively applying the projectors onto each of the constraint sets. This has a guaranteed convergence when the sets of interest are subspaces [25]. Moreover, the method weakly converges to a point on the intersection when the constraint sets are closed and convex [13]. While there are other projection methods, we confine ourselves mainly to alternating projections, and the DR algorithm which we now introduce.

**Definition 2.4.** Given two nonempty subsets \(A\) and \(B\) of \(\mathcal{H}\), the DR operator \(T_{A,B}\) is defined as
\[
T_{A,B} := I + R_B R_A.
\]

It is worth noting (see [7, Equations (20)–(23)]) that, if \(P_A\) is single-valued, then
\[
T_{A,B} = I - P_A + P_B R_A \quad \text{and} \quad P_A(\text{Fix } T_{A,B}) = A \cap B.
\]

If \(A\) and \(B\) are closed convex subsets of \(\mathcal{H}\) with \(A \cap B \neq \emptyset\), then, for any \(x_0 \in \mathcal{H}\), the sequence \((x_n)_{n \in \mathbb{N}}\) generated by \(x_{n+1} = T_{A,B}(x_n)\) converges weakly to a point \(x^* \in \text{Fix } T_{A,B}\), and the shadow sequence \((P_A(x_n))_{n \in \mathbb{N}}\) converges weakly to \(P_A(x^*) \in A \cap B\) [28, 36].

### 2.2. Convergence of fixed point iterations

Most of the projection algorithms that we have already mentioned can be cast as fixed point iterations. That is, for some starting point, a sequence is generated by repeated applications of the operator at hand, ideally to attain a fixed point in the limit. In this subsection, we will recall the relevant notions as well as the propositions necessary to establish convergence of projection algorithms.
Definition 2.5. Let $C \subseteq \mathcal{H}$ and let $T : C \rightarrow \mathcal{H}$. The mapping $T$ is said to be

(i) **nonexpansive** if, for all $x, y \in C$,

$$
\|T(x) - T(y)\| \leq \|x - y\|.
$$

(ii) **firmly nonexpansive** if, for all $x, y \in C$,

$$
\|T(x) - T(y)\|^2 + \|(I - T)(x) - (I - T)(y)\|^2 \leq \|x - y\|^2;
$$

(iii) **$\alpha$-averaged** if $\alpha \in (0, 1)$ and there exists a nonexpansive operator $R : C \rightarrow \mathcal{H}$ such that

$$
T = (1 - \alpha) I + \alpha R,
$$

or equivalently, if $\alpha \in (0, 1)$ and, for all $x, y \in C$,

$$
\|T(x) - T(y)\|^2 + \frac{1 - \alpha}{\alpha} \|(I - T)(x) - (I - T)(y)\|^2 \leq \|x - y\|^2.
$$

It follows from these definitions that $T$ is firmly nonexpansive if and only if it is $1/2$-averaged. Moreover, if $T$ is $\alpha$-averaged, then it is nonexpansive and also $\beta$-averaged with $\beta \in (\alpha, 1)$.

Example 2.6. Let $C$ be a nonempty closed convex subset of $\mathcal{H}$. Then

(i) $P_C$ is firmly nonexpansive.

(ii) $R_C$ is nonexpansive.

Proof. This follows from [6, Proposition 4.16 and Corollary 4.18].

It is also easy to establish that the composition of two nonexpansive operators is again nonexpansive. Also, other averaged maps may be obtained from convex combinations and compositions of already known averaged maps.

Proposition 2.7. Let $C \subseteq \mathcal{H}$ and let $T_j : C \rightarrow \mathcal{H}$ be $\alpha_j$-averaged for each $j \in J := \{1, 2, \ldots, r\}$. Define the operator $T : C_1 \times C_2 \times \cdots \times C_r \rightarrow \mathcal{H}^r$ by

$$
T(x) = (T_1(x_1), T_2(x_2), \ldots, T_r(x_r))
$$

for all $x = (x_1, x_2, \ldots, x_r) \in C_1 \times C_2 \times \cdots \times C_r$. Then following statements hold.

(i) $\sum_{j \in J} \lambda_j T_j$ is $\alpha$-averaged with $\alpha = \sum_{j \in J} \lambda_j \alpha_j$, whenever $\lambda_j > 0$ and $\sum_{j \in J} \lambda_j = 1$;

(ii) $T_r T_{r-1} \cdots T_1$ is $\alpha$-averaged with

$$
\alpha = \left(1 + \left(\sum_{j \in J} \frac{\alpha_j}{1 - \alpha_j}\right)^{-1}\right)^{-1}.
$$

Proof. For (i), see [6, Proposition 4.42]. For (ii), see [6, Proposition 4.46].

The next proposition establishes averagedness as well as characterizes the fixed point set of operators that are coordinate-wise averaged.

Proposition 2.8. Let $C_j \subseteq \mathcal{H}$ and let $T_j : C_j \rightarrow \mathcal{H}$ be $\alpha_j$-averaged for each $j \in J := \{1, 2, \ldots, r\}$. Define the operator $T : C_1 \times C_2 \times \cdots \times C_r \rightarrow \mathcal{H}^r$ by

$$
T(x) = (T_1(x_1), T_2(x_2), \ldots, T_r(x_r))
$$

for all $x = (x_1, x_2, \ldots, x_r) \in C_1 \times C_2 \times \cdots \times C_r$. Then following statements hold.
(i) $T$ is $\alpha$-averaged with $\alpha = \max_{j \in J} \alpha_j$.

(ii) Fix $T = \text{Fix}T_1 \times \text{Fix}T_2 \times \cdots \times \text{Fix}T_r$.

Proof. (i): Let $x, y \in C_1 \times C_2 \times \cdots \times C_r$. Since $\alpha = \max_{j \in J} \alpha_j \in (0, 1)$, we also have that $T_j$ is $\alpha$-averaged for each $j \in J$. Using Definition 2.5 and (3), we obtain

\[
\|T(x) - T(y)\|^2 + \frac{1 - \alpha}{\alpha} \| (I - T)(x) - (I - T)(y) \|^2
\]

\[
= \sum_{j=1}^r \|T_j(x_j) - T_j(y_j)\|^2 + \frac{1 - \alpha}{\alpha} \sum_{j=1}^r \| (I - T_j)(x_j) - (I - T_j)(y_j) \|^2
\]

\[
= \sum_{j=1}^r \left(\|T_j(x_j) - T_j(y_j)\|^2 + \frac{1 - \alpha}{\alpha} \| (I - T_j)(x_j) - (I - T_j)(y_j) \|^2 \right)
\]

\[
\leq \sum_{j=1}^r \|x_j - y_j\|^2
\]

\[
= \|x - y\|^2.
\]

Thus, $T$ is $\alpha$-averaged.

(ii): This is immediate from the definition. \qed

The following proposition provides a useful criterion for convergence of fixed point iterations.

**Proposition 2.9** (Opial’s theorem). Let $T : \mathcal{H} \to \mathcal{H}$ be $\alpha$-averaged with $\text{Fix}T \neq \emptyset$. Then, for any $x_0 \in \mathcal{H}$, the sequence $(x_n)_{n \in \mathbb{N}}$ generated by $x_{n+1} = T(x_n)$ converges weakly to a point $x^* \in \text{Fix}T$.

Proof. See [33], or set $\lambda_n = 1$ for all $n$ in [6, Proposition 5.16]. \qed

**Proposition 2.10.** Let $C$ be a nonempty subset of $\mathcal{H}$ and let $T_j : C \to C$ be $\alpha_j$-averaged for each $j \in \{1, 2, \ldots, r\}$ such that $\bigcap_{j=1}^r \text{Fix}T_j \neq \emptyset$. Then $\text{Fix}(T_r \cdots T_2 T_1) = \bigcap_{j=1}^r \text{Fix}T_j$.

Proof. See [6, Corollary 4.51]. \qed

### 2.3. Product space reformulation

The **product space reformulation** rewrites a many-set feasibility problem into a two-set feasibility problem [35]. Given $K_1, K_2, \ldots, K_r \subseteq \mathcal{H}$, with corresponding projectors $P_{K_1}, P_{K_2}, \ldots, P_{K_r}$, the sets $C$ and $D$ in the product Hilbert space $\mathcal{H}^r$ are defined by

\[
C := K_1 \times K_2 \times \cdots \times K_r \quad \text{and} \quad D := \{(x_1, x_2, \ldots, x_r) \in \mathcal{H}^r : x_1 = x_2 = \cdots = x_r\}.
\]

The $r$-set feasibility problem is equivalent to the two-set feasibility problem on $C$ and $D$ in the sense that

\[
x^* \in \bigcap_{j=1}^r K_j \iff x^* := (x^*, x^*, \ldots, x^*) \in C \cap D.
\]

Furthermore, the projectors onto $C$ and $D$ are given by

\[
P_C(x) = P_{K_1}(x_1) \times P_{K_2}(x_2) \times \cdots \times P_{K_r}(x_r) \quad \text{and}
\]

\[
(6a)
\]
for any $x = (x_1, x_2, \ldots, x_r) \in \mathcal{H}_r$; see, e.g., [6, Proposition 29.3 and Proposition 26.4(iii)]. Note that $D$ is a closed subspace of $\mathcal{H}$, and $C$ is a closed convex set if and only if $K_1, K_2, \ldots, K_r$ are closed and convex.

The product space reformulation allows us to use MAP and DR even when the number of constraint sets is greater than two.

3. Constraint reduction for feasibility problems

The main objective of this section is to introduce a constraint reduction reformulation for the $r$-set feasibility problem defined in (1). Before describing the new reformulation, we first prove a generalization of Theorem 2.3. This will be important in defining a particular case where the resulting operator arising from the constraint reduction reformulation for DR method will have a guaranteed convergence.

3.1. Projectors onto intersections

The following theorem extends Theorem 2.3, which applies for two closed subspaces of $\mathcal{H}$, to the setting of a closed affine subspace and a proximinal subset.

**Theorem 3.1.** Let $A$ be a closed affine subspace and $B$ be a proximinal subset of $\mathcal{H}$. Consider the following statements.

(i) $P_A(B) \subseteq B$,
(ii) $P_A(B) = A \cap B$,
(iii) $P_B(A) \subseteq A$,
(iv) $P_B(A) = A \cap B$,
(v) $P_B P_A = P_{A \cap B}$.

Then (i) $\implies$ (ii) $\implies$ (iii) $\implies$ (iv) $\implies$ (v). Moreover, if $B$ is convex, then all statements are equivalent.

**Proof.** (i) $\implies$ (ii): If $P_A(B) \subseteq B$, then $P_A(B) \subseteq A \cap B = P_A(A \cap B) \subseteq P_A(B)$, which yields $P_A(B) = A \cap B$.

(ii) $\implies$ (iii): Assume that $P_A(B) = A \cap B$. Take any $a \in A$, any $b \in P_B(a) \subseteq B$, and set $p = P_A(b) \in P_A(B) = A \cap B$. Since $A$ is a closed affine subspace, Proposition 2.2(ii) gives $\langle a - p, b - p \rangle = 0$, which implies that

$$||a - b||^2 = ||a - p||^2 + ||b - p||^2.$$  

As $b \in P_B(a)$ and $p \in B$, it holds that $||a - b|| \leq ||a - p||$. Combining with the above equality yields $||b - p||^2 = 0$, so $b = p \in A$. Since $a$ was chosen arbitrarily, we deduce that $P_B(A) \subseteq A$.

(iii) $\implies$ (iv): This follows by interchanging the roles of $A$ and $B$ in the proof of “(i) $\implies$ (ii)”. 

7
(iv) \(\implies\) (v): Assume that \(P_B(A) = A \cap B\). Fix \(x \in \mathcal{H}\), let any \(b \in P_BP_A(x)\) and any \(c \in P_{A \cap B}(x)\). Then \(b \in P_B(A) = A \cap B\) and also \(c \in A \cap B\). Setting \(a = P_A(x)\), we have \(b \in P_B(a)\). Since \(A\) is a closed affine subspace, Proposition 2.2(ii) implies that, for all \(z \in B\), \(\langle x - a, a - z \rangle = 0\), which yields
\[
\|x - b\|^2 = \|x - a\|^2 + \|a - b\|^2 \quad \text{and} \quad \|x - c\|^2 = \|x - a\|^2 + \|a - c\|^2.
\]

In addition, \(\|a - b\| \leq \|a - c\|\) since \(b \in P_B(a)\) and \(c \in B\). Therefore, \(\|x - b\| \leq \|x - c\| = d_{A \cap B}(x)\), which together with \(b \in A \cap B\) implies that \(\|x - b\| = \|x - c\|\) and \(b \in P_{A \cap B}(x)\). From this, we also obtain that \(\|a - c\| = \|a - b\| = d_B(a)\), and hence that \(c \in P_B(a)\), which means \(c \in P_BP_A(x)\).

Since \(x, b\) and \(c\) were chosen arbitrarily, we deduce that \(P_BP_A = P_{A \cap B}\).

(v) \(\implies\) (i): Assume that \(B\) is convex and that \(P_BP_A = P_{A \cap B}\). Fix \(x \in B\) and set \(y := P_BP_A(x) = P_{A \cap B}(x)\). Then \(y \in A \cap B\). Since \(A\) is a closed affine subspace and \(y \in A\), Proposition 2.2(ii) gives
\[
\langle P_A(x) - y, P_A(x) - x \rangle = 0. \tag{7}
\]
Moreover, since \(x \in B\) and \(y = P_BP_A(x)\), Proposition 2.2(i) yields
\[
\langle P_A(x) - y, x - y \rangle \leq 0. \tag{8}
\]
Adding (7) and (8), we obtain
\[
\langle P_A(x) - y, P_A(x) - y \rangle \leq 0,
\]
which implies that \(P_A(x) = y = P_BP_A(x)\). Thus, \(P_A(B) \subseteq P_B(P_A(B)) \subseteq B\).

We remark that if \(B\) is not convex, then we need not have \(P_A(B) \subseteq B\) when \(P_BP_A = P_{A \cap B}\). For a counterexample, we refer to Figure 1a. Here we take \(A = \{(x, x) : x \in \mathbb{R}\}\) and \(B = \{(x, y) \in \mathbb{R}^2 : 1 \leq x^2 + y^2 \leq 4\}\). It is easy to check that \(A\) is a subspace, \(B\) is nonconvex and \(P_BP_A = P_{A \cap B}\). However, \(x_0 \in B\) and \(P_A(x_0) \notin B\).

![Figure 1: Counterexamples](image)

Figure 1: Counterexamples: The figure on the left shows that \(P_A(x_0) \notin B\) even if \(P_BP_A = P_{A \cap B}\) for a nonconvex set \(B\). The figure on the right illustrates that even when \(B\) is convex and \(P_BP_A = P_{A \cap B}\), it does not follow that \(P_A\) and \(P_B\) commute.

Additionally, if \(B\) is convex and any of the equivalent statements is true, then it does not follow that \(P_A P_B = P_B P_A\). In particular, \(P_A P_B\) need not be equal to \(P_{A \cap B}\). To visualize
this, refer to Figure 1b where we redefined $B = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 4\}$ to show that $P_A P_B(x_0) = P_{A \cap B}(x_0)$.

**Example 3.2.** We now consider the following examples to illustrate the previous theorem.

(i) Let $A, B \subseteq \mathbb{R}^2$ where

\[ A = \{(x, x) : x \in \mathbb{R}\}, \quad B = \{(x, y) : |x| + |y| \leq 1\}. \]

In this example, $A$ is a subspace and $B$ is a closed convex subset of $\mathbb{R}^2$. Further, $P_B P_A \subseteq A$. Thus, all statements in Theorem 3.1 hold. In particular, $P_B P_A = P_{A \cap B}$. Refer to Figure 2a.

(ii) Let $A, B \subseteq \mathbb{R}^2$ where

\[ A = \{(x, 0) : x \in \mathbb{R}\}, \quad B = \left\{ (x, y) : \sqrt{|x|} + \sqrt{|y|} \leq 1 \right\}. \]

Here, $A$ is a subspace and $B$ is a nonconvex set. This is a particular example where $B$ is nonconvex but all of the statements in Theorem 3.1 still hold. Refer to Figure 2b.

Figure 2: The plot on the left shows that $P_B P_A = P_{A \cap B}$ which is equivalent to all other statements in Theorem 3.1 because $B$ is convex. The figure on the right is an example satisfying all statements in Theorem 3.1 even though $B$ is nonconvex.

### 3.2. Constraint reduction reformulation

The feasibility problem defined in (1) may be solved using a projection algorithm that is applicable to an $r$-set feasibility problem. In particular, one may employ the product DR and the product MAP by defining the product spaces $C$ and $D$ as in (4a) and (4b), respectively.

We now introduce a constraint reduction reformulation which also rewrites an $r$-set feasibility problem into a two-set. This is formalized in the following definition.
Definition 3.3 (Constraint reduction reformulation). Let \( K_1, K_2, \ldots, K_r \) be subsets of \( \mathcal{H} \). The constraint reduction reformulation of the \( r \)-set feasibility problem in (1) is the two-set feasibility problem given by

\[
\text{find } x := (x_1, x_2, \ldots, x_{r-1}) \in V \cap W \subseteq \mathcal{H}^{r-1},
\]

where \( V \) and \( W \) denote the reduced product space constraints given by

\[
V := K_1 \times K_2 \times \cdots \times K_{r-2} \times (K_{r-1} \cap K_r) \subseteq \mathcal{H}^{r-1}, \quad (9a)
\]

\[
W := \{(x_1, x_2, \ldots, x_{r-1}) \in \mathcal{H}^{r-1} : x_1 = x_2 = \cdots = x_{r-1}\}. \quad (9b)
\]

The associated mappings \( Q_V \) and \( P_W \) on \( \mathcal{H}^{r-1} \) are defined as

\[
Q_V(x) := P_{K_1}(x_1) \times P_{K_2}(x_2) \times \cdots \times P_{K_{r-2}}(x_{r-2}) \times P_{K_{r-1}}P_{K_r}(x_{r-1}), \quad (10a)
\]

\[
P_W(x) := \left( \frac{1}{r-1} \sum_{j=1}^{r-1} x_j, \frac{1}{r-1} \sum_{j=1}^{r-1} x_j, \ldots, \frac{1}{r-1} \sum_{j=1}^{r-1} x_j \right). \quad (10b)
\]

This new reformulation can be viewed as a modification of Pierra’s product space reformulation where two constraint sets in the product space \( C \) in (4a) are replaced with their intersection. In particular, \( K_{r-1} \cap K_r \) replaces \( K_{r-1} \times K_r \) in the definition of \( V \). The operator \( Q_V \) is defined to take the role of \( P_C \) by replacing \( P_{K_{r-1}} \) and \( P_{K_r} \) with the composition \( P_{K_{r-1}}P_{K_r} \). This operator is, in general, not the projector onto \( V \). Comparing further the two reformulations, we note that \( W \) is actually a subspace with dimension one less than that of \( D \) defined in (4b), and \( P_W \) is the projector onto \( W \) which takes the role of \( P_D \) defined in (6b). We also note that when \( K_r = \mathcal{H} \), the constraint reduction reformulation for \( r \) sets reduces to Pierra’s product reformulation for \( r-1 \) sets.

We remark that \( V, W \) and their associated mappings may be reformulated differently to allow for the intersection of other pairs of constraint sets. This will further cut down the dimension of the reduced product space constraints and the ambient Hilbert space. For simplicity of exposition, we focus on the set in Definition 3.3, but our results extend to the more general case.

As the following lemma shows, the constraint reduction reformulation still enjoys the equivalence statement (5) satisfied by the product space reformulation.

Lemma 3.4. Let \( K_1, K_2, \ldots, K_r \) be subsets of \( \mathcal{H} \), and consider \( V \) and \( W \) as defined in (9a) and (9b), respectively. Then

\[
x^* \in \bigcap_{j=1}^r K_j \iff x^* := (x^*, x^*, \ldots, x^*) \in V \cap W.
\]

Proof. If \( x^* \in \bigcap_{j=1}^r K_j \), then \( x^* \in K_j \) for all \( j \in \{1, 2, \ldots, r - 2\} \) and \( x^* \in K_{r-1} \cap K_r \). Consequently, \( x^* \in V \cap W \). The reverse implication is straightforward.

We now apply the constraint reduction reformulation to the method of alternating projections to deduce our first constraint reduced algorithm.

Constraint reduction for MAP

Definition 3.5. Let \( K_1, K_2, \ldots, K_r \) be proximinal subsets of \( \mathcal{H} \). The constraint-reduced MAP operator, denoted by \( S \), is defined by

\[
S := P_WQ_V,
\]

where \( Q_V \) and \( P_W \) are the operators defined in (10a) and (10b), respectively.
In the next theorem, we show global convergence of constraint-reduced MAP in the convex setting.

**Theorem 3.6.** Let $K_1, K_2, \ldots, K_r$ be closed convex subsets of $\mathcal{H}$ with nonempty intersection. Then the following statements hold.

(i) Fix $S = V \cap W = \{(x_1, x_2, \ldots, x_{r-1}) : x_1 = x_2 = \cdots = x_{r-1} \in \bigcap_{j=1}^r K_j\}$.

(ii) $S$ is 3/4-averaged. If, in addition, $P_{K_{r-1}}(K_r) \subseteq K_r$ and $K_r$ is affine, then $S$ is 2/3-averaged.

(iii) For any $x_0 \in \mathcal{H}^{-1}$, the sequence $(x_n)_{n \in \mathbb{N}}$ generated by $x_{n+1} = S(x_n)$ converges weakly to $x^* = (x^*, x^*, \ldots, x^*) \in V \cap W$ with $x^* \in \bigcap_{j=1}^r K_j$.

**Proof.** (i): We first note that $P_{K_j}$ is 1/2-averaged for each $j \in \{1, 2, \ldots, r\}$ by Example 2.6, and then that $P_{K_{r-1}}P_{K_r}$ is 2/3-averaged by Proposition 2.7(ii). Viewing $P_{K_j}$ as 2/3-averaged for each $j \in \{1, 2, \ldots, r-2\}$, we deduce from Proposition 2.8(i) that $Q_V$ is 2/3-averaged.

Since $\text{Fix } P_{K_j} = K_j$ for each $j \in \{1, 2, \ldots, r\}$ and $\bigcap_{j=1}^r K_j \neq \emptyset$, we have $\text{Fix } P_{K_{r-1}}P_{K_r} = K_{r-1} \cap K_r$ by Proposition 2.10, and then $\text{Fix } Q_V = K_1 \times \cdots \times K_{r-2} \times (K_{r-1} \cap K_r) = V$ by Proposition 2.8(ii). Noting that $W$ is also closed convex set, we have from Example 2.6 that $P_W$ is 1/2-averaged. Moreover, $\text{Fix } P_W = W$ and, by Lemma 3.4,

$$V \cap W = \left\{(x_1, x_2, \ldots, x_{r-1}) : x_1 = x_2 = \cdots = x_{r-1} \in \bigcap_{j=1}^r K_j \right\} \neq \emptyset.$$  

Applying Proposition 2.10 again to $Q_V$ and $P_W$ gives us $\text{Fix } S = \text{Fix } Q_V \cap \text{Fix } P_W = V \cap W$.

(ii): As shown in (i), $Q_V$ is 2/3-averaged and $P_W$ is 1/2-averaged. By applying Proposition 2.7(ii), $S$ is 3/4-averaged. Let us now assume that $P_{K_{r-1}}(K_r) \subseteq K_r$ and $K_r$ is affine. Theorem 3.1 yields $P_{K_{r-1}}P_{K_r} = P_{K_{r-1} \cap K_r}$. This makes $P_{K_{r-1}}P_{K_r}$ and $Q_V$ both 1/2-averaged. Consequently, $S$ is 2/3-averaged as given again by Proposition 2.7(ii).

(iii): We have from (i) that $\text{Fix } S = V \cap W \neq \emptyset$. Since $S$ is 3/4-averaged by (ii), invoking Proposition 2.9 yields the desired result. □

**Remark 3.7.** Under the assumptions of Theorem 3.6(ii), the operator $P_{K_{r-1}}P_{K_r}$ is not 1/2-averaged in general, even when $K_{r-1}$ and $K_r$ are both closed subspaces of $\mathbb{R}^2$ [5, Example 4.2.5]. As a consequence, the operator $Q_V$ is not 1/2-averaged in general.

**Constraint reduction reformulation for DR**

**Definition 3.8.** Let $K_1, K_2, \ldots, K_r$ be proximinal subsets of $\mathcal{H}$. The constraint-reduced DR operator, denoted by $T$, is defined by

$$T := I - P_W + Q_V R_W = I - P_W + Q_V (2P_W - I),$$

where $Q_V$ and $P_W$ are the operators defined in (10a) and (10b), respectively.

We reiterate that $Q_V$ is not necessarily a projector onto $V$, so that the classic convergence results for DR (or product DR) do not easily follow for $T$. Although a similar characterization of its fixed points still holds, we do not have a general convergence result analogous to Theorem 3.6 for the constraint-reduced DR. But in particular cases where we know more about the structure of $K_{r-1}$ and $K_r$, we can prove convergence.
Theorem 3.9. Let $K_1, K_2, \ldots, K_r$ be proximinal subsets of $\mathcal{H}$ with nonempty intersection. Suppose that $P_{K_{r-1}}(K_r) \subseteq K_r$. Then the following statements hold.

(i) $P_W(\text{Fix} T) = V \cap W$. In particular, if $(x_1,x_2,\ldots,x_{r-1}) \in \text{Fix} T$, then
\[
\frac{1}{r-1} \sum_{j=1}^{r-1} x_j \in \bigcap_{j=1}^{r} K_j.
\]

(ii) If $K_r$ is affine, then $T = (1+R_VR_W)/2$ coincides with the DR operator for $W$ and $V$.

(iii) If $K_1, \ldots, K_{r-1}$ are convex and $K_r$ is affine, then $T$ is firmly nonexpansive. Consequently, for any $x_0 \in \mathcal{H}^r$, the sequence $(x_n)_{n \in \mathbb{N}}$ generated by $x_{n+1} = T(x_n)$ converges weakly to a point $x^* = (x_1^*, x_2^*, \ldots, x_{r-1}^*) \in \text{Fix} T$. Moreover, writing $x_n = (x_{1,n}, x_{2,n}, \ldots, x_{r-1,n})$, the sequence $\left(\frac{1}{r-1} \sum_{j=1}^{r-1} x_{j,n}\right)_{n \in \mathbb{N}}$ converges weakly to
\[
\frac{1}{r-1} \sum_{i=1}^{r-1} x_{i}^* \in \bigcap_{j=1}^{r} K_j.
\]

Proof. (i): First, it follows from $P_{K_{r-1}}(K_r) \subseteq K_r$ that $P_{K_{r-1}}P_{K_r}(\mathcal{H}) = P_{K_{r-1}}(K_r) \subseteq K_{r-1} \cap K_r$, and thus,
\[Q_V(\mathcal{H}^r) \subseteq K_1 \times \cdots \times K_{r-2} \times K_{r-1} \cap K_r = V.
\]
Let $x \in \text{Fix} T$. Then $x \in T(x) = x - P_W(x) + Q_VR_W(x)$, which implies that $P_W(x) \in Q_VR_W(x) \subseteq Q_V(\mathcal{H}^r) \subseteq V$. Therefore, $P_W(x) \in V \cap W$. We deduce that $P_W(\text{Fix} T) \subseteq V \cap W$. On the other hand, it is straightforward to see that $V \cap W \subseteq \text{Fix} T$, which yields $V \cap W = P_W(V \cap W) \subseteq P_W(\text{Fix} T)$. Hence, $P_W(\text{Fix} T) = V \cap W$.

Now, if $x = (x_1,x_2,\ldots,x_{r-1}) \in \text{Fix} T$, then
\[
P_W(x) = \left(\frac{1}{r-1} \sum_{j=1}^{r-1} x_j, \frac{1}{r-1} \sum_{j=1}^{r-1} x_j, \ldots, \frac{1}{r-1} \sum_{j=1}^{r-1} x_j\right) \in P_W(\text{Fix} T) = V \cap W,
\]
and the conclusion follows from Lemma 3.4.

(ii): Assume that $K_r$ is affine. Since $K_r$ is proximinal, it is closed (see [19, Chapter 3]), and we have that $K_r$ is a closed affine subspace. Using Theorem 3.1, $P_{K_{r-1}}P_{K_r} = P_{K_{r-1} \cap K_r}$, and so $Q_V = P_V$ is the projector onto $V$. This implies that $T = (1 - P_W + P_VR_W) = (1+R_VR_W)/2$ is the DR operator for $W$ and $V$.

(iii): Assume that $K_1, \ldots, K_{r-1}$ are convex and $K_r$ is affine. By (ii), $T = (1+R_VR_W)/2$. Since every proximinal set in a Hilbert space is closed (see [19, Chapter 3]), we derive that $V$ is convex and closed. As $W$ is also convex and closed, by Example 2.6, $R_WR_V$ is nonexpansive and hence $T$ is 1/2-averaged, i.e., firmly nonexpansive.

Finally, since $\bigcap_{j=1}^{r} K_j \neq \emptyset$, Lemma 3.4 implies that $V \cap W \neq \emptyset$. The weak convergence of $(x_n)_{n \in \mathbb{N}}$ to $x^* \in \text{Fix} T$ follows from Proposition 2.9, see also [28, Theorem 1]. We also derive from [36, Theorem 1] that $(P_W(x_n))_{n \in \mathbb{N}}$ converges weakly to $P_W(x^*) \in P_W(\text{Fix} T) = V \cap W$. This completes the proof.

We wish to highlight that the constraint reduction reformulation for closed convex sets $K_1, K_2, \ldots, K_r$ with additional assumptions that $P_{K_{r-1}}(K_r) \subseteq K_r$ and that $K_r$ is a closed affine subspace, coincides with a non-standard application of the product space reformulation since $P_{K_{r-1}}P_{K_r} = P_{K_{r-1} \cap K_r}$ by Theorem 3.1. On the other hand, if we lift the convexity assumption
on at least the set $K_{r-1}$ but assuming it is proximal, then we still have $P_{K_{r-1}}P_{K_r} = P_{K_{r-1} \cap K_r}$ by Theorem 3.1. In this case, the projector is no longer guaranteed to be nonexpansive and thus the convergence results for constraint-reduced operators like $S$ or $T$ do not necessarily follow. As we will see in the next section, local convergence in nonconvex settings can still be guaranteed by replacing convexity with set regularity notions.

### 3.3. Local convergence of constraint reduced algorithms

In this subsection, $\mathcal{H}$ is finite-dimensional. Then a nonempty set in $\mathcal{H}$ is proximal if and only if it is closed; see [6, Corollary 3.15]. Let $C$ be a nonempty closed subset of $\mathcal{H}$. The limiting normal cone to $C$ at $x \in C$ (see [32, Definition 1.1(ii) and Theorem 1.6]) can be given by

$$N_C(x) = \left\{ \lim_{n \to +\infty} \lambda_n(z_n - x_n) : \lambda_n \geq 0, x_n \to x \text{ with } x_n \in P_C(z_n) \right\}.$$  

Recall from [27] that $C$ is superregular at a point $x \in C$ if, for any $\varepsilon > 0$, there exists $\delta > 0$ such that, for all $y, z \in C \cap \mathbb{B}(x; \delta)$ and all $u \in N_C(z)$,

$$\langle u, y - z \rangle \leq \varepsilon \|u\| \|y - z\|.$$  

A family of sets $\{K_1, K_2, \ldots, K_r\}$ in $\mathcal{H}$ is said to be

(i) linearly regular around $x \in \mathcal{H}$ if there exist $\kappa \geq 0$ and $\delta > 0$ such that, for all $z \in \mathbb{B}(x; \delta)$,

$$d_{K_1 \cap K_2 \cap \cdots \cap K_r}(z) \leq \kappa \max\{d_{K_1}(z), d_{K_2}(z), \ldots, d_{K_r}(z)\}.$$  

(ii) strongly regular at $x \in \mathcal{H}$ if

$$u_1 + u_2 + \cdots + u_r = 0 \text{ with } u_j \in N_{K_j}(x) \implies u_1 = u_2 = \cdots = u_r = 0.$$  

When $r = 2$, the strong regularity condition can be written as

$$N_{K_1}(x) \cap (-N_{K_2}(x)) = \{0\}.$$  

Interested readers can find more discussion on linear regularity and strong regularity in [4, 15, 16, 26, 27, 29].

**Proposition 3.10.** Let $\{K_1, K_2, \ldots, K_r\}$ be a family of sets in $\mathcal{H}$. The following statements hold.

(i) If $K_j$ is superregular at $x_j \in K_j$ for each $j \in \{1, \ldots, r\}$, then the product set $C := K_1 \times K_2 \times \cdots \times K_r \subseteq \mathcal{H}^r$ is superregular at $x := (x_1, x_2, \ldots, x_r) \in C$.

(ii) If $K_j$ is superregular at $x \in K := \bigcap_{j=1}^r K_j$ for each $j \in \{1, \ldots, r\}$ and $\{K_1, K_2, \ldots, K_r\}$ is strongly regular at every $z$ near $x$, then the intersection set $K$ is superregular at $x$.

**Proof.** Let $\varepsilon > 0$.

(i): Since $K_j$ is superregular at $x_j$, there exists $\delta_j > 0$ such that, for all $y, z \in K_j \cap \mathbb{B}(x_j; \delta_j)$ and all $u_j \in N_{K_j}(z_j)$, we have

$$\langle u_j, y_j - z_j \rangle \leq \varepsilon \|u_j\| \|y_j - z_j\|.$$  

(11)
Set $\delta = \min_{j=1,\ldots,r} \delta_j$. Let $y = (y_1, \ldots, y_r), z = (z_1, \ldots, z_r) \in C \cap B(x; \delta)$ and $u = (u_1, \ldots, u_n) \in N_C(z) = N_{K_1}(z) \times \cdots \times N_{K_r}(z)$. Then (11) followed by the Cauchy–Schwarz inequality yields

$$
\langle u, y - z \rangle = \sum_{j=1}^{r} \langle u_j, y_j - z_j \rangle \leq \varepsilon \sum_{j=1}^{r} \|u_j\| \|y_j - z_j\| \leq \varepsilon \left( \sum_{j=1}^{r} \|u_j\|^2 \right)^{1/2} \left( \sum_{j=1}^{r} \|y_j - z_j\|^2 \right)^{1/2} = \varepsilon \|u\| \|y - z\|
$$

which establishes the result.

(ii): Since $K_j$ is superregular at $x$, there exists $\delta > 0$ such that, for all $y, z \in K \cap B(x; \delta) \subseteq K_j \cap B(x; \delta)$ and all $u_j \in N_{K_j}(z)$, we have

$$
\langle u_j, y - z \rangle \leq \varepsilon \|u_j\| \|y - z\|
$$

Let $u \in N_K(z)$ be arbitrary. By assumption, shrinking $\delta$ if necessary, $\{K_1, K_2, \ldots, K_r\}$ is strongly regular at $z$ and, by [32, Corollary 3.37], $u = \sum_{j=1}^{r} u_j$ with some $u_j \in N_{K_j}(z)$. We then derive from [16, Proposition 2.4] and [32, Theorem 1.6] the existence of $\zeta > 0$ independent of $u_j$’s and $u$ such that

$$
\|u\| = \left\| \sum_{j=1}^{r} u_j \right\| \geq \zeta \sum_{j=1}^{r} \|u_j\|.
$$

Therefore,

$$
\langle u, y - z \rangle = \sum_{j=1}^{r} \langle u_j, y - z \rangle \leq \varepsilon \sum_{j=1}^{r} \|u_j\| \|y - z\| \leq \frac{\varepsilon}{\zeta} \|u\| \|y - z\|
$$

which completes the proof. \qed

**Proposition 3.11.** Let $\{K_1, K_2, \ldots, K_r\}$ be a family of sets in $\mathcal{H}$ and set

$$
C := K_1 \times K_2 \times \cdots \times K_r \quad \text{and} \quad D := \{(x_1, x_2, \ldots, x_r) \in \mathcal{H}^r : x_1 = x_2 = \cdots = x_r\}.
$$

Then the following statements hold.

(i) $\{K_1, K_2, \ldots, K_r\}$ is linearly regular around $x \in \mathcal{H}$ if and only if $\{C, D\}$ is linearly regular around $(x, x, \ldots, x) \in \mathcal{H}^r$.

(ii) $\{K_1, K_2, \ldots, K_r\}$ is strongly regular at $x \in \mathcal{H}$ if and only if $\{C, D\}$ is strongly regular at $(x, x, \ldots, x) \in \mathcal{H}^r$.

**Proof.** (i): Set $K := K_1 \cap K_2 \cap \cdots \cap K_r$. We first have that, for all $z = (z, z, \ldots, z) \in D$,

$$
d_{C \cap D}^2(z) = \inf_{y = (y, y, \ldots, y) \in C \cap D} \|z - y\|^2 = r \inf_{y \in K} \|z - y\|^2 = rd_{K}^2(z) \quad (12)
$$

and, since $P_C(z) = P_{K_1}(z) \times P_{K_2}(z) \times \cdots \times P_{K_r}(z)$,

$$
d_{C}^2(z) = d_{K_1}^2(z) + d_{K_2}^2(z) + \cdots + d_{K_r}^2(z). \quad (13)
$$

Assume that $\{K_1, K_2, \ldots, K_r\}$ is linearly regular around $x \in \mathcal{H}$. Then there exist $\kappa \geq 0$ and $\delta > 0$ such that, for all $z \in B(x; \delta)$, we have

$$
d_K(z) \leq \kappa \max\{d_{K_1}(z), d_{K_2}(z), \ldots, d_{K_r}(z)\}. \quad (14)
$$

14
Set $x := (x, x, \ldots, x) \in \mathcal{H}^r$ and let $y \in \mathbb{B}(x; \sqrt{t}\delta/2)$ and $z = P_D(y)$. Noting that $x \in D$, we have

$$||z - x|| \leq ||y - z|| + ||y - x|| \leq 2||y - x|| \leq \sqrt{t}\delta.$$ 

Thus, $z = (z, z, \ldots, z) \in D$ with $z \in \mathbb{B}(x; \delta)$. It follows from (12), (13), and (14) that

$$d_{C \cap D}(z) = \max\{d_{C}(z), d_{D}(z)\} \leq \sqrt{t}\kappa d_{C}(y) + ||y - z|| = \sqrt{t}\kappa(d_{C}(y) + d_{D}(y)).$$

We deduce that

$$d_{C \cap D}(y) \leq d_{C \cap D}(z) + ||y - z|| = d_{C \cap D}(z) + d_{D}(y) \leq \sqrt{t}\kappa d_{C}(y) + (1 + \sqrt{t})d_{D}(y) \leq (1 + 2\sqrt{t})\max\{d_{C}(y), d_{D}(y)\},$$

which implies the linear regularity of $\{C, D\}$ around $x$.

Conversely, assume that $\{C, D\}$ is linearly regular around $x = (x, x, \ldots, x) \in \mathcal{H}^r$, i.e., there exist $\kappa \geq 0$ and $\delta > 0$ such that, for all $z \in \mathbb{B}(x; \delta)$,

$$d_{C \cap D}(z) \leq \kappa \max\{d_{C}(z), d_{D}(z)\}.$$

Let $z \in \mathbb{B}(x; \delta/\sqrt{t})$. Then $z := (z, z, \ldots, z) \in D \cap \mathbb{B}(x; \delta)$ and the above inequality implies

$$d_{C \cap D}(z) \leq \kappa d_{C}(z).$$

Thus, by using (12) and (13), we deduce linear regularity of $\{K_1, K_2, \ldots, K_r\}$ around $x$.

(ii): For all $x = (x, x, \ldots, x) \in D$, we have from [32, Proposition 1.2] that

$$N_C(x) = N_{K_1}(x) \times N_{K_2}(x) \times \cdots \times N_{K_r}(x). \tag{15}$$

and from, e.g., [6, Proposition 26.4(ii)] that

$$N_D(x) = \{(u_1, u_2, \ldots, u_r) \in \mathcal{H}^r : u_1 + u_2 + \cdots + u_r = 0\}. \tag{16}$$

Assume that $\{K_1, K_2, \ldots, K_r\}$ is strongly regular at $x \in \mathcal{H}$. Set $x := (x, x, \ldots, x) \in \mathcal{H}^r$ and let $u \in N_C(x) \cap (-N_D(x))$. In view of (15), we can write $u = (u_1, u_2, \ldots, u_r)$ with $u_j \in N_{K_j}(x)$. Since $u \in -N_D(x)$, it follows from (16) that $u_1 + u_2 + \cdots + u_r = 0$. By the strong regularity of $\{K_1, K_2, \ldots, K_r\}$, we have $u_1 = u_2 = \cdots = u_r = 0$, and so $u = 0$. Altogether, we have $N_C(x) \cap (-N_D(x)) = \{0\}$, and thus $\{C, D\}$ is strongly regular at $x$.

Conversely, assume that $\{C, D\}$ is strongly regular at $x = (x, x, \ldots, x) \in \mathcal{H}^r$ and assume that

$$u_1 + u_2 + \cdots + u_r = 0 \quad \text{with} \quad u_j \in N_{K_j}(x).$$

Then $u := (u_1, u_2, \ldots, u_r) \in N_C(x)$ due to (15) and, in turn, (16) implies that $u \in -N_D(x)$, so $u \in N_C(x) \cap (-N_D(x)) = \{0\}$. We therefore have that $u_1 = u_2 = \cdots = u_r = 0$, which proves the strong regularity of $\{K_1, K_2, \ldots, K_r\}$ at $x$.

Recall that a sequence $(x_n)_{n \in \mathbb{N}}$ is said to converge $R$-linearly to a point $x^*$ if there exist $\rho \in [0, 1)$ and $\sigma > 0$ such that, for all $n \in \mathbb{N}$,

$$||x_n - x^*|| \leq \sigma \rho^n.$$

**Theorem 3.12.** Let $K_1, K_2, \ldots, K_{r-1}$ be closed subsets and $K_r$ be a closed affine subspace of $\mathcal{H}$ such that $\bigcap_{j=1}^r K_j \neq \emptyset$ and $P_{K_{r-1}}(K_r) \subseteq K_r$. Suppose that $K_1, \ldots, K_{r-2}$, and $K_{r-1} \cap K_r$ are superregular at a point $\overline{x} \in \bigcap_{j=1}^r K_j$. Set $\overline{x} := (\overline{x}, \overline{x}, \ldots, \overline{x}) \in \mathcal{H}^{r-1}$. Then the following statements hold.
If \( \{K_1, \ldots, K_{r-2}, K_{r-1} \cap K_r\} \) is linearly regular around \( \bar{x} \), then, whenever the starting point is sufficiently close to \( \bar{x} \), the sequence \( (x_n)_{n \in \mathbb{N}} \) generated by \( x_{n+1} \in S(x_n) \) converges \( R \)-linearly to a point \( x^* = (x^*, x^*, \ldots, x^*) \in V \cap W \) with \( x^* \in \bigcap_{j=1}^{r} K_j \).

(ii) If \( \{K_1, \ldots, K_{r-2}, K_{r-1} \cap K_r\} \) is strongly regular at \( \bar{x} \), then, whenever the starting point is sufficiently close to \( \bar{x} \), the sequence \( (x_n)_{n \in \mathbb{N}} \) generated by \( x_{n+1} \in T(x_n) \) converges \( R \)-linearly to a point \( x^* = (x^*, x^*, \ldots, x^*) \in V \cap W \) with \( x^* \in \bigcap_{j=1}^{r} K_j \).

Proof. We first derive from Lemma 3.4 that \( \bar{x} \in V \cap W \) and from Proposition 3.10(i) that \( V \) is superregular at \( \bar{x} \). Since \( P_{K_{r-1}}(K_r) \subseteq K_r \) and \( K_r \) is a closed affine subspace, Theorem 3.1 implies that \( P_{K_{r-1}} P_{K_r} = P_{K_{r-1} \cap K_r} \). In turn, \( Q_V = P_V \).

(i): We have \( S = P_W P_V \) and, by Proposition 3.11(i), \( \{V, W\} \) is linearly regular around \( \bar{x} \). Now, since \( V \) is superregular at \( \bar{x} \) and \( W \) is convex, applying [15, Corollary 5.12(i)(b)] with \( \lambda = \mu = \alpha = 1 \), we get the conclusion.

(ii): According to Theorem 3.9(ii), \( T = (1 + R_W R_V) / 2 \). By Proposition 3.11(ii), \( \{V, W\} \) is strongly regular at \( \bar{x} \). Noting that \( V \) is superregular at \( \bar{x} \) and \( W \) is convex, and using [15, Corollary 5.12(i)(a)] with \( \lambda = \mu = 2 \) and \( \alpha = 1 \) (see also [34, Theorem 4.3]), we complete the proof.

4. Application: wavelet construction

A wavelet \( \psi \) on the line is a function whose dyadic dilation and integer translations form an orthonormal basis for \( L^2(\mathbb{R}, \mathbb{C}) \). The utility of wavelets in analyzing and synthesizing signals relies on certain wavelet properties like compact support and regularity. The earliest examples of compactly supported smooth wavelets with orthonormal shifts were first achieved by Daubechies [18] through the multiresolution analysis (MRA) introduced by Mallat and Meyer [30, 31]. The methods employed by Daubechies are heavily reliant on complex analysis techniques that are not readily extendable to higher dimensions.

Recently, wavelet construction has been formulated as a feasibility problem [22, 23, 24]. The product space DR and MAP, along with other projection algorithms, have been successfully employed to solve the wavelet feasibility problem. The product DR was observed to yield both already known and unseen examples of wavelets on the line consistently. This approach has also been extended to produce nonseparable wavelets on the plane which required a higher number of constraint sets. In certain applications in signal and image processing, the efficiency of these wavelets requires additional properties including real-valuedness, symmetry, and cardinality [20]. Unfortunately, the inclusion of more constraints also requires additional product space dimensions. As the number of constraints gets large, the size of formulation becomes computationally intractable. It is on this ground that we want to evade an additional dimension by exploiting the property in (2) whenever it is viable.

We also remark that there are theoretical obstructions to obtain wavelets with the desired properties. Except for the case of Haar wavelet, there exists no symmetric, real-valued wavelets with orthonormal shifts, and compact support [18]. However, if we remove the real-valuedness condition, we may be able to obtain complex-valued scaling function and wavelet with perfect symmetry properties. Similarly, there exist no continuous, cardinal wavelets with compact support, and orthogonal shifts [38]. These theoretical obstructions may also be circumvented, without completely ruling out the desirable benefits of perfect symmetry or cardinality, by seeking for near-symmetry or near-cardinality [20].

In this section, we recall the wavelet feasibility problem and verify that a pair of its constraint
polynomials of the form \( m \) forcing \( \sigma \) to be compactly supported on \( \psi \). Shifts and dilates of \( \sigma \) is the "row swap" matrix. Additionally, a necessary condition for the orthonormality of the uniform samples and the coefficient matrices \( A \) determined by \( M \) is that \( \sigma \psi \). Finding the coefficients of these filters is key to constructing a \((\phi, \psi)\) pair.

MRA conditions and design criteria

A consistency condition arises from the definition of \( U(\xi) \), that is, \( U(\xi + 1/2) = \sigma U(\xi) \) where \( \sigma \) is the "row swap" matrix. Additionally, a necessary condition for the orthonormality of the shifts and dilates of \( \psi \) is that \( m_0(0) = 1 \) and \( U(\xi) \) is unitary almost everywhere. For \( \phi \) and \( \psi \) to be compactly supported on \([0, M - 1]\), we seek to impose that \( m_0 \) and \( m_1 \) be trigonometric polynomials of the form \( m_0(\xi) = \sum_{k=0}^{M-1} h_k e^{2\pi i k \xi} \) and \( m_1(\xi) = \sum_{k=0}^{M-1} g_k e^{2\pi i k \xi} \). Consequently, \( U(\xi) = \sum_{k=0}^{M-1} \sum_{j=0}^{M-1} A_k \psi^{2\pi i j \xi} \) with each \( A_k \in \mathbb{C}^{2 \times 2} \). The regularity criterion can be achieved by forcing \( \frac{d}{d\xi} U(0) \) to be diagonal for all \( 0 < l \leq \frac{M-2}{2} \). To ensure that we obtain real-valued scaling and wavelet functions, we require \( U(\xi) = U(-\xi) \). Finally, if \( U(\xi)^\dagger \) denotes a copy of \( U(\xi) \) with negated off-diagonal entries and \( \phi \) is symmetric about the center of support, then \( U(\xi) = e^{2\pi i(M-1)\xi} U(\xi)^\dagger \).

Discretisation by uniform sampling

The compact support condition allows us to write \( U(\xi) \) as a matrix-valued trigonometric polynomial of degree \( M - 1 \). And because a trigonometric polynomial of degree \( M - 1 \) is determined by \( M \) distinct points, we discretise \( U(\xi) \) by a uniform sampling at \( M \) points \( \{\xi_j\}_{j=0}^{M-1} \subseteq [0, 1] \). If \( U_j = U(\frac{j}{M}) \), then the sampling procedure produces an ensemble \( U = (U_0, U_1, \ldots, U_{M-1}) \in (\mathbb{C}^{2 \times 2})^M \) of matrices. Moreover, the coefficient matrices \( A_k \) may be obtained from the ensembles by an \( M \)-point discrete Fourier transform, that is,

\[
A_k = (\mathcal{F}_M U)_k = \frac{1}{M} \sum_{j=0}^{M-1} U_j e^{-2\pi i j k / M},
\]

which is also invertible to recover back \( U_j = (\mathcal{F}_M^{-1} A)_j \). This establishes a connection between the uniform samples and the coefficient matrices \( A_k \) of \( U(\xi) \).

Wavelet properties encoded on the ensembles

The consistency condition is imposed on the ensemble of samples to satisfy \( U_j + \frac{M}{2} = \sigma U_j \) for all \( j \in \{0, 1, \ldots, M - 1\} \). On the other hand, unitarity of each sample \( U_j = U(\frac{j}{M}) \) for \( j \in \)
In terms of the sample matrices and \( U \) for \( j \in \{0, 1, \ldots, M - 1\} \), is sufficient for \( U(\xi) \) to be unitary almost everywhere. However, it transpires that forcing \( U(\xi) \) to be unitary at 2M samples, uniformly chosen to be \( U(\frac{\pi}{2M}) \) and \( U(\frac{2\pi}{2M}) \), for \( j \in \{0, 1, \ldots, M - 1\} \), is sufficient for \( U(\xi) \) to be unitary almost everywhere.

Incidentally, given \( U = (U(\frac{\pi}{M}))_{j=0}^{M-1} \), the other M samples written to form an ensemble \( \tilde{U} \) may be obtained from \( U \) using \( \tilde{U} = F_M \chi_M F_M(U) \), where \( (\chi_M)_j = e^{2\pi ij/M} \) for \( j \in \{0, 1, \ldots, M - 1\} \).

In terms of the sample matrices \( U_j \), the regularity condition is imposed by forcing \( \sum_{j=0}^{M-1} j^\ell A_j \) to be diagonal, where

\[
\sum_{j=0}^{M-1} j^\ell A_j = \frac{1}{M} \sum_{j=0}^{M-1} \alpha_{\ell j} U_k \quad \text{and} \quad \alpha_{\ell j} = \frac{1}{M} \sum_{j=0}^{M-1} j^\ell e^{-2\pi \mathbf{i} j/M}.
\]

For real-valuedness, the ensembles must satisfy \( U_j = \overline{U_{M-j}} \) for \( j \in \{1, 2, \ldots, \frac{M}{2}\} \). Lastly, we require \( U_j = e^{2\pi \mathbf{i} (M-1)j/M} \overline{U_{M-j}} \) for all \( j \in \{1, 2, \ldots, \frac{M}{2}\} \) to meet the symmetry condition.

**Wavelet construction as a feasibility problem**

Let \((\mathbb{C}^{2 \times 2})_M^\sigma\) denote the collection of ensembles in \((\mathbb{C}^{2 \times 2})^M\) that satisfy the consistency condition. Further, let \( \mathcal{U}(2) \) denote the collection of all 2-by-2 unitary matrices. We define \( C_1, C_2, C_3, C_4^{(R)}, C_4^{(S)} \subseteq (\mathbb{C}^{2 \times 2})_M^\sigma \), where \( M \) is even and \( M \geq 4 \), as follows.

\[
C_1 := \left\{ U : U_0 = \begin{bmatrix} 1 & 0 \\ 0 & z \end{bmatrix}, |z| = 1, U_j \in \mathcal{U}(2), j \in \{0, 1, \ldots, M/2\} \right\}, \tag{18a}
\]

\[
C_2 := \left\{ U : (\mathcal{F}_M \chi_M \mathcal{F}_M^{-1}(U))_j \in \mathcal{U}(2), j \in \{0, 1, \ldots, M/2\} \right\}, \tag{18b}
\]

\[
C_3 := \left\{ U : \sum_{j=0}^{M-1} \alpha_{\ell j} U_k \in \text{diag}(\mathbb{C}^{2 \times 2}), 1 \leq \ell \leq (M-2)/2 \right\}, \tag{18c}
\]

\[
C_4^{(R)} := \left\{ U : U_j = \overline{U_{M-j}}, j \in \{1, 2, \ldots, M/2\} \right\}, \tag{18d}
\]

\[
C_4^{(S)} := \left\{ U : U_j = e^{2\pi \mathbf{i} (M-1)j/M} \overline{U_{M-j}}, j \in \{1, 2, \ldots, M/2\} \right\}. \tag{18e}
\]

**Problem 1** (Real-valued wavelets). *The problem to construct real-valued smooth orthogonal is to find an ensemble \( U = (U_0, \ldots, U_{M-1}) \in \bigcap_{k=1}^{3} C_k \cap C_4^{(R)} \subseteq (\mathbb{C}^{2 \times 2})_M^\sigma \).*

**Problem 2** (Symmetric wavelets). *The problem to construct symmetric smooth orthogonal wavelet is to find an ensemble \( U = (U_0, \ldots, U_{M-1}) \in \bigcap_{k=1}^{3} C_k \cap C_4^{(S)} \subseteq (\mathbb{C}^{2 \times 2})_M^\sigma \).*

We remark that \( C_1 \) and \( C_2 \) are nonconvex subsets of \((\mathbb{C}^{2 \times 2})_M^\sigma\), and every ensemble in both \( C_1 \) and \( C_2 \) will satisfy the unitarity condition. The subspaces \( C_3, C_4^{(R)}, \) and \( C_4^{(S)} \) are constraint sets for regularity, real-valuedness, and symmetry, respectively. The projectors onto \( C_1, C_2, \) and \( C_3 \) are computed in [22, Section 6.3] and those onto \( C_4^{(R)} \) and \( C_4^{(S)} \) are referred to [20, Section 3]. We will show that \( C_4^{(R)} \) and \( C_4^{(S)} \) are both invariant under the projector onto \( C_1 \) which we recall in the next proposition.

**Proposition 4.1.** Let \( \mathcal{U} = (U_0, U_1, \ldots, U_{M-1}) \in (\mathbb{C}^{2 \times 2})_M^\sigma \) and \( \tilde{U} := \{\tilde{U}_0, \tilde{U}_1, \ldots, \tilde{U}_{M-1}\} \in PC_1(\mathcal{U}) \). Suppose further that \( z \) is the \((2, 2)\)-entry of \( U_0 \) and that \( U_j = X_j \Sigma_j Y_j^* \) is a singular value decomposition for \( U_j \) where \( j \in \{1, 2, \ldots, M - 1\} \setminus \left\{ \frac{M}{2} \right\} \). Then

\[
\tilde{U}_0 = \begin{bmatrix} 1 & 0 \\ 0 & \tilde{z} \end{bmatrix}, \quad \tilde{U}_{M/2} = \sigma \tilde{U}_0, \quad \text{and} \quad \tilde{U}_j = X_j \Sigma_j Y_j^* \quad \text{for} \quad j \in \{1, 2, \ldots, M - 1\} \setminus \left\{ \frac{M}{2} \right\}.
\]
We emphasize that the ensembles in $P_{C_1}(U)$ do satisfy the consistency condition [22, Lemma 6.3.6]. We now verify two important relations among the constraint sets. These relations will allow us to an appropriate pair of constraint sets for applying the constraint reduction reformulation to the wavelet feasibility problem.

**Theorem 4.2.** Let $C_1$, $C_4^{(R)}$, and $C_4^{(S)}$ be as defined in (18). Then the following statements hold.

(i) $P_{C_1}(C_4^{(R)}) \subseteq C_4^{(R)}$.

(ii) $P_{C_1}(C_4^{(S)}) \subseteq C_4^{(S)}$.

**Proof.** (i): Let $U \in C_4^{(R)}$ and $\tilde{U} \in P_{C_1}(U)$. Then $U_j = U_{M-j}$ and $U_j = \sigma U_j + \frac{m}{2}$ for $j \in \{0, 1, \ldots, \frac{M}{2}\}$. Consequently, $U_0$ and $U_{M_j}$ have real entries. We deduce from Proposition 4.1 that $\tilde{U}_0$ and $\tilde{U}_{M_j}$ will also have real entries. For $j \in \{1, 2, \ldots, \frac{M}{2} - 1\}$, again by Proposition 4.1, $\tilde{U}_j = X_jY_j^*$, where $U_j = X_j\Sigma_jY_j^*$ is a singular value decomposition for $U_j$. Moreover,

$$U_{M-j} = \bar{U}_j = \bar{X}_j\Sigma_jY_j^* = \bar{X}_j\Sigma_jY_j^*$$

is a singular value decomposition for $U_{M-j}$, and so $\tilde{U}_{M-j} = \bar{X}_jY_j^* = \bar{U}_j$. Therefore, $\tilde{U} \in C_4^{(R)}$.

(ii): Let $U \in C_4^{(S)}$ and $\tilde{U} \in P_{C_1}(U)$. Then $U_j = e^{-2\pi i(M-1)j/M}U_{M-j}$ and $U_j = \sigma U_j + \frac{m}{2}$ for $j \in \{0, 1, \ldots, \frac{M}{2}\}$. In particular, $U_0 = U_0^\dagger$ and $U_{M_j} = -U_{M_j}^\dagger$. We know from Proposition 4.1 that $\tilde{U}_0$ and $\tilde{U}_{M_j}$ are diagonal and deduce that $\tilde{U}_0 = \tilde{U}_0^\dagger$ and $\tilde{U}_{M_j} = \sigma \tilde{U}_0 = -(\sigma \tilde{U}_0)^\dagger = -\tilde{U}_{M_j}^\dagger$. For $j \in \{1, 2, \ldots, \frac{M}{2} - 1\}$, we also learn from Proposition 4.1 that $\tilde{U}_j = X_jY_j^*$, where $U_j = X_j\Sigma_jY_j^*$ is a singular value decomposition for $U_j$. Denote $s = e^{-2\pi i(M-1)j/M}$ and $\tau = \text{diag}(-1, 1) \in \mathbb{C}^{2 \times 2}$. Then

$$U_{M-j} = e^{-2\pi i(M-1)j/M}U_j^\dagger = s\tau U_j^{\tau^*} = (s\tau X_j)\Sigma_j(Y_j^*\tau^*)$$

is a singular value decomposition for $U_{M-j}$ since $|s| = 1$. Hence,

$$\tilde{U}_{M-j} = (s\tau X_j)(Y_j^*\tau^*) = s(\tau X_jY_j^*\tau^*) = s(X_jY_j^*)^\dagger = e^{-2\pi i(M-1)j/M}\tilde{U}_j^\dagger,$$

and we deduce that $\tilde{U} \in C_4^{(S)}$. 

The results in Theorem 4.2 further justify our choice of $C_1$ and $C_4^{(R)}$ as the pair of constraints to replace with their intersection for constraint reduction reformulation of Problem 1. Similarly, the pair of $C_1$ and $C_4^{(S)}$ is the natural choice for Problem 2. We will solve these problems in the next subsection.

We note that a solution $U = (U_0, \ldots, U_{M-1})$ of Problem 1 or 2 contains the $M$ samples of $U(\xi)$ from which we recover the coefficients $A_k$ using (17). Consequently, the coefficients of the scaling filter $m_0(\xi)$ and wavelet filter $m_1(\xi)$ may be easily pulled out from the $A_k$’s. Through the *cascade algorithm* applied to the coefficients of $m_0$ and $m_1$, we may be able to plot the scaling function $\phi$ and wavelet $\psi$, respectively.
4.2. Numerical experiments

The wavelet feasibility problems defined in Problems 1–2 can be straightforwardly reformulated to a two-set feasibility problem using the product space reformulation defined in Section 2.3. The product DR and MAP are then employable to solve the two-set problem. Alternatively, we may apply the constraint-reduction reformulation to the problems at hand. We abuse notation by consistently denoting the reduced product space constraints as in Definition 3.3 for both problems.

**Constraint-reduction reformulation for Problem 1:** The product space constraints for obtaining real-valued wavelets are defined by

\[
V := \left( C_1 \cap C_4^{(R)} \right) \times C_2 \times C_3 \subseteq \left( (\mathbb{C}^{2\times2})^M \right)^3,
\]

\[
W := \left\{ (U_j)_{j=1}^3 \in \left( (\mathbb{C}^{2\times2})^M \right)^3 : U_1 = U_2 = U_3 \right\}.
\]

**Constraint-reduction reformulation for Problem 2:** The product space constraints for obtaining symmetric wavelets are defined by

\[
V := \left( C_1 \cap C_4^{(S)} \right) \times C_2 \times C_3 \subseteq \left( (\mathbb{C}^{2\times2})^M \right)^3,
\]

\[
W := \left\{ (U_j)_{j=1}^3 \in \left( (\mathbb{C}^{2\times2})^M \right)^3 : U_1 = U_2 = U_3 \right\}.
\]

The associated operators \( Q_V \) and \( P_W \) for both the two new problems are defined similar to what appeared in Definition 3.3.

We now solve the wavelet problem corresponding to the case where \( M = 6 \). We compare the performance of product DR against constraint-reduced DR, and the performance of product MAP against constraint-reduced MAP. Henceforth, we let \( (x_n)_{n \in \mathbb{N}} \) be the sequence of iterates generated by a projection algorithm. In all numerical implementations, we consider a tolerance \( \varepsilon := 10^{-9} \). For the DR variant, we use the stopping criterion given by \( \| Q_V P_W (x_n) - P_W (x_n) \| < \varepsilon \) which when satisfied indicates that \( P_W (x_n) \) can be declared as a feasible point. Similarly for constraint-reduced MAP, we set a stopping criterion \( \| Q_V (x_n) - x_n \| < \varepsilon \) to decide that the iterate \( x_n \) lies on the intersection of \( V \) and \( W \). We consider a projection algorithm to have solved our feasibility problem if and when it attains a point that satisfies the stopping criterion within the cutoff of 50,000 iterates. For our numerical results, we provide statistics on the number of iterations which we mainly consider as performance measure. Additionally, we comment on the versatility of an algorithm in tackling the nonconvex wavelet feasibility problem by counting the number of times it solves a particular problem initialized at 1,000 ensembles that satisfy the consistency condition and with complex entries having real and imaginary parts chosen from uniformly distributed random number in the interval \((0,1)\).

**Real-valued wavelets**

To construct real-valued wavelets, we need to deal with the first feasibility problem. We employ the product DR and MAP, and their constraint-reduced variants to solve the appropriate version of Problem 1. We only compare the performance of product DR with that of constraint-reduced DR, doing the same for MAP. Thus, we are essentially comparing the robustness of the product space and constraint reduction reformulations.

Table 1 shows the performance of DR in solving Problem 1 using product space and constraint reduction reformulations. For this problem, the product DR solved more cases than...
constraint-reduced DR. In cases where both algorithms solved the feasibility problem, the constraint-reduced DR used up lesser number of iterations 97% of the time. This suggests that the constraint-reduced DR outperforms product DR in terms of the number of iterations. This claim is supported by the computed mean and median number of iterations for the contraint-reduced DR which are less than that of the product DR.

| Algorithm                  | cases solved | solved alone | solved by both | when solved by both |
|----------------------------|--------------|--------------|----------------|---------------------|
|                            | wins | mean  | median         |                     |
| Product DR                 | 619  | 262   | 357            | 11                  |
| Constraint-reduced DR      | 497  | 140   | 357            | 346                 |

Table 1: Statistics on the performance of product DR and constraint-reduced DR for wavelet feasibility problems with real-valuedness constraint.

| Algorithm                  | cases solved | solved alone | solved by both | when solved by both |
|----------------------------|--------------|--------------|----------------|---------------------|
|                            | wins | mean  | median         |                     |
| Product MAP                | 264  | 95    | 169            | 0                   |
| Constraint-reduced MAP     | 235  | 66    | 169            | 169                 |

Table 2: Statistics on the performance of product MAP and constraint-reduced MAP for wavelet feasibility problems with real-valuedness constraint.

Similarly, Table 2 summarizes our results when MAP is used to solve the feasibility problem through product space and constraint reduction reformulations. In our statistics, the two algorithms performed closely in terms of their efficacy to solve the feasibility problem – with product MAP solving few more problems than the constraint-reduced version. In cases where both algorithms solved the feasibility problem, the constraint-reduced MAP always has lesser number of iterations, outperforming the product MAP. These are reflected in the computed mean and median number of iterations for both algorithms.

21
In contradistinction, the MAP is not really as robust as DR in solving the feasibility problem at hand as suggested by the total number of cases that each solved. Figure 3 shows an example of real-valued scaling function-wavelet pair generated by solving Problem 1. This wavelet is exactly Daubechies’ $\psi_3$ wavelet which is known to have the maximal number of vanishing moments for its length of support [18, Chapter 5]. Other solutions may be obtained by lowering the requirement on regularity.

**Symmetric wavelets**

For constructing symmetric wavelets, we solve the second feasibility problem. We employ the product DR and MAP, and their constraint-reduced variants to solve the product space and constraint-reduced versions of Problem 2, respectively.

![Figure 4: Complex-valued compactly supported smooth scaling function and wavelet with symmetry properties obtained by solving Problem 2.](image)

Table 3 summarizes the performance of DR in solving Problem 2 using product space and constraint reduction reformulations. The constraint-reduced DR solved every test case while product DR failed in 44 cases. In instances where both algorithms converged, the constraint-reduced DR used up lesser number of iterations 78% of the time. This suggests that the constraint-reduced DR outperforms product DR as also reflected in the computed mean and median number of iterations.

| Algorithm               | cases solved | solved alone | solved by both | when solved by both |
|-------------------------|--------------|--------------|----------------|---------------------|
|                         | wins | mean | median |
| Product DR              | 956  | 206  | 6969  | 5941    |
| Constraint-reduced DR   | 1000 | 750  | 5439  | 4754    |

Table 3: Statistics on the performance of product DR and constraint-reduced DR for wavelet feasibility problems with symmetry constraint.

Similarly, Table 4 highlights our results for cases where MAP is used to solve the feasibility problem through product space and constraint reduction reformulations. In our statistics, the two algorithms solved all test cases with constraint-reduced MAP incurring lesser number of
Algorithm cases solved solved alone solved by both when solved by both wins mean median
Product MAP 1000 0 1000 2 5528 5648
Constraint-reduced MAP 1000 0 1000 998 4157 4232

Table 4: Statistics on the performance of product MAP and constraint-reduced MAP for wavelet feasibility problems with symmetry constraint.

iterations 99% of the time. This suggests that constraint-reduced MAP outperforms product MAP in this sense as can also be seen in the computed mean and median number of iterations for both algorithms.

It is also noteworthy that based on our statistics, MAP’s variants are more effective than DR’s in finding symmetric wavelets. Figure 4 shows an example of a real-valued scaling function-wavelet pair generated by solving Problem 2.

5. Conclusions

We have introduced a constraint reduction reformulation for converting many-set feasibility problems into two-set problems. It provides an equivalent formulation of many-set feasibility problems by defining the reduced product constraints with lower dimension than that of Pierra’s product space technique. Our new reformulation gives rise to constraint-reduced variants of any projection algorithm that can be used to solve two-set feasibility problems. We have presented a global convergence analysis for the constraint-reduced variants of DR and MAP in the convex setting, and a local convergence analysis in a nonconvex setting. In carrying out the analysis for the constraint-reduced DR, we have generalized a well-known result which guarantees that the composition of two projectors onto subspaces is again a projector onto the intersection. Even when the constraint sets do not possess the additional structure required, the constraint-reduced variants of projection algorithms still serve as useful heuristics for solving nonconvex feasibility problems.

The required property among the constraint sets for the convergence of constraint-reduced DR appear exactly in the wavelet feasibility problems so it provided us a suitable venue for numerical implementations of the new reformulation technique. In certain cases, the performance of constraint-reduced DR and MAP has been seen as improvement over their usual product variants.

Acknowledgments

The authors are thankful for the helpful insight of Scott Lindstrom on Theorem 3.1. The authors were partially supported by the Australian Research Council through grants DP160101537 (MND, NDD, and JAH), DP190100555 (MND), and DE200100063 (MKT).

References

[1] F. J. Aragón Artacho, J. M. Borwein, and M. K. Tam, Douglas–Rachford feasibility methods for matrix completion problems, ANZIAM J., 55 (2014), pp. 299–326.
[2] _____, Recent results on Douglas–Rachford methods for combinatorial optimization problems, J. Optim. Theory App., 163 (2014), pp. 1–30.

[3] F. J. Aragón Artacho, R. Campoy, and M. K. Tam, The Douglas–Rachford algorithm for convex and nonconvex feasibility problems, Math. Method Oper. Res., (2019), pp. 1–40.

[4] H. H. Bauschke and J. M. Borwein, On projections algorithms for solving convex feasibility problems, SIAM Rev., 38 (1996), pp. 367–426.

[5] H. H. Bauschke, J. M. Borwein, and A. S. Lewis, The method of cyclic projections for closed convex sets in Hilbert space, Contemp. Math., 204 (1997), pp. 1–38.

[6] H. H. Bauschke and P. L. Combettes, Convex Analysis and Monotone Operator Theory in Hilbert Spaces, Springer, Cham, 2017.

[7] H. H. Bauschke and M. N. Dao, On the finite convergence of the Douglas–Rachford algorithm for solving (not necessarily convex) feasibility problems in Euclidean spaces, SIAM J. Optim., 27 (2017), pp. 507–537.

[8] H. H. Bauschke, D. Noll, and H. M. Phan, Linear and strong convergence of algorithms involving averaged nonexpansive operators, J. Math. Anal. Appl., 421 (2015), pp. 1–20.

[9] 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, Springer, New York, 2011, pp. 93–109.

[10] J. M. Borwein and M. K. Tam, A cyclic Douglas–Rachford iteration scheme, J. Optim. Theory Appl., 160 (2014), pp. 1–29.

[11] _____, The cyclic Douglas–Rachford feasibility method for inconsistent feasibility problems, J. Nonlinear Convex A., 16 (2015), pp. 537–584.

[12] _____, Reflection methods for inverse problems with applications to protein conformation determination, in Generalized Nash Equilibrium Problems, Bilevel Programming and MPEC, Springer, 2017, pp. 83–100.

[13] L. Bregman, The method of successive projection for finding a common point of convex sets, Sov. Math. Dok., 6 (1965), pp. 688–692.

[14] W. Cheney and A. A. Goldstein, Proximity maps for convex sets, P. Am. Math. Soc., 10 (1959), pp. 448–450.

[15] M. N. Dao and H. M. Phan, Linear convergence of the generalized Douglas–Rachford algorithm for feasibility problems, J. Global Optim., 72 (2018), pp. 443–474.

[16] _____, Linear convergence of projection algorithms, Math. Oper. Res., 44 (2019), pp. 715–738.

[17] M. N. Dao and M. K. Tam, A Lyapunov-type approach to convergence of the Douglas–Rachford algorithm for a nonconvex setting, J. Global Optim., 73 (2019), pp. 83–112.

[18] I. Daubechies, Ten Lectures on Wavelets, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1992.

[19] F. Deutsch, Best Approximation in Inner Product Spaces, Springer-Verlag, New York, USA, 2001.
[20] N. D. Dizon, J. A. Hogan, and J. D. Lakey, *Optimization in the construction of nearly cardinal and nearly symmetric wavelets*, in 13th International conference on Sampling Theory and Applications (SampTA), IEEE, 2019, pp. 1–4.

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

[22] D. J. Franklin, *Projective Algorithms for Non-separable Wavelets and Clifford Fourier Analysis*, PhD thesis, The University of Newcastle (Australia), 2018.

[23] D. J. Franklin, J. A. Hogan, and M. K. Tam, *Higher-dimensional wavelets and the Douglas-Rachford algorithm*, in 13th International Conference on Sampling Theory and Applications (SampTA), IEEE, 2019, pp. 1–4.

[24] ——, *A Douglas–Rachford construction of non-separable continuous compactly supported multidimensional wavelets*, arXiv preprint arXiv:2006.03302, (2020).

[25] I. Halperin, *The product of projection operators*, Acta. Sci. Math. (Szeged), 23 (1962), pp. 96–99.

[26] A. Y. Kruger, *About regularity of collections of sets*, Set-Valued Anal., 14 (2006), pp. 187–206.

[27] A. S. Lewis, D. R. Luke, and J. Malick, *Local linear convergence for alternating and averaged nonconvex projections*, Found. Comput. Math., 9 (2009), pp. 485–513.

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

[29] D. R. Luke, N. H. Thao, and M. K. Tam, *Quantitative convergence analysis of iterated expansive, set-valued mappings*, Math. Oper. Res., 43 (2018), pp. 1143–1176.

[30] S. Mallat, *Multiresolution approximations and wavelet orthonormal bases of $L_2(\mathbb{R})$*, T. A. Math. Soc., 315 (1989), pp. 69–87.

[31] Y. Meyer, *Wavelets and Operators*, Cambridge University Press, Cambridge, UK, 1993.

[32] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation I*, Springer, Berlin, 2006.

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

[34] H. M. Phan, *Linear convergence of the Douglas–Rachford method for two closed sets*, Optimization, 65 (2016), pp. 369–385.

[35] G. Pierra, *Decomposition through formalization in a product space*, Math. Program., 28 (1984), pp. 96–115.

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

[37] J. von Neumann, *Functional Operators Volume II: The Geometry of Orthogonal Spaces*, Princeton University Press, New Jersey, USA, 1950.

[38] X. Xia and Z. Zhang, *On sampling theorem, wavelets, and wavelet transforms*, IEEE T. Signal Proces., 41 (1993), pp. 3524–3535.