The Block-wise Circumcentered–Reflection Method

Roger Behling · J.-Yunier Bello-Cruz · Luiz-Rafael Santos

Abstract The elementary Euclidean concept of circumcenter has recently been employed to improve two aspects of the classical Douglas–Rachford method for projecting onto the intersection of affine subspaces. The so-called circumcentered–reflection method is able to both accelerate the average reflection scheme by the Douglas–Rachford method and cope with the intersection of more than two affine subspaces. We now introduce the technique of circumcentering in blocks, which, more than just an option over the basic algorithm of circumcenters, turns out to be an elegant manner of generalizing the method of alternating projections. Linear convergence for this novel block-wise circumcenter framework is derived and illustrated numerically. Furthermore, we prove that the original circumcentered–reflection method essentially finds the best approximation solution in one single step if the given affine subspaces are hyperplanes.

Keywords Accelerating convergence · Best approximation problem · Circumcenter scheme · Douglas–Rachford method · Linear and finite convergence · Method of alternating projections.

Mathematics Subject Classification (2010) 49M27 · 65K05 · 65B99 · 90C25

1 Introduction

We consider the important feasibility problem of projecting onto the intersection of affine subspaces, frequently also referred to as *best approximation problem*. Let $\{U_i\}_{i \in \mathbf{I}}$ be a family of finitely many affine subspaces in $\mathbb{R}^n$ with $\mathbf{I} := \{1, 2, \ldots, m\}$ and m fixed (we require no relation between n and m). The intersection of the family is denoted by $S := \bigcap_{i \in \mathbf{I}} U_i$ (which we assume nonempty) and the problem we are interested in consists of projecting a given point $z \in \mathbb{R}^n$ onto S. Equivalently,

$$\min_{s \in S} \|z - s\|. \tag{1}$$

Here and throughout the paper, $\langle \cdot, \cdot \rangle$ stands for the Euclidean inner product and $\| \cdot \|$ is the induced norm. The best approximation problem (1) has the unique solution $P_S(z)$, where P_S denotes the Euclidean projection onto S. Problem (1) can, of course, be rewritten as a convex quadratic program with objective function $\frac{1}{2}\|z - s\|^2$ and equality constraints, as each U_i is an affine subspace. Note that S itself is an affine subspace. Note also that the classical problem of finding the least-norm solution of a system of linear equations is a particular case of (1).

Roger Behling ✉
School of Applied Mathematics, Fundação Getúlio Vargas
Rio de Janeiro, RJ – 22250-900, Brazil. E-mail: rogerbehling@gmail.com

J.-Yunier Bello-Cruz
Department of Mathematical Sciences, Northern Illinois University.
DeKalb, IL – 60115-2828, USA. E-mail: yunierbello@niu.edu

Luiz-Rafael Santos
Department of Mathematics, Federal University of Santa Catarina.
Blumenau, SC – 88040-900, Brazil. E-mail: l.r.santos@ufsc.br

Reflection and projection type methods are celebrated tools for solving a variety of feasibility problems, including (1), and they remain trendy due to their balance between good performance and simplicity (see, *e.g.*, [6]). Probably the two most famous and standard among these methods are the Douglas–Rachford method (DRM), or averaged alternating reflection method (see, *e.g.*, [3]); and the method of alternating projections (MAP) which is also known as von Neumann's or Kaczmarz' algorithm (see, *e.g.*, [5, 29]). Upon our ideas presented in [13, 14], we devote this work to study a circumcenter type method related to both DRM and MAP.

Suitable DRM and MAP schemes determine the solution of the best approximation problem (1). DRM and MAP only use knowledge provided by projections onto individual sets, which often leads to a desirable low computational cost per iteration. Nonetheless, slow convergence due to zig-zag or spiral behavior are usually inherent to these classical methods (see, *e.g.*, [3, 4, 7]). In order to minimize spiralness of Douglas–Rachford sequences to a certain extent, we have introduced the circumcentered–reflection method (CRM). This was firstly done in [13] for problem (1) with two sets, that is, $m = 2$. In this case, if we have a current iterate say $z \in \mathbb{R}^n$, DRM moves us to $z_{DRM} := \frac{1}{2}(\mathrm{Id} + R_{U_2}R_{U_1})(z)$, whereas MAP provides $z_{MAP} := P_{U_2}P_{U_1}(z)$. The symbol Id denotes the identity operator and $R_{U_i} := 2P_{U_i} - \mathrm{Id}$ is the reflection operator onto U_i. We proposed the iteration $z_{CRM} := \mathrm{circumcenter}\{z, R_{U_1}(z), R_{U_2}R_{U_1}(z)\}$, where z_{CRM} fulfills two properties: **(i)** it lies on the affine subspace defined by z, $R_{U_1}(z)$ and $R_{U_2}R_{U_1}(z)$, which we denote by $\mathrm{aff}\{z, R_{U_1}(z), R_{U_2}R_{U_1}(z)\}$ and, **(ii)** z_{CRM} is equidistant to z, $R_{U_1}(z)$ and $R_{U_2}R_{U_1}(z)$, therefore the use of the term *circumcenter*.

The resulting algorithm significantly outperforms DRM and MAP numerically as presented in [13]. This numerical performance of CRM, together with the deficiency of DRM in dealing with more than two sets (see [1, Example 2.1] and some modifications [17, 18] for DRM), motivated our theoretical study in [14]. The circumcenter schemes we came up with are already in the attention of specialists of the field (see Aragón Artacho *et al* [2], Bauschke *et al* [10–12], Lindstrom and Sims [26] and Ouyang [27]) and questions on the possibility of successful behavior in more general and more important settings are arising. It is worth emphasizing that DRM handles satisfactorily some highly relevant kinds of problems related to nonconvex and inconsistent feasibility problems involving (affine) subspaces (see, for instance, [8, 9, 19, 24, 25]). This suggests a promising behavior of circumcenter-type methods for these kinds of problems since CRM may be seen as a geometrical improvement of DRM.

The linear convergence of the circumcentered–reflection method (CRM) was established in [14] for solving problem (1) with $m \geq 2$ affine subspaces. Since the computation of a circumcenter requires the resolution of a suitable $m \times m$ linear system, this might not be of negligible computational cost for large m. To avoid this drawback for problems where the computation of z_{CRM} is simply too demanding, we propose in the present work the Block-wise Circumcentered–Reflection Method (Bw-CRM) by using an arbitrary ordered partition of the indices $\{1, 2, \ldots, m\}$, which contains the original CRM described above as a particular realization. Moreover, two elegant connections of this scheme with MAP follow. These nice interpretations further indicate a possible potential of the proposed method for solving problems more general than (1) (even nonconvex), where some affine structure remains, though.

The presentation of this paper is as follows. Definitions, basic facts and important auxiliary results are presented in Section 2. Still in Section 2, we introduce the notion of *best approximation mapping* along with properties of these mappings, which are key to our work. In Section 3, we formally introduce Bw-CRM. The global Q-Linear convergence of Bw-CRM for problem (1) is proven in Section 3.1. Connections between Bw-CRM and MAP are briefly discussed in Section 3.2. In Section 4 we prove the curious CRM feature of solving problem (1) in only one step when the correspondent affine subspaces are hyperplanes. Numerical illustrations are presented in Section 5. In Section 6 we provide a summary of our work and new ideas for future investigation.

2 Preliminary and auxiliary results

Let us review the definition of Friedrichs angle and provide key results needed in sequel.

Definition 1 (Friedrichs angle) The *cosine of the Friedrichs angle* between affine subspaces V and W with nonempty intersection is given by

$$c_F := \sup \left\{ \langle v, w \rangle \mid v \in \hat{V} \cap (\hat{V} \cap \hat{W})^{\perp},\ w \in \hat{W} \cap (\hat{V} \cap \hat{W})^{\perp},\ \|v\| \leq 1,\ \|w\| \leq 1 \right\}.$$

Here, $\hat{V}$ and $\hat{W}$ are subspaces given by $V - \hat{z}$ and $W - \hat{z}$, respectively, where $\hat{z} \in V \cap W$ is arbitrary but fixed, and the $\perp$ operation provides the correspondent orthogonal subspace.

In the above definition, it is easy to check that c_F does not depend on the choice of $\hat{z}$. Moreover, it is well known that $0 \leq c_F < 1$, for any two intersecting affine subspaces. See fundamental properties of the Friedrichs angle in [20, Theorem 13] and [21, Lemma 9.5], for instance.

For clearer presentation of our results, we introduce the concept of *best approximation mapping* (BAM).

Definition 2 (best approximation mapping) Let $V \neq \varnothing$ be a given affine subspace in $\mathbb{R}^n$. We say that $G_V : \mathbb{R}^n \to \mathbb{R}^n$ is a *best approximation mapping* with respect to V (for short V-BAM) if

(i) $P_V(G_V(z)) = P_V(z)$, for all $z \in \mathbb{R}^n$; and

(ii) there exists a constant $r_V \in [0, 1)$ such that $\|G_V(z) - P_V(z)\| \leq r_V \|z - P_V(z)\|$, for all $z \in \mathbb{R}^n$.

Note that the projection operator P_V is a V-BAM. Indeed, if $G_V = P_V$, for any $z \in \mathbb{R}^n$ and all $r_V \geq 0$, we have $P_V(G_V(z)) = P_V(P_V(z)) = P_V(z)$ and $\|G_V(z) - P_V(z)\| = \|P_V(z) - P_V(z)\| = 0 \leq r_V \|z - P_V(z)\|$. In general, it is easy to see that $G_V = (1 - \alpha) \operatorname{Id} + \alpha P_V$ with $0 < \alpha < 2$ is a V-BAM with $r_V = |1 - \alpha| \in [0, 1)$. Nonetheless, Definition 2 allows for non-affine mappings. Later we will see and use the fact that the circumcenter operator defined in [14] is a *best approximation mapping*, even though it is usually non-affine.

Simple manipulations provide an immediate consequence of Definition 2.

Proposition 1 *Let G_V be a V-BAM with constant $r_V \in [0, 1)$. For any $z \in \mathbb{R}^n$ and $\ell \in \mathbb{N}$, $P_V(G_V^\ell(z)) = P_V(z)$ and $(G_V^k(z))_{k \in \mathbb{N}}$ converges to $P_V(z)$ with linear rate r_V.*

Proof Let $z \in \mathbb{R}^n$ and $\ell \in \mathbb{N}$. Then, Definition 2(i) implies

$$P_V(G_V^\ell(z)) = P_V(G_V(G_V^{\ell-1}(z))) = P_V(G_V^{\ell-1}(z)) = \cdots = P_V(G_V(z)) = P_V(z).$$

Moreover,

$$
\begin{aligned}
\|G_V^k(z) - P_V(z)\| &= \|G_V(G_V^{k-1}(z)) - P_V(G_V^{k-1}(z))\| \\
&\leq r_V \|G_V^{k-1}(z) - P_V(G_V^{k-1}(z))\| \leq \cdots \leq r_V^{k-1} \|G_V(z) - P_V(G_V(z))\| \\
&= r_V^{k-1} \|G_V(z) - P_V(z)\| \\
&\leq r_V^k \|z - P_V(z)\|,
\end{aligned}
$$

proving the proposition. $\qquad\square$

The main purpose of this section is to study the composition of best approximation mappings. In order to do this, we state and prove an auxiliary result on adjacent angles.

Proposition 2 *Let $u, v \in \mathbb{R}^n$ be nonzero vectors forming an angle $\gamma \in [0, \frac{\pi}{2}]$. If a nonzero vector $w \in \mathbb{R}^n$ forms an angle $\beta \in [0, \pi]$ with v and the angle ϕ between w and u is such that $\phi \in [0, \frac{\beta}{2}]$, then $\gamma \geq \frac{\beta}{2}$.*

Proof Assume, without loss of generality, that $\|u\| = \|v\| = \|w\| = 1$. Then,

$$\cos \gamma = \langle u, v \rangle, \ \cos \beta = \langle v, w \rangle \ \text{and} \ \cos \phi = \langle u, w \rangle.$$

Also, we have $\cos \gamma \geq 0$ and $\cos \phi \geq 0$.

If $\beta = 0$, $\gamma \geq \frac{\beta}{2}$, trivially. Moreover, $\beta = \pi$ if, and only if, $w = -v$. In this case,

$$0 \leq \cos \phi = \langle u, w \rangle = \langle u, -v \rangle = -\langle u, v \rangle = -\cos \gamma \leq 0.$$

Thus, $\cos \gamma = 0$ and $\gamma = \frac{\pi}{2} = \frac{\beta}{2}$.

For the rest of the proof, let $\beta \in (0, \pi)$ and consider the following convex optimization problem

$$
\begin{aligned}
\min_x \quad & -\langle v, x \rangle \\
\text{s.t.} \quad & \frac{1}{2}\|x\|^2 - \frac{1}{2} \leq 0 \\
& \cos \frac{\beta}{2} - \langle w, x \rangle \leq 0.
\end{aligned}
\tag{2}
$$

By Weierstrass, this problem has a solution as the objective function is continuous and the feasible set is compact. Note that $\frac{1}{2}\left(1 + \cos\frac{\beta}{2}\right) w$ is a Slater point, that is, it fulfills both constraints strictly because $\cos\frac{\beta}{2} < 1$ as $\beta \in (0, \pi)$. Therefore, x is a solution of (2) if, and only if, it satisfies the KKT conditions

$$-v + \mu_1 x - \mu_2 w = 0 \tag{3a}$$

$$\|x\| - 1 \le 0 \tag{3b}$$

$$\cos\frac{\beta}{2} - \langle w, x \rangle \le 0 \tag{3c}$$

$$\mu_1 \ge 0, \ \mu_2 \ge 0 \tag{3d}$$

$$\mu_1 \left(\|x\| - 1\right) = 0 \tag{3e}$$

$$\mu_2 \left(\cos\frac{\beta}{2} - \langle w, x \rangle\right) = 0, \tag{3f}$$

where μ_1 and μ_2 are Lagrange multipliers.

We claim that $x^* := \frac{v+w}{\|v+w\|}$, which is well defined since $\beta \ne \pi$, is a KKT point for (2). Condition (3a) is satisfied with $\mu_1^* = \|v + w\|$ and $\mu_2^* = 1$. These multipliers yield (3d) strictly. Trivially, since $\|x^*\| = 1$, condition (3b) holds sharply and (3e) follows as well. Obviously, from $\langle v, v \rangle = \langle w, w \rangle = 1$, we have

$$\langle w, x^* \rangle = \left\langle w, \frac{v+w}{\|v+w\|} \right\rangle = \frac{\langle w, v \rangle + \langle w, w \rangle}{\|v+w\|} = \frac{\langle v, v \rangle + \langle v, w \rangle}{\|v+w\|} = \left\langle v, \frac{v+w}{\|v+w\|} \right\rangle = \langle v, x^* \rangle. \tag{4}$$

Then,

$$2 \langle w, x^* \rangle = \langle v, x^* \rangle + \langle w, x^* \rangle = \left\langle v, \frac{v+w}{\|v+w\|} \right\rangle + \left\langle w, \frac{v+w}{\|v+w\|} \right\rangle$$

$$= \frac{\langle v+w, v+w \rangle}{\|v+w\|} = \frac{\|v+w\|^2}{\|v+w\|} = \|v+w\|$$

$$= \sqrt{\|v\|^2 + \|w\|^2 + 2\langle v, w \rangle}$$

$$= \sqrt{2 + 2\cos\beta} = 2\sqrt{\frac{1 + \cos\beta}{2}}$$

$$= 2\cos\frac{\beta}{2}, \tag{5}$$

that is, condition (3c) holds sharply and yields (3f).

Note that u is a feasible point for (2). In fact, by assumption $\|u\| = 1$ and $0 \le \phi \le \frac{\beta}{2}$, which means that $\langle w, u \rangle = \cos\phi \ge \cos\frac{\beta}{2}$.

Finally, using the definition of γ, the optimality of x^*, (4) and (5) we derive

$$\cos\gamma = \langle v, u \rangle \le \langle v, x^* \rangle = \langle w, x^* \rangle = \cos\frac{\beta}{2}.$$

Hence, $\gamma \ge \frac{\beta}{2}$, proving the lemma. $\qquad\square$

We now start to address the composition of best approximation mappings. The next result is the keystone of our analysis.

Lemma 1 (composition of two best approximation mappings) *Let us consider two affine subspaces V and W of $\mathbb{R}^n$ with nonempty intersection $V \cap W$. Then, the composition of a V-BAM and a W-BAM is a $(V \cap W)$-BAM.*

Proof Let $G_V : \mathbb{R}^n \to \mathbb{R}^n$ and $G_W : \mathbb{R}^n \to \mathbb{R}^n$ be two *best approximation mappings* with respect to V and W, respectively, and corresponding constants $0 \le r_V < 1$ and $0 \le r_W < 1$.

In order to prove item (i) of Definition 2 for the composition $G := G_W \circ G_V$ w.r.t. $V \cap W$ we are going to combine Pythagoras equations with properties of projections. Note that we have to prove that $P_{V \cap W}(G(z)) = P_{V \cap W}(z)$ for all $z \in \mathbb{R}^n$.

Let us take an arbitrary, but fixed, $z \in \mathbb{R}^n$ and set $\hat{z} := P_{V \cap W}(z)$. The definition of $\hat{z}$ implies that $\hat{z} \in V \cap W$. In particular, $\hat{z} \in V$ and we have

$$\|z - \hat{z}\|^2 = \|z - P_V(z)\|^2 + \|P_V(z) - \hat{z}\|^2. \tag{6}$$

Since $P_{V \cap W}(P_V(z)) \in V$, we can write

$$\|z - P_{V \cap W}(P_V(z))\|^2 = \|z - P_V(z)\|^2 + \|P_V(z) - P_{V \cap W}(P_V(z))\|^2. \tag{7}$$

Of course, $\|z - \hat{z}\| \leq \|z - P_{V \cap W}(P_V(z))\|$. Using this fact and subtracting (7) from (6) yields $\|P_V(z) - \hat{z}\| \leq \|P_V(z) - P_{V \cap W}(P_V(z))\|$. By uniqueness of projections onto closed convex sets, we conclude that

$$\hat{z} = P_{V \cap W}(P_V(z)). \tag{8}$$

The fact that both $\hat{z}$ and $P_{V \cap W}(G_V(z))$ lie in V allows us to derive further Pythagoras relations

$$\|G_V(z) - \hat{z}\|^2 = \|G_V(z) - P_V(G_V(z))\|^2 + \|P_V(G_V(z)) - \hat{z}\|^2$$

and

$$\|G_V(z) - P_{V \cap W}(G_V(z))\|^2 = \|G_V(z) - P_V(G_V(z))\|^2 + \|P_V(G_V(z)) - P_{V \cap W}(G_V(z))\|^2.$$

Since G_V is a V-BAM, it holds that $P_V(G_V(z)) = P_V(z)$ and the previous equations reduce to

$$\|G_V(z) - \hat{z}\|^2 = \|G_V(z) - P_V(z)\|^2 + \|P_V(z) - \hat{z}\|^2 \tag{9}$$

and

$$\|G_V(z) - P_{V \cap W}(G_V(z))\|^2 = \|G_V(z) - P_V(z)\|^2 + \|P_V(z) - P_{V \cap W}(G_V(z))\|^2. \tag{10}$$

As proved above in (8), $\hat{z} = P_{V \cap W}(P_V(z))$, which implies that $\|P_V(z) - \hat{z}\| \leq \|P_V(z) - P_{V \cap W}(G_V(z))\|$. This inequality, together with (9) and (10), gives us $\|G_V(z) - \hat{z}\| \leq \|G_V(z) - P_{V \cap W}(G_V(z))\|$. Then, as $P_{V \cap W}(G_V(z))$ is uniquely defined and $\hat{z} \in V \cap W$, we must have

$$P_{V \cap W}(G_V(z)) = \hat{z}. \tag{11}$$

Our proof towards item (i) of Definition 2 continues with similar arguments, now regarding W. By Pythagoras we get

$$\|G_V(z) - P_{V \cap W}(G_V(z))\|^2 = \|G_V(z) - P_W(G_V(z))\|^2 + \|P_W(G_V(z)) - P_{V \cap W}(G_V(z))\|^2$$

and

$$\|G_V(z) - P_{V \cap W}(G_W(G_V(z)))\|^2 = \|G_V(z) - P_W(G_V(z))\|^2 + \|P_W(G_V(z)) - P_{V \cap W}(G_W(G_V(z)))\|^2.$$

Since we proved that $P_{V \cap W}(G_V(z)) = \hat{z}$, taking into account that G_W is a W-BAM, which provides $P_W(G_W(G_V(z))) = P_W(G_V(z))$, and bearing in mind that $G(z) = G_W(G_V(z))$, we can rewrite the equations above as

$$\|G_V(z) - \hat{z}\|^2 = \|G_V(z) - P_W(G(z))\|^2 + \|P_W(G(z)) - \hat{z}\|^2 \tag{12}$$

and

$$\|G_V(z) - P_{V \cap W}(G(z))\|^2 = \|G_V(z) - P_W(G(z))\|^2 + \|P_W(G(z)) - P_{V \cap W}(G(z))\|^2. \tag{13}$$

From the definition of Euclidean projection it follows that $\|G_V(z) - \hat{z}\| \leq \|G_V(z) - P_{V \cap W}(G_V(z))\|$, because $\hat{z}$ realizes the distance of $G_V(z)$ to $V \cap W$. This, combined with (12) and (13), leads to $\|P_W(G(z)) - \hat{z}\| \leq \|P_W(G(z)) - P_{V \cap W}(G(z))\|$.

We can derive two additional Pythagoras relations

$$\|G(z) - \hat{z}\|^2 = \|G(z) - P_W(G(z))\|^2 + \|P_W(G(z)) - \hat{z}\|^2 \tag{14}$$

and

$$\|G(z) - P_{V \cap W}(G(z))\|^2 = \|G(z) - P_W(G(z))\|^2 + \|P_W(G(z)) - P_{V \cap W}(G(z))\|^2. \tag{15}$$

We have just seen that $\|P_W(G(z)) - \hat{z}\| \leq \|P_W(G(z)) - P_{V \cap W}(G(z))\|$, which together with (14) and (15) yields $\|G(z) - \hat{z}\| \leq \|G(z) - P_{V \cap W}(G(z))\|$. Hence, $P_{V \cap W}(G(z)) = \hat{z} = P_{V \cap W}(z)$, which fulfills condition (i) of Definition 2 for G w.r.t. $V \cap W$.

Let us now address item (ii) of Definition 2 for G w.r.t. $V \cap W$. We have to prove that there exists a nonnegative constant $0 \leq r_{V \cap W} < 1$ so that, for all $z \in \mathbb{R}^n$,

$$\|G(z) - P_{V \cap W}(z)\| \leq r_{V \cap W} \|z - P_{V \cap W}(z)\|. \tag{16}$$

Again, let $z \in \mathbb{R}^n$ be arbitrary, fixed and $\hat{z} = P_{V \cap W}(z)$. If $z = \hat{z}$, (16) is fulfilled for any nonnegative constant as $G(z)$ will be equal to $P_{V \cap W}(z)$. In fact, G_V being a V-BAM, together with $z = \hat{z}$, gives us

$$\|G_V(z) - \hat{z}\| = \|G_V(z) - P_V(z)\| \le r_V \|z - P_V(z)\| = r_V \|z - \hat{z}\| = 0.$$

Thus, $G_V(z) = \hat{z}$. On the other hand, G_W is a W-BAM, so

$$\|G(z) - z\| = \|G(z) - \hat{z}\| = \|G_W(G_V(z)) - \hat{z}\| = \|G_W(\hat{z}) - \hat{z}\| = \|G_W(\hat{z}) - P_W(\hat{z})\| \le r_W \|\hat{z} - P_W(\hat{z})\| = 0.$$

This means that, if $z = \hat{z}$, $G(z) = z$ and the left-hand side of (16) is equal to zero and this inequality holds for any nonnegative constant $r_{V \cap W}$.

Therefore, from now on, assume $z \ne \hat{z}$. We will construct $r_{V \cap W}$ upon the constants $r_V \in [0, 1)$, $r_W \in [0, 1)$, and $c_F \in [0, 1)$, the latter the cosine of the Friedrichs angle $\theta_F \in (0, \frac{\pi}{2}]$ between V and W.

It will be key to look at the angle α between vectors $z - \hat{z}$ and $P_V(z) - \hat{z}$. Note first that $\alpha \in [0, \frac{\pi}{2}]$, since from (6) the triangle with vertices z, $\hat{z}$ and $P_V(z)$ has a right angle at $P_V(z)$. Also, of course,

$$\cos \alpha = \frac{\|P_V(z) - \hat{z}\|}{\|z - \hat{z}\|}. \tag{17}$$

Moreover, by using equation (9), the V-BAM hypothesis $\|G_V(z) - P_V(z)\| \le r_V \|z - P_V(z)\|$, equation (6) and that $\cos \alpha \le 1$, we conclude that

$$\begin{aligned}
\|G_V(z) - \hat{z}\|^2 &= \|G_V(z) - P_V(z)\|^2 + \|P_V(z) - \hat{z}\|^2 \\
&\le r_V^2 \|z - P_V(z)\|^2 + \|P_V(z) - \hat{z}\|^2 \\
&= r_V^2 \left(\|z - \hat{z}\|^2 - \|P_V(z) - \hat{z}\|^2 \right) + \|P_V(z) - \hat{z}\|^2 \\
&= r_V^2 \|z - \hat{z}\|^2 + (1 - r_V^2) \|P_V(z) - \hat{z}\|^2 \\
&= r_V^2 \|z - \hat{z}\|^2 + (1 - r_V^2) \cos^2 \alpha \|z - \hat{z}\|^2 \tag{18} \\
&\le \|z - \hat{z}\|^2. \tag{19}
\end{aligned}$$

Now, we split our analysis in two cases: $\alpha \in [\frac{\theta_F}{2}, \frac{\pi}{2}]$; $\alpha \in [0, \frac{\theta_F}{2})$.

Case 1: $\alpha \in [\frac{\theta_F}{2}, \frac{\pi}{2}]$.

In this case, $\cos \alpha \le \cos \frac{\theta_F}{2}$. This, combined with (18) and the fact that $\cos \frac{\theta_F}{2} = \sqrt{\frac{1 + c_F}{2}}$ provides

$$\begin{aligned}
\|G_V(z) - \hat{z}\|^2 &\le r_V^2 \|z - \hat{z}\|^2 + (1 - r_V^2) \cos^2 \alpha \|z - \hat{z}\|^2 \\
&\le r_V^2 \|z - \hat{z}\|^2 + (1 - r_V^2) \cos^2 \frac{\theta_F}{2} \|z - \hat{z}\|^2 \\
&= \left(r_V^2 + (1 - r_V^2) \frac{1 + c_F}{2} \right) \|z - \hat{z}\|^2. \tag{20}
\end{aligned}$$

Since $r_V^2 < 1$ and $\frac{1 + c_F}{2} < 1$, we have $(1 - r_V^2) \frac{1 + c_F}{2} < (1 - r_V^2)$. Then,

$$r_V^2 + (1 - r_V^2) \frac{1 + c_F}{2} < r_V^2 + (1 - r_V^2) = 1.$$

Since G_W is a W-BAM, we have $P_W(G_W(G_V(z))) = P_W(G_V(z))$ and $\|G_W(G_V(z)) - P_W(G_V(z))\| \le r_W \|G_V(z) - P_W(G_V(z))\|$, with $r_W \in [0, 1)$. So, we can write $\|G(z) - P_W(G(z))\| \le \|G_V(z) - P_W(G(z))\|$, which combined with (12) and (14), gives us

$$\|G(z) - \hat{z}\| \le \|G_V(z) - \hat{z}\|.$$

Hence, this inequality and (20) imply that

$$\|G(z) - \hat{z}\| \le r_1 \|z - \hat{z}\|,$$

with $r_1 \in [0, 1)$, given by

$$r_1 := \sqrt{r_V^2 + (1 - r_V^2) \frac{1 + c_F}{2}}. \tag{21}$$

Case 2: $\alpha \in [0, \frac{\theta_F}{2})$.

In this case we initially consider the triangle of vertexes $G_V(z)$, $\hat{z}$ and $P_W(G_V(z))$. Since G_V is a V-BAM, $P_V(G_V(z)) = P_V(z)$. We will be particularly interested in the angle ϕ between $G_V(z) - \hat{z}$ and $P_V(z) - \hat{z}$, when these vectors are nonzero. The vector $P_V(z) - \hat{z}$ is automatically nonzero, because of $\alpha < \frac{\pi}{2}$ and (17). If the vector $G_V(z) - \hat{z}$ is zero, we get the desired result as shown below.

Suppose $G_V(z) = \hat{z}$, then $G(z) = G_W(G_V(z)) = G_W(\hat{z})$ and it is easy to verify that $G_W(\hat{z}) = \hat{z}$. Indeed,

$$\|G_W(\hat{z}) - \hat{z}\| = \|G_W(\hat{z}) - P_W(\hat{z})\| \le r_W\|\hat{z} - P_W(\hat{z})\| = \|\hat{z} - \hat{z}\| = 0.$$

So, $G(z) = \hat{z}$ and the left-hand side of (16) is equal to zero and this inequality is fulfilled for any nonnegative constant $r_{V \cap W}$.

Assume for the rest of the proof that $G_V(z) \ne \hat{z}$. Thus,

$$\cos\phi = \frac{\|P_V(z) - \hat{z}\|}{\|G_V(z) - \hat{z}\|} \ge \frac{\|P_V(z) - \hat{z}\|}{\|z - \hat{z}\|} = \cos\alpha,$$

where the inequality is due to (19). Therefore, $0 \le \phi \le \alpha$ and, consequently, $\phi \in [0, \frac{\theta_F}{2})$.

We consider now another triangle, the one of vertexes $P_V(z)$, $\hat{z}$ and $P_W(G_V(z))$. If the vertexes $\hat{z}$ and $P_W(G_V(z))$ coincide, we get the following bound:

$$\begin{aligned}
\|G(z) - \hat{z}\| &= \|G_W(G_V(z)) - \hat{z}\| \\
&= \|G_W(G_V(z)) - P_W(G_V(z))\| \\
&\le r_W\|G_V(z) - P_W(G_V(z))\| \\
&\le r_W\|G_V(z) - \hat{z}\| \\
&\le r_W\|z - \hat{z}\|,
\end{aligned} \tag{22}$$

where we used, respectively, the definition of G, the current assumption $P_W(G_V(z)) = \hat{z}$, the hypothesis that G_W is a W-BAM, the fact that $\hat{z}$ lies in W and (19).

For the rest of the proof, assume also that $P_W(G_V(z)) \ne \hat{z}$ and define β, the angle between the nonzero vectors $P_V(z) - \hat{z}$ and $P_W(G_V(z)) - \hat{z}$. It is easy to see that the former belongs to $\hat{V}$ and the latter belongs to $\hat{W}$, where $\hat{V}$ and $\hat{W}$ are the subspaces given by $V - \hat{z}$ and $W - \hat{z}$, respectively. Also, recall from (8) that $P_{V \cap W}(P_V(z)) = \hat{z}$ and therefore $P_V(z) - \hat{z} \in (\hat{V} \cap \hat{W})^\perp$.

We rewrite (12) using the W-BAM property $P_W(G(z)) = P_W(G_W(G_V(z))) = P_W(G_V(z))$ as

$$\|G_V(z) - \hat{z}\|^2 = \|G_V(z) - P_W(G_V(z))\|^2 + \|P_W(G_V(z)) - \hat{z}\|^2$$

and Pythagoras can be employed as

$$\|G_V(z) - P_{V \cap W}(P_W(G_V(z)))\|^2 = \|G_V(z) - P_W(G_V(z))\|^2 + \|P_W(G_V(z)) - P_{V \cap W}(P_W(G_V(z)))\|^2.$$

On the one hand, $\|P_W(G_V(z)) - P_{V \cap W}(P_W(G_V(z)))\| \le \|P_W(G_V(z)) - \hat{z}\|$. On the other hand, $\|G_V(z) - \hat{z}\| \le \|G_V(z) - P_{V \cap W}(P_W(G_V(z)))\|$ because we have already seen in (11) that $P_{V \cap W}(G_V(z)) = \hat{z}$. Hence, $P_{V \cap W}(P_W(G_V(z))) = \hat{z}$ and $P_W(G_V(z)) - \hat{z} \in (\hat{V} \cap \hat{W})^\perp$.

We can then use the definition of the cosine of the Friedrichs angle θ_F between V and W and get

$$\cos\beta = \left\langle \frac{P_W(G_V(z)) - \hat{z}}{\|P_W(G_V(z)) - \hat{z}\|}, \frac{P_V(z) - \hat{z}}{\|P_V(z) - \hat{z}\|} \right\rangle \le c_F = \cos\theta_F,$$

which provides $\beta \in [\theta_F, \pi]$.

By now we have the nonzero vectors $G_V(z) - \hat{z}$, $P_V(z) - \hat{z}$ and $P_W(G_V(z)) - \hat{z}$. The vectors $G_V(z) - \hat{z}$ and $P_V(z) - \hat{z}$ form angle $\phi \in [0, \frac{\theta_F}{2})$, vectors $P_V(z) - \hat{z}$ and $P_W(G_V(z)) - \hat{z}$ form angle $\beta \in [\theta_F, \pi]$. Let γ be the angle between vectors $G_V(z) - \hat{z}$ and $P_W(G_V(z)) - \hat{z}$. Obviously, by Pythagoras, $\gamma \in [0, \frac{\pi}{2}]$ and

$$\cos\gamma = \frac{\|P_W(G_V(z)) - \hat{z}\|}{\|G_V(z) - \hat{z}\|}. \tag{23}$$

More than that, from Proposition 2, we conclude that $\gamma \in [\frac{\beta}{2}, \frac{\pi}{2}]$. In particular, we get $\gamma \ge \frac{\beta}{2} \ge \frac{\theta_F}{2}$ and hence

$$\cos\gamma \le \cos\frac{\theta_F}{2} = \sqrt{\frac{1 + c_F}{2}}. \tag{24}$$

Then, enforcing similar arguments as in Case 1, we obtain

$$
\begin{aligned}
\|G(z) - \hat{z}\|^2 &= \|G_W(G_V(z)) - \hat{z}\|^2 \\
&= \|G_W(G_V(z)) - P_W(G_V(z))\|^2 + \|P_W(G_V(z)) - \hat{z}\|^2 \\
&\leq r_W^2 \|G_V(z) - P_W(G_V(z))\|^2 + \|P_W(G_V(z)) - \hat{z}\|^2 \\
&= r_W^2 (\|G_V(z) - P_W(G_V(z))\|^2 + \|P_W(G_V(z)) - \hat{z}\|^2) + (1 - r_W^2)\|P_W(G_V(z)) - \hat{z}\|^2 \\
&= r_W^2 \|G_V(z) - \hat{z}\|^2 + (1 - r_W^2)\|P_W(G_V(z)) - \hat{z}\|^2 \\
&= r_W^2 \|G_V(z) - \hat{z}\|^2 + (1 - r_W^2)\cos^2 \gamma \|G_V(z) - \hat{z}\|^2 \\
&\leq \left(r_W^2 + (1 - r_W^2)\frac{1 + c_F}{2} \right) \|G_V(z) - \hat{z}\|^2 \\
&\leq \left(r_W^2 + (1 - r_W^2)\frac{1 + c_F}{2} \right) \|z - \hat{z}\|^2.
\end{aligned}
\tag{25}
$$

The first line corresponds to the definition of G, the second is by Pythagoras and the third holds because G_W is a W-BAM. The fourth line is a rearrangement of terms, followed by Pythagoras in the fifth. Then, (23) and (24) are employed respectively. At last, we used (19).

Analogously to the proof that $r_1 = \sqrt{r_V^2 + (1 - r_V^2)\frac{1 + c_F}{2}}$ is strictly smaller than 1, we can see that

$$
r_2 := \sqrt{r_W^2 + (1 - r_W^2)\frac{1 + c_F}{2}} < 1.
$$

Finally, we can gather Cases 1 and 2. From (21), (22) and (25), we have $\|G(z) - \hat{z}\| \leq r_{V \cap W}\|z - \hat{z}\|$ for all $z \in \mathbb{R}^n$, with $r_{V \cap W} \in [0, 1)$ given by $r_{V \cap W} := \max\{r_1, r_W, r_2\} = \max\{r_1, r_2\}$.   □

We are going to see next that Lemma 1 can be extended to the case of ℓ affine subspaces, with ℓ being any positive integer.

Theorem 1 (finite composition of best approximation mappings) *Let us consider an indexed family of ℓ affine subspaces* $\mathbf{W} = \{W_1, W_2, \ldots, W_\ell\}$ *of* $\mathbb{R}^n$ *with nonempty intersection* S_ℓ. *Assume that each* $G_{W_j} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ *($j = 1, \ldots, \ell$) is W_j-BAM. Then, $G := G_{W_\ell} \circ \cdots \circ G_{W_2} \circ G_{W_1}$ is a S_ℓ-BAM.*

Proof The proof follows by an induction argument on ℓ, the number of affine subspaces.

If $\ell = 1$, we have $G = G_{W_1}$ and then G is a S_ℓ-BAM.

Assume the result for a fixed ℓ. Let $\widehat{\mathbf{W}} := \mathbf{W} \cup \{W_{\ell+1}\}$, where $W_{\ell+1}$ is an affine subspace such that it has nonempty intersection $S_{\ell+1}$ with S_ℓ, and let $G_{W_{\ell+1}}$ be a $W_{\ell+1}$-BAM. Employing Lemma 1 with S_ℓ and $W_{\ell+1}$ playing the role of V and W, respectively, and $\widehat{G}$ and $G_{W_{\ell+1}}$ playing the role of G_V and G_W, respectively, we get that $\widehat{G} := G \circ G_{W_{\ell+1}}$ is a $S_{\ell+1}$-BAM.   □

In the next section, we define the block-wise circumcenter operator and will prove that it is a *best approximation mapping*.

3 The block-wise circumcentered–reflection method

The main purpose of this paper is applying the recently developed circumcentered–reflection method (CRM) [14] to solve problem (1) by taking advantage of a block-wise structure. This idea may be beneficial in certain problems coming from the discretization of partial differential equations as we describe and illustrate in our numerical section. We remind that CRM iterates by taking an ordered round of successive reflections onto affine subspaces and then it chooses the new iterate by means of equidistance to the reflected points, which explains the usage of the geometric term *circumcenter*.

Let us give the definition of the circumcenter of a block of finitely many affine subspaces.

Definition 3 (circumcentered-reflection for a block) Let $\mathcal{B} := (U_1, U_2, \ldots, U_q)$ be a block of ordered affine subspaces, where $q \geq 1$ is a fixed integer. Suppose also that the intersection $S_{\mathcal{B}} := \bigcap_{i=1}^q U_i$ is nonempty. The circumcenter of the block $\mathcal{B}$ at the point $z \in \mathbb{R}^n$ is denoted by $C_{\mathcal{B}}(z)$ and defined by the following properties:

(i) $C_{\mathcal{B}}(z) \in W_z := \text{aff}\{z, R_{U_1}(z), R_{U_2} R_{U_1}(z), \ldots, R_{U_q} \cdots R_{U_2} R_{U_1}(z)\}$;

(ii) $\|z - C_{\mathcal{B}}(z)\| = \|R_{U_1}(z) - C_{\mathcal{B}}(z)\| = \cdots = \|R_{U_q} \cdots R_{U_2} R_{U_1}(z) - C_{\mathcal{B}}(z)\|$.

It is worth noting that the order in which reflections are composed affects the outcome circumcenter. If not said otherwise, we use increasing order of indices for the computation of a circumcenter.

Before presenting the definition of the block-wise circumcentered–reflection method (Bw-CRM), we list two consequences of results from Lemma 3.1 of [14] that will be at the core of our convergence analysis for Bw-CRM.

Lemma 2 (good definition of CRM) *Consider a block of affine subspaces $\mathcal{B} = (U_1, U_2, \ldots, U_q)$ with $S_{\mathcal{B}} = \cap_{i=1}^{q} U_i$ nonempty. For any $z \in \mathbb{R}^n$, $C_{\mathcal{B}}(z)$ is well and uniquely defined.*

Proof See [14, Lemma 3.1]. □

The circumcenter, as above, is the intersection of suitable bisectors. Its computation requires the resolution of a $q \times q$ linear system of equations. Details can be found in [14, p. 161] and [10, Theorem 4.1].

The previous lemma established that the circumcenter is well defined. We now recall that the circumcenter operator is a BAM.

Theorem 2 (circumcenter operator is a BAM) *Consider a block of affine subspaces $\mathcal{B} = (U_1, U_2, \ldots, U_q)$ with $S_{\mathcal{B}} = \cap_{i=1}^{q} U_i$ nonempty. Then, there exists a constant $r_{\mathcal{B}} \in [0, 1)$ so that*

$$\|C_{\mathcal{B}}(z) - P_{S_{\mathcal{B}}}(z)\| \le r_{\mathcal{B}} \|z - P_{S_{\mathcal{B}}}(z)\|,$$

for all $z \in \mathbb{R}^n$. Moreover, $P_{S_{\mathcal{B}}}(C_{\mathcal{B}}(z)) = P_{S_{\mathcal{B}}}(z)$.

Proof See [14, Lemma 3.2]. □

The previous theorem says that $C_{\mathcal{B}}$ is a $S_{\mathcal{B}}$-BAM. In order to define our new circumcenter scheme, consider the following terminology.

Definition 4 (block partition) We say that $\mathbf{B} = (\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_p)$ is an ordered collection of blocks (with cardinality p) for the ordered affine subspaces $U_1, U_2, \ldots, U_m$ if we can write $\mathcal{B}_1 = (U_{q_0+1}, U_{q_0+2}, \ldots, U_{q_1})$, $\mathcal{B}_2 = (U_{q_1+1}, U_{q_1+2}, \ldots, U_{q_2})$, $\ldots$, $\mathcal{B}_p = (U_{q_{p-1}+1}, U_{q_{p-1}+2}, \ldots, U_{q_p})$, with $q_0 = 0$ and $q_p = m$. We assume that every block $\mathcal{B}_i$ has size $q_i - q_{i-1} \ge 1$, $i = 1, \ldots, p$.

Note that in the previous definition we are simply selecting subsets of subspaces based on a partition of the set of indices, illustrated below

$$\mathbf{I} = \{1, 2, \ldots, m\} = \{\underbrace{q_0 + 1, \ldots, q_1}_{\text{1st block indices}}, \underbrace{q_1 + 1, \ldots, q_2}_{\text{2nd block indices}}, \underbrace{q_2 + 1, \ldots, q_3}_{\text{3rd block indices}}, \ldots, \underbrace{q_{p-1} + 1, \ldots, q_p}_{p\text{-th block indices}}\}.$$

We now define the block-wise circumcentered-reflection operator.

Definition 5 (block-wise circumcentered-reflection) Let $\mathbf{B} = (\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_p)$ be an ordered collection of blocks for the affine subspaces $U_1, U_2, \ldots, U_m$ and assume that increasing index order is taken for both blocks and subspaces. Then, for a point $z \in \mathbb{R}^n$ we define the *block-wise circumcentered-reflection* step $C_{\mathbf{B}}(z)$ by

$$C_{\mathbf{B}}(z) := C_{\mathcal{B}_p} \circ C_{\mathcal{B}_{p-1}} \circ \cdots \circ C_{\mathcal{B}_2} \circ C_{\mathcal{B}_1}(z).$$

A key result is presented in Section 3.1. It establishes linear convergence of the sequence $\left(C_{\mathbf{B}}^k(z)\right)_{k \in \mathbb{N}}$ to $P_S(z)$. Our proof that Bw-CRM provides a sequence converging linearly to the solution of the best approximation problem (1) depends on some further auxiliary results, derived in the next section.

In the following section, the circumcenter operators for each block $C_{\mathcal{B}_j}$ will play the role of the *best approximation mappings* G_{W_j}'s. Furthermore, $C_{\mathbf{B}}$ will play the role of G in Theorem 1.

3.1 Linear convergence of the block-wise circumcentered–reflection method

Now, we summarize our result on Bw-CRM. Remind that $\mathbf{B} = (\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_p)$ is a fixed ordered collection of ordered p blocks for the affine subspaces $U_1, U_2, \ldots, U_m$. Recall the notation $S := \cap_{i=1}^m U_i$ and $C_{\mathbf{B}}$ for the block-wise circumcentered-reflection operator regarding $\mathbf{B}$. Due to the last auxiliary result, we easily derive linear convergence of Bw-CRM for solving problem (1). Next, we formally state that $C_{\mathbf{B}}$ is a *best approximation mapping* with respect to S.

Theorem 3 (block-wise operator is a BAM) *Let $C_{\mathbf{B}}$ be the block-wise circumcentered-reflection operator regarding $\mathbf{B}$. Then, there exists a constant $r_{\mathbf{B}} \in [0, 1)$ so that*

$$\|C_{\mathbf{B}}(z) - P_S(z)\| \le r_{\mathbf{B}}\|z - P_S(z)\|,$$

for all $z \in \mathbb{R}^n$. Moreover, $P_S(C_{\mathbf{B}}(z)) = P_S(z)$ and the convergence of $(C_{\mathbf{B}}^k(z))_{k \in \mathbb{N}}$ is linear to the unique solution $P_S(z)$, i.e.,

$$\lim_{k \to \infty} C_{\mathbf{B}}^k(z) = P_S(z).$$

Furthermore, the global Q-linear rate is $r_{\mathbf{B}} \in [0, 1)$, i.e., for all $k \in \mathbb{N}$,

$$\|C_{\mathbf{B}}^k(z) - P_S(z)\| \le r_{\mathbf{B}}^k\|z - P_S(z)\|.$$

Proof Due to Definition 5, $C_{\mathbf{B}}$ is a composition of circumcenter operators, which of each is a BAM (Theorem 1) and thus, by Theorem 2, it is itself a BAM. The claims on the sequence $(C_{\mathbf{B}}^k(z))_{k \in \mathbb{N}}$ follow then from Proposition 1. □

3.2 Connections between Bw-CRM and MAP

Based on our papers [13, 14] and on the previous results, we briefly discuss now some curious connections between Bw-CRM and the method of alternating projections (MAP).

Our concept of *best approximation mapping* is, by definition, a relaxation of a projection operator. With that said, the first relation between Bw-CRM and MAP we want to point out is that Bw-CRM happens to be a *best approximation mapping*, as proven in the last section. Furthermore, the well known linear convergence of MAP for a finite number of intersecting affine subspaces [21, Theorems 9.31 and 9.33] follows as an immediate consequence of the result on best approximation mappings stated in Theorem 1.

Another connection between Bw-CRM and MAP follows from the fact that the projection of a point onto a closed convex set can be seen as the circumcenter regarding the given point and its reflection onto the corresponding set. In other words, if you have a point $z \in \mathbb{R}^n$ and a closed convex set U, then $P_U(z) = \text{circumcenter}\{z, R_U(z)\}$ because $P_U(z) \in \text{aff}\{z, R_U(z)\}$ and $\|z - P_U(z)\| = \|R_U(z) - P_U(z)\|$. Therefore, considering the notation from the previous section, we can observe that when all blocks $\mathcal{B}_i$'s have cardinality 1, i.e., $p = m$ and $\mathcal{B}_i = (U_i)$ for all $i = 1, \ldots, m$, we have that $C_{\mathcal{B}_i}$ is precisely the orthogonal projector onto U_i. Hence, the block-wise circumcentered-reflection operator $C_{\mathbf{B}} := C_{\mathcal{B}_p} \circ \cdots \circ C_{\mathcal{B}_2} \circ C_{\mathcal{B}_1}$ coincides with the MAP operator $P_{U_m} \circ \cdots \circ P_{U_2} \circ P_{U_1}$.

In addition to having the aforementioned connections to MAP, we will see next that the full-block Bw-CRM, i.e., CRM itself, serves as a projector when the multi-set intersection regards only hyperplanes. CRM indeed finds the projection of any given point onto the intersection of hyperplanes in one single step. Perhaps, such a feature might be useful in the implementation of projection methods.

4 One step convergence of CRM for hyperplane intersection

The initial motivation in the development of our first circumcenter scheme in [13] was defining a method that could handle the trivial problem of finding the intersection of two crossing lines in $\mathbb{R}^2$ in one step. In the present section, this is done in dimension n for hyperplanes.

The key ingredient that enables the full block Bw-CRM (original CRM) to converge in only one step for hyperplane intersection is that the orthogonal subspace to a given nonempty hyperplane always has dimension one. Interestingly, the first clues on this one-step convergence were indicated by our numerical experiments. Thanks to them we came up with the following results.

Lemma 3 (one step convergence for full block Bw-CRM) *Consider $\mathcal{H} = (H_1, H_2, \ldots, H_p)$ where H_i's are hyperplanes with nonempty intersection $S_{\mathcal{H}}$ and let $C_{\mathcal{H}}$ be the CRM operator regarding $\mathcal{H}$. If $z \in \mathbb{R}^n$ is so that for all $i = 1, 2, \ldots, p$ we have $R_{H_i} \cdots R_{H_1}(z) \notin H_i$, then the circumcenter $C_{\mathcal{H}}(z)$ is already the projection of z onto $S_{\mathcal{H}}$.*

Proof Without loss of generality, we assume that H_i's are subspaces, as their intersection $S_{\mathcal{H}}$ is nonempty.

It was proven in [14, Lemma 3.1] that $C_{\mathcal{H}}(z)$ is precisely the projection of $P_{S_{\mathcal{H}}}(z)$ onto

$$W_z = \mathrm{aff}\{z, R_{H_1}(z), R_{H_2}R_{H_1}(z), \ldots, R_{H_p} \cdots R_{H_2}R_{H_1}(z)\}.$$

Therefore, by considering the subspace $\hat{W}_z := W_z - C_{\mathcal{H}}(z)$, we have

$$P_{S_{\mathcal{H}}}(z) - C_{\mathcal{H}}(z) \perp \hat{W}_z. \tag{26}$$

Let $v_1 := R_{H_1}(z) - z$, $v_2 := R_{H_2}R_{H_1}(z) - R_{H_1}(z), \ldots, v_p := R_{H_p} \cdots R_{H_2}R_{H_1}(z) - R_{H_{p-1}} \cdots R_{H_2}R_{H_1}(z)$. Clearly, $v_i \in \hat{W}_z$, for $i = 1, \ldots, p$ and $\hat{W}_z = \mathrm{span}\{v_1, v_2, \ldots, v_p\}$. Also, from the definition of reflection, we have $v_i \perp H_i$, for all $i = 1, \ldots, p$.

By taking into account the hypothesis $R_{H_i} \cdots R_{H_1}(z) \notin H_i$, it is straightforward to conclude that all v_i's are not zero. Then, since each H_i is a hyperplane, we have

$$\mathrm{span}\{v_i\} = H_i^{\perp}, \ i = 1, \ldots, p.$$

Now, linear algebra gives us

$$\begin{aligned}
\hat{W}_z &= \mathrm{span}\{v_1, v_2, \ldots, v_p\} \\
&= \mathrm{span}\{v_1\} + \mathrm{span}\{v_2\} + \cdots + \mathrm{span}\{v_p\} \\
&= H_1^{\perp} + H_2^{\perp} + \cdots + H_p^{\perp} \\
&= \mathrm{span}\{H_1^{\perp} \cup H_2^{\perp} \cup \cdots \cup H_p^{\perp}\} \\
&= (H_1 \cap H_2 \cap \cdots \cap H_p)^{\perp} \\
&= S_{\mathcal{H}}^{\perp}
\end{aligned}$$

and from (26) we have

$$P_{S_{\mathcal{H}}}(z) - C_{\mathcal{H}}(z) \perp S_{\mathcal{H}}^{\perp}. \tag{27}$$

We have shown in [13, 14] that $P_{S_{\mathcal{H}}}(C_{\mathcal{H}}(z)) = P_{S_{\mathcal{H}}}(z)$ and because $P_{S_{\mathcal{H}}}(C_{\mathcal{H}}(z)) - C_{\mathcal{H}}(z)$ is orthogonal to $S_{\mathcal{H}}$, it follows that

$$P_{S_{\mathcal{H}}}(z) - C_{\mathcal{H}}(z) \perp S_{\mathcal{H}}. \tag{28}$$

The combination of (27) and (28) implies that $P_{S_{\mathcal{H}}}(z) - C_{\mathcal{H}}(z) = 0$, that is,

$$C_{\mathcal{H}}(z) = P_{S_{\mathcal{H}}}(z),$$

which completes the proof. □

We observe that one can easily construct an example with two lines playing the role of hyperplanes in $\mathbb{R}^2$ violating the hypothesis in Lemma 3 for certain initial points, where indeed the one step convergence of CRM is lost. We might then ask if at least finite convergence of CRM can always be expected in the case of hyperplane intersection. Although we lean towards a positive answer to this interesting theoretical question, we note that it is essentially irrelevant. There are at least two reasons for that. The first is that violating $R_{H_i} \cdots R_{H_1}(z) \notin H_i$ is completely "*bad luck*". More formally, one can actually show that the set $\{z \in \mathbb{R}^n \mid R_{H_i} \cdots R_{H_1}(z) \notin H_i, \forall i = 1, 2, \ldots, p\}$ is dense in $\mathbb{R}^n$ (see further comments at the end of the section). The second reason why having z in the complement of the previous set, namely *bad luck*, is not really an issue, is that we can derive a simple and cheap procedure to rewrite our best approximation problem in an equivalent way such that CRM solves the reformulation in one single step. Next we describe this procedure upon a lemma.

Lemma 4 (procedure for dealing with *bad luck*) *Consider $\mathcal{H} = (H_1, H_2, \ldots, H_p)$ where the H_i's are hyperplanes with nonempty intersection $S_{\mathcal{H}}$. Let $z \in \mathbb{R}^n$ and assume the existence of a smallest index $\check{\imath}$ in $\{1, 2, \ldots, p\}$ for which $\check{z} := R_{H_{\check{\imath}}} \cdots R_{H_2}R_{H_1}(z) \in H_{\check{\imath}}$. Denote by $\check{a}$ any given non-null orthogonal vector to the hyperplane $H_{\check{\imath}}$ and let us write $z_{rep} := z + tR_{H_1}R_{H_2} \cdots R_{H_{\check{\imath}-2}}R_{H_{\check{\imath}-1}}R_{H_{\check{\imath}}}(\check{a})$, where "rep" stands for the idea of replacement of z. Then, for all real number t we have $P_{S_{\mathcal{H}}}(z_{rep}) = P_{S_{\mathcal{H}}}(z)$ and for all non-null t sufficiently close to zero it holds that $R_{H_i} \cdots R_{H_2}R_{H_1}(z_{rep}) \notin H_i$ for $i = 1, \ldots, \check{\imath}$.*

Proof Without loss of generality, assume that the hyperplanes $H_1, H_2, \ldots, H_p$ are subspaces, as their intersection $S_{\mathcal{H}}$ is nonempty. The fact that reflections onto subspaces preserve the correspondent best approximation solution is a trivial consequence of Pythagoras and the definition and affinity of the reflections. So, the projections onto $S_{\mathcal{H}}$ of all the points $R_{H_i} \cdots R_{H_2} R_{H_1}(z)$ with $i = 1, 2, \ldots, p$ is given by $P_{S_{\mathcal{H}}}(z)$. This holds in particular for $\check{z}$. By construction, $t\check{a}$ is orthogonal to H_i, hence we conclude using Pythagoras again that for all real number t the projection of $\check{z} + t\check{a}$ onto $S_{\mathcal{H}}$ is also given by $P_{S_{\mathcal{H}}}(z)$. Now, it is easy to see that z_{rep} is defined by reflections of $\check{z} + t\check{a}$ onto H_i's starting backwards from the index $\check{i}$ until 1. Indeed, remind that $z_{rep} := z + t R_{H_1} R_{H_2} \cdots R_{H_{\check{i}-2}} R_{H_{\check{i}-1}} R_{H_{\check{i}}}(\check{a})$, thus

$$R_{H_1}(z_{rep}) = R_{H_1}(z + t R_{H_1} R_{H_2} \cdots R_{H_{\check{i}-2}} R_{H_{\check{i}-1}} R_{H_{\check{i}}}(\check{a}))$$

Using the linearity of the reflection R_{H_1} and the fact that $R_{H_1} R_{H_1} = \mathrm{Id}$, we get

$$R_{H_1}(z_{rep}) = R_{H_1}(z) + t R_{H_2} \cdots R_{H_{\check{i}-2}} R_{H_{\check{i}-1}} R_{H_{\check{i}}}(\check{a}).$$

Employing this argument successively for R_{H_2} until $R_{H_{\check{i}}}$ implies that

$$R_{H_{\check{i}}} \cdots R_{H_2} R_{H_1}(z_{rep}) = R_{H_{\check{i}}} \cdots R_{H_2} R_{H_1}(z) + t\check{a},$$

that is,

$$R_{H_{\check{i}}} \cdots R_{H_2} R_{H_1}(z_{rep}) = \check{z} + t\check{a}.$$

It follows that the projections of z_{rep} and $\check{z} + t\check{a}$ onto $S_{\mathcal{H}}$ must coincide. Hence, $P_{S_{\mathcal{H}}}(z_{rep}) = P_{S_{\mathcal{H}}}(z)$.

For all non-null t we have $R_{H_{\check{i}}} \cdots R_{H_2} R_{H_1}(z_{rep}) = \check{z} + t\check{a} \notin H_{\check{i}}$ as $\check{a}$ is non-null and orthogonal to $H_{\check{i}}$. This gives the lemma if $\check{i} = 1$. So, assume from now on that $\check{i} > 1$. It remains to show that $R_{H_i} \cdots R_{H_2} R_{H_1}(z_{rep}) \notin H_i$ for $i = 1, \ldots, \check{i} - 1$ if we take a non-null t with sufficiently small modulus. That follows easily by hypothesis together with continuity of reflections and Euclidean distance to hyperplanes. By the definition of $\check{i}$ we have that $\mathrm{dist}(R_{H_1}(z), H_1) \neq 0$. Therefore, by continuity in t of the function

$$f_1(t) := \mathrm{dist}(R_{H_1}(z_{rep}), H_1) = \mathrm{dist}(R_{H_1}(z + t R_{H_1} R_{H_2} \cdots R_{H_{\check{i}-2}} R_{H_{\check{i}-1}} R_{H_{\check{i}}}(\check{a})), H_1)$$

we must have a whole interval $[-t_1, t_1]$, with $t_1 > 0$ for which $f_1(t) \neq 0$. Intervals $[-t_i, t_i]$ with $t_i > 0$ like the previous one can be derived in the same way for the remaining indices $i = 2, \ldots, \check{i} - 1$ by considering the functions

$$f_i(t) := \mathrm{dist}(R_{H_i} \cdots R_{H_2} R_{H_1}(z_{rep}), H_i) = \mathrm{dist}(R_{H_i} \cdots R_{H_2} R_{H_1}(z + t R_{H_1} R_{H_2} \cdots R_{H_{\check{i}-2}} R_{H_{\check{i}-1}} R_{H_{\check{i}}}(\check{a})), H_i).$$

Let $[-\check{t}, \check{t}]$ represent the smallest of these intervals. We then have that $R_{H_i} \cdots R_{H_2} R_{H_1}(z_{rep}) \notin H_i$ for $i = 1, \ldots, \check{i}$ if z_{rep} is defined by means of a parameter t belonging to $[-\check{t}, \check{t}]$. □

Note that the previous lemma does not necessarily lead us to a point z_{rep} under the conditions of Lemma 3, we only have an improvement with respect to the index $\check{i}$. Nevertheless, if the *rep* operation defined in Lemma 4 is applied successively at most $p - \check{i}$ times, we get a new initial point say z_{REP} so that $P_{S_{\mathcal{H}}}(z) = P_{S_{\mathcal{H}}}(z_{REP})$ and we have $R_{H_i} \cdots R_{H_2} R_{H_1}(z_{REP}) \notin H_i$ for all $i = 1, 2, \ldots, p$. That is, z_{REP} satisfies the conditions of Lemma 3 while keeping $P_{S_{\mathcal{H}}}(z)$ as the best approximation solution. This means that the full block Bw-CRM, which is the original CRM, is categorically always able to find the solution of the best approximation problem (1) in one single step for hyperplane intersection. Let us state this as a theorem.

Theorem 4 (one step convergence of CRM) *Let $\mathcal{H} = (H_1, H_2, \ldots, H_p)$ where H_i's are hyperplanes with nonempty intersection $S_{\mathcal{H}}$, $C_{\mathcal{H}}$ be the CRM operator regarding $\mathcal{H}$ and $z \in \mathbb{R}^n$ be given. Then, CRM finds the projection of z onto $S_{\mathcal{H}}$ in one single step (with eventual use of z_{REP} as described above).*

We remind that the probability of having to employ the *rep* procedure is zero. This is due to the fact that the set of points $z \in \mathbb{R}^n$ so that $R_{H_i} \cdots R_{H_1}(z) \notin H_i$ for all $i = 1, 2, \ldots, p$ is dense in $\mathbb{R}^n$. The density holds because any $z \in \mathbb{R}^n$ violating the aforementioned conditions can be approximated by a sequence of correspondent z_{rep}'s coming from sufficiently shrinking the size of $t \neq 0$ from Lemma 4. In any case, note that the *rep* procedure is implementable. One only needs to consider a backtracking search on the parameter t, reflect onto hyperplanes (which can be done by closed formula) and check pertinence to these hyperplanes.

To finalize the discussion in this section, we would like to present some further remarks.

We want to note that one can consider trivial examples showing that the conditions for one-step convergence in Lemma 3, although sufficient, are not necessary. CRM will converge in one single step whenever the successive reflections generate an affine space of dimension $n - r$, where r is the dimension of the intersection of the given

subspaces. One could have the dimension $n - r$ even if the given subspaces are not hyperplanes and also under the *bad luck* of getting reflected points precisely on them.

Our last remark is on possible finite convergence of CRM for hyperplane intersection without employing the *rep* procedure at all. Although omitting the proof, we notice that CRM converges in at most 3 steps with no *rep* procedure for the intersection of 2 hyperplanes in $\mathbb{R}^n$. The challenging question for more than 2 hyperplanes is left open. Also, we intend to investigate under which conditions one has finite convergence for CRM, when the subspaces are not all hyperplanes.

5 Numerical illustrations

The geometric nature of Bw-CRM can be used as a tool for solving some classical problems, *e.g.*, the least squares problem, the minimum-norm least-squares (rank deficient) problems, the least-norm solutions of undetermined system and under-determined large-scale linear systems, which are particular instances of problem (1). In this section, we illustrate the performance of Bw-CRM to solve two related problems: an application in computed tomography and the minimum-norm least square problem. We run all the numerical experiments in `Julia` language [15].

5.1 Application in Computed Tomography

Reconstruction of images in Computed Tomography (CT) can be addressed by approximately solving linear systems of equations coming from the discretization of suitable inverse problems. Algebraic reconstruction techniques (ART), which are basically MAP type methods, are usually employed to solve those linear systems as not much accuracy is needed for a solution representing a reasonable image for medical purposes [23, Chapter 11].

In this subsection, we solve a problem $As = b$, whose solution provides the well known Shepp-Logan phantom head [28]. This is a standard synthetic image that serves as the model of a human head and is used for testing image reconstruction algorithms. The data for the matrix A and the vector b were generated using `AIR Tools II`, a package by Hansen and Jørgensen [22], and imported to be used in the `Julia` implementation. In this case, A has 5732 rows and 2500 columns. The package also provides the exact 50×50 pixel Shepp-Logan image, which is represented as $\hat{z} \in \mathbb{R}^{2500}$.

In our experiments, we use Bw-CRM and look at the quality of image reconstructions after a fixed budget of 10 iterations. The affine subspaces under consideration are the hyperplanes given by each row of $As = b$. These affine subspaces are distributed in blocks, where each block contains q hyperplanes, except maybe for the last one which contains $(5732 \bmod q)$ hyperplanes. We exhibit in Table 1 the residue and distance to the actual solution of each version of Bw-CRM, where Bw-CRM-q indicates that the block size used is q — or $(5732 \bmod q)$ and the time in seconds of which method. Remind that Bw-CRM-1 is MAP. It is worth noting that Bw-CRM-16, Bw-CRM-64 and Bw-CRM-256 all beat Bw-CRM-1 (MAP) both in iterations to achieve the same residue.

Table 1: Bw-CRM applied to CT – Matrix size: 5732×2500 – Budget of 10 iterations.

Method-Block size	$\|Az_{10} - b\|$	$\|z_{10} - \hat{z}\|$	CPU (s)
Bw-CRM-1 (MAP)	3.0321×10^1	1.3816	5.3876
Bw-CRM-16	2.8590×10^1	1.3382	8.5242
Bw-CRM-64	4.2602	1.0332	6.1665
Bw-CRM-256	7.1039×10^{-1}	2.7423×10^{-1}	8.7073

In Figure 1 we display the original solution and each reconstruction by Bw-CRM for $q = 1, 16, 64, 256$. The best solution is achieved by Bw-CRM-256 at the price of solving 22 symmetric positive definite linear systems of size 256 and 1 of size 100, as $5732 = 22 \cdot 256 + 100$.
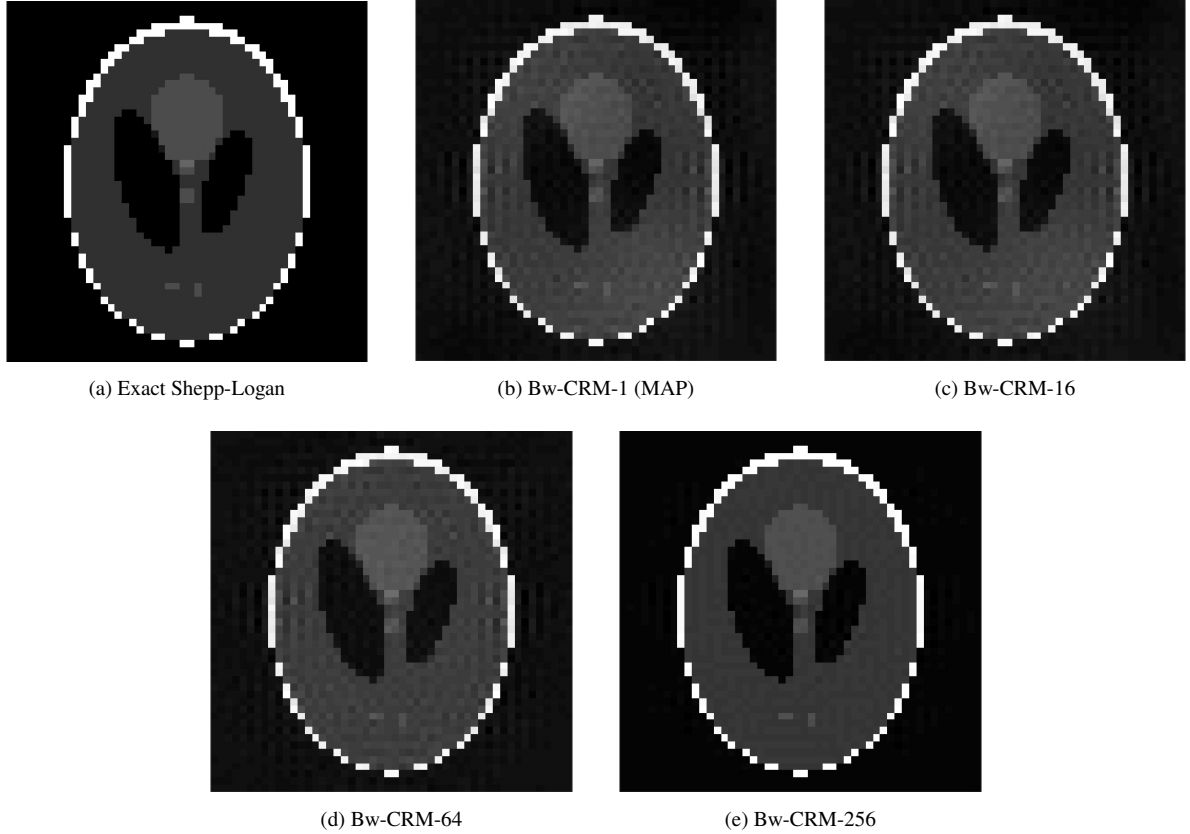
<div align="center">(a) Exact Shepp-Logan   (b) Bw-CRM-1 (MAP)   (c) Bw-CRM-16</div>

<div align="center">(d) Bw-CRM-64   (e) Bw-CRM-256</div>

Fig. 1: CT image reconstructions of Shepp-Logan phantom of size 50×50.

5.2 Solving a least norm problem

A direct application of Bw-CRM is to solve the following optimization problem: Find $\hat{z} \in \mathbb{R}^n$, the solution of

$$\min \|z - s\|, \text{ subject to } As = b, \tag{29}$$

where $A \in \mathbb{R}^{p \times n}(p \leq n)$, $b \in \mathbb{R}^p$ and z a given vector. The solution $\hat{z}$ is the closest point to z that lies in the intersection $S_{\mathcal{H}}$ of the hyperplanes in $\mathcal{H} := (H_1, H_2, \ldots, H_p)$, where H_i is given by the solutions of the i-th equation of $As = b$, that is, $\hat{z}$ is the projection of z onto $S_{\mathcal{H}}$.

As shown in Section 4, Bw-CRM, when applied to solving this problem by taking the p individual hyperplanes forming the equations (as the main block $\mathcal{H}$), finds the solution $\hat{z}$ in just one iteration — hatring some *bad luck*, as already discussed. If we set $z = 0$, thus $C_{\mathcal{H}}(0) = \hat{z}$ and problem (29) becomes the minimum norm of under-determined system problem (MNP). It is well-known that if A has full rank we can solve (29) by using the Moore-Penrose pseudo-inverse of A, as $\hat{z} = z + A^T (AA^T)^{-1}(b - Az)$.

In order to illustrate various possible choices of blocks for Bw-CRM, we solve problem (29) using matrix coming from a finite element modeling, called `FIDAP005`, and available at Matrix Market [16]. The matrix A is given by selecting respectively the first 12, 24 and 27 rows of `FIDAP005`, b is the correspondent vector of ones and we take $z = 0$. The structure of the entire sparse matrix `FIDAP005` is shown in Figure 2. Next, we show the results for Bw-CRM in Tables 2 to 4, where each subspace under consideration is given by a row equation of $As = b$. The different size of block choices are displayed in the first column of the tables, followed by the number of blocks, the number of projections/reflections, the number of iterations, the norm of the residue and the CPU time, in seconds. The stopping criterion was having the norm of the residue smaller than the labeled tolerance `tol`. Note that Table 4 presents the results where the sparse block structure of matrix `FIDAP005` is explored.

As expected by the results of Section 4, the full block Bw-CRM converges in one iteration for the hyperplane intersection problems above. Note that we have to be careful when looking at the CPU time as it depends on the inner linear system solver for finding circumcenters. What we can say, though, is that the number of iterations tends
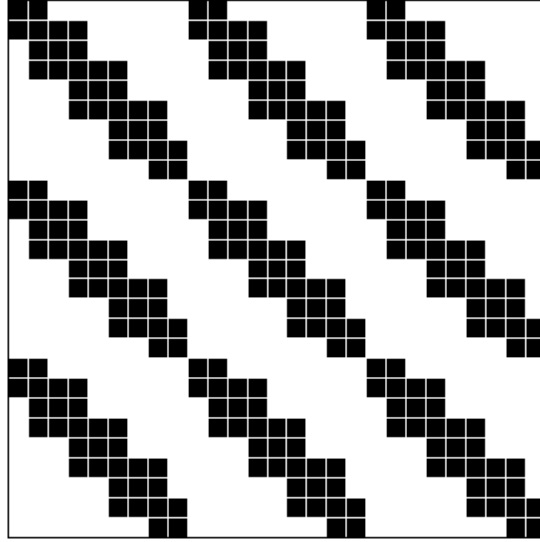
Fig. 2: Matrix `FIDAP005` sparsity structure.

Table 2: Results for Bw-CRM – Matrix size: 12×27 – `tol` $= 10^{-5}$.

Method-Block size	Blocks	Proj/Reflec	Iter	$\|Az_{\text{Iter}} - b\|$	CPU (s)
Bw-CRM-1 (MAP)	12	180	15	4.2323×10^{-6}	6.2563×10^{-4}
Bw-CRM-2	6	156	13	7.1863×10^{-6}	8.7344×10^{-4}
Bw-CRM-3	4	180	15	5.2967×10^{-6}	9.3172×10^{-4}
Bw-CRM-4	3	120	10	8.5466×10^{-6}	5.7481×10^{-4}
Bw-CRM-6	2	132	11	3.4566×10^{-6}	6.1370×10^{-4}
Bw-CRM-12 (CRM)	1	12	1	7.8280×10^{-14}	9.9924×10^{-5}

Table 3: Results for Bw-CRM – Matrix size: 24×27 – `tol` $= 10^{-5}$.

Method-Block size	Blocks	Proj/Reflec	Iter	$\|Az_{\text{Iter}} - b\|$	CPU (s)
Bw-CRM-1 (MAP)	24	12 048	502	9.8869×10^{-6}	6.9701×10^{-2}
Bw-CRM-2	12	11 904	496	9.8041×10^{-6}	1.2950×10^{-1}
Bw-CRM-3	8	11 160	465	9.9053×10^{-6}	9.9697×10^{-2}
Bw-CRM-4	6	9576	399	9.9067×10^{-6}	6.5870×10^{-2}
Bw-CRM-6	4	11 880	495	9.8665×10^{-6}	8.9289×10^{-2}
Bw-CRM-8	3	10 440	435	9.7581×10^{-6}	7.0528×10^{-2}
Bw-CRM-12	2	8328	347	9.7967×10^{-6}	8.1303×10^{-2}
Bw-CRM-24 (CRM)	1	24	1	1.4852×10^{-12}	1.4166×10^{-4}

Table 4: Results for Bw-CRM – Matrix size: 27×27 – `tol` $= 10^{-3}$.

Method-Block size	Blocks	Proj/Reflec	Iter	$\|Az_{\text{Iter}} - b\|$	CPU (s)
BW-CRM-1 (MAP)	27	3 992 166	147 858	9.9998×10^{-4}	3.0667×10^{1}
BW-CRM-3	9	3 448 602	127 726	9.9997×10^{-4}	2.6326×10^{1}
BW-CRM-9	3	3 209 355	118 865	9.9999×10^{-4}	2.9115×10^{1}
BW-CRM-27 (CRM)	1	27	1	6.9229×10^{-10}	1.5987×10^{-4}

to slightly increase as the number of blocks increase. It would be interesting to investigate whether there exists a sort of optimal block size, with respect to particular instances.

In contrast to the feasible set of the problems regarding Tables 2 and 3, the feasible set of the problem addressed in Table 4 reduces to a singleton. Even though it is known that MAP suffers from zig-zag behavior, we got surprised with the huge amount of iterations that it took to converge in the case of Table 4. We have established connections between Bw-CRM and MAP in Section 3.2 and unfortunately it seems that, in the case, when MAP performs poorly this is inherited by Bw-CRM, except for the full block Bw-CRM. This is a motivation for future investigation on randomized order of subspaces or blocks for Bw-CRM, as randomized versions of MAP performs a lot better [29].

6 Concluding remarks

We presented new notions and results regarding circumcenter schemes for projecting a given point onto the (nonempty) intersection of a finite number of affine subspaces. Circumcenter iterations were introduced in [13] and shown to provide a better bond between reflections than the one considered in the classical Douglas-Rachford approach. The results in [14] improved [13] by enabling the Circumcentered-Reflection Method (CRM) to deal with $m > 2$ affine subspaces. In the present article we also dealt with more than two sets. We defined the Block-wise Circumcentered-Reflection Method (Bw-CRM), which considers the m affine subspaces in blocks. More precisely, we composed circumcenter operators along a partition of the indices $1, 2, \ldots, m$. In this way, the original circumcenter method from [14] can be seen as Bw-CRM with one full block, where this block contains all m affine subspaces. It was interesting that by considering Bw-CRM with m blocks, *i.e.*, the case where each block contains exactly one affine subspace, we recovered the famous method of alternating projections (MAP). Linear convergence for any blocks choice of Bw-CRM was proven. Our proof was carried out in a unified fashion thanks to the introduction of a new concept, the one of *best approximation mapping*. In addition to deriving theoretical linear convergence of Bw-CRM, numerical experiments were run. For the numerical tests we considered blocks with homogeneous cardinality in order to investigate the relation between speed of convergence (time/complexity) and number of blocks in Bw-CRM. The experiments also indicated what became a curious result in this paper: it turns out that CRM (Bw-CRM with one full block) finds the projection of any given point onto the intersection of hyperplanes in one single step.

This work contributed not only with a deeper understanding of circumcenter type methods, we think that our results represent another step towards using circumcenters in other settings. Our future research will be focused on enforcing circumcenter iterations for solving the nonconvex problem: Find $z \in S$ with

$$S = U \cap V,$$

where $U = \bigcup_{i=1}^{m} U_i$, with U_i, for each i, being a subspaces and V being an affine subspace. This problem contains as a particular case the nonconvex sparse affine feasibility problem for which DRM and MAP fail to converge globally. We have strong convictions based on initial numerical tests and some preliminary proofs that a (block-wise) circumcenter method can perform very well (global convergence) for this kind of affine-structured problem.

References

1. Aragón Artacho, F.J., Borwein, J.M., Tam, M.K.: Recent Results on Douglas–Rachford Methods for Combinatorial Optimization Problems. J. Optim. Theory Appl. **163**(1), 1–30 (2013)
2. Aragón Artacho, F.J., Campoy, R., Tam, M.K.: The Douglas-Rachford Algorithm for Convex and Nonconvex Feasibility Problems. arXiv (1904.09148) (2019)
3. Bauschke, H.H., Bello-Cruz, J.Y., Nghia, T.T.A., Phan, H.M., Wang, X.: The rate of linear convergence of the Douglas–Rachford algorithm for subspaces is the cosine of the Friedrichs angle. J. Approx. Theory **185**, 63–79 (2014)
4. Bauschke, H.H., Bello-Cruz, J.Y., Nghia, T.T.A., Phan, H.M., Wang, X.: Optimal Rates of Linear Convergence of Relaxed Alternating Projections and Generalized Douglas-Rachford Methods for Two Subspaces. Numer. Algorithms **73**(1), 33–76 (2016)
5. Bauschke, H.H., Borwein, J.M.: On the convergence of von Neumann's alternating projection algorithm for two sets. Set-Valued Anal. **1**(2), 185–212 (1993)
6. Bauschke, H.H., Borwein, J.M.: On Projection Algorithms for Solving Convex Feasibility Problems. SIAM Rev. **38**(3), 367–426 (2006)
7. Bauschke, H.H., Deutsch, F.R., Hundal, H., Park, S.H.: Accelerating the Convergence of the Method of Alternating Projections. Trans. Amer. Math. Soc. **355**(9), 3433–3461 (2003)
8. Bauschke, H.H., Luke, D.R., Phan, H.M., Wang, X.: Restricted Normal Cones and the Method of Alternating Projections: Theory. Set-Valued Var. Anal. **21**(3), 431–473 (2013)
9. Bauschke, H.H., Moursi, W.M.: The Douglas–Rachford Algorithm for Two (Not Necessarily Intersecting) Affine Subspaces. SIAM J. Optim. **26**(2), 968–985 (2016)
10. Bauschke, H.H., Ouyang, H., Wang, X.: On circumcenters of finite sets in Hilbert spaces. Linear Nonlinear Anal. **4**(2), 271–295 (2018)
11. Bauschke, H.H., Ouyang, H., Wang, X.: Circumcentered methods induced by isometries. arXiv (1908.11576) (2019)
12. Bauschke, H.H., Ouyang, H., Wang, X.: On circumcenter mappings induced by nonexpansive operators. Pure and Applied Functional Analysis (in press)
13. Behling, R., Bello-Cruz, J.Y., Santos, L.R.: Circumcentering the Douglas–Rachford method. Numer. Algorithms **78**(3), 759–776 (2018)
14. Behling, R., Bello-Cruz, J.Y., Santos, L.R.: On the linear convergence of the circumcentered-reflection method. Oper. Res. Lett. **46**(2), 159–162 (2018)
15. Bezanson, J., Edelman, A., Karpinski, S., Shah, V.B.: Julia: A Fresh Approach to Numerical Computing. SIAM Rev. **59**(1), 65–98 (2017)

16. Boisvert, R.F., Pozo, R., Remington, K., Barrett, R.F., Dongarra, J.J.: Matrix Market: a web resource for test matrix collections. In: R.F. Boisvert (ed.) Quality of Numerical Software, pp. 125–137. Springer, Boston, MA, Boston, MA (1997)
17. Borwein, J.M., Tam, M.K.: A Cyclic Douglas–Rachford Iteration Scheme. J. Optim. Theory Appl. **160**(1), 1–29 (2014)
18. Borwein, J.M., Tam, M.K.: The cyclic Douglas-Rachford method for inconsistent feasibility problems. J. Nonlinear Convex Anal. **16**(4), 573–584 (2015)
19. Demanet, L., Zhang, X.: Eventual linear convergence of the Douglas-Rachford iteration for basis pursuit. Math. Comp. **85**(297), 209–238 (2016)
20. Deutsch, F.R.: The Angle Between Subspaces of a Hilbert Space. In: S.P. Singh (ed.) Approximation Theory, Wavelets and Applications, pp. 107–130. Springer, Dordrecht, Dordrecht (1995)
21. Deutsch, F.R.: Best Approximation in Inner Product Spaces. CMS Books in Mathematics. Springer, New York, NY (2001)
22. Hansen, P.C., Jørgensen, J.S.: AIR Tools II: algebraic iterative reconstruction methods, improved implementation. Numer. Algorithms **79**(1), 107–137 (2017)
23. Herman, G.T.: Fundamentals of computerized tomography, 2 edn. Advances in Pattern Recognition. Springer, Dordrecht (2009)
24. Hesse, R., Luke, D.R.: Nonconvex Notions of Regularity and Convergence of Fundamental Algorithms for Feasibility Problems. SIAM J. Optim. **23**(4), 2397–2419 (2013)
25. Hesse, R., Luke, D.R., Neumann, P.: Alternating Projections and Douglas-Rachford for Sparse Affine Feasibility. IEEE Trans. Signal Process. **62**(18), 4868–4881 (2014)
26. Lindstrom, S.B., Sims, B.: Survey: Sixty Years of Douglas–Rachford. arXiv (1809.07181) (2018)
27. Ouyang, H.: Circumcenter operators in Hilbert spaces. Master's thesis, University of British Columbia, Okanagan (2018)
28. Shepp, L.A., Logan, B.F.: The Fourier reconstruction of a head section. IEEE Trans. Nucl. Sci. **21**(3), 21–43 (1974)
29. Strohmer, T., Vershynin, R.: A Randomized Kaczmarz Algorithm with Exponential Convergence. J Fourier Anal Appl **15**(2), 262–278 (2008)