

# INFEASIBILITY AND ERROR BOUND IMPLY FINITE CONVERGENCE OF ALTERNATING PROJECTIONS\*

ROGER BEHLING<sup>†</sup>, YUNIER BELLO-CRUZ<sup>‡</sup>, AND LUIZ-RAFAEL SANTOS<sup>§</sup>

**Abstract.** This paper combines two ingredients in order to get a rather surprising result on one of the most studied, elegant and powerful tools for solving convex feasibility problems, the method of alternating projections (MAP). Going back to names such as Kaczmarz and von Neumann, MAP has the ability to track a pair of points realizing minimum distance between two given closed convex sets. Unfortunately, MAP may suffer from arbitrarily slow convergence, and sublinear rates are essentially only surpassed in the presence of some Lipschitzian error bound, which is our first ingredient. The second one is a seemingly unfavorable and unexpected condition, namely, infeasibility. For two non-intersecting closed convex sets satisfying an error bound, we establish finite convergence of MAP. In particular, MAP converges in finitely many steps when applied to a polyhedron and a hyperplane in the case in which they have empty intersection. Moreover, the farther the target sets lie from each other, the fewer are the iterations needed by MAP for finding a best approximation pair. Insightful examples and further theoretical and algorithmic discussions accompany our results, including the investigation of finite termination of other projection methods.

**Key words.** Convex feasibility problem, Infeasibility, Error bound, Finite convergence, Alternating projection.

**AMS subject classifications.** 47N10, 49M27, 65K05, 90C25

**1. Introduction.** The method of alternating projections (MAP) has a remarkable impact in so many areas of Mathematics and is one of the main classical tools for solving convex feasibility problems. A broad class of problems in Applied Mathematics is effectively solved by MAP [34]. Definitely one of a kind, MAP is not only capable of tracking a point in the intersection of given closed convex sets  $X, Y \subset \mathbb{R}^n$ , it delivers a replacement of an actual solution when  $X$  and  $Y$  do not intersect. Such a replacement comes in the form of a pair  $(\bar{x}, \bar{y}) \in X \times Y$ , often called *best approximation pair* to  $X$  and  $Y$ , as it minimizes the Euclidean distance between these two sets.

It is well-known that MAP converges globally whenever the distance between  $X$  and  $Y$  is attainable. It is also worth mentioning that, in view of Pierra's famous product space reformulation [48], feasibility problems involving a finite number of sets can be narrowed down to seeking a common point to two sets  $X$  and  $Y$ .

The present work focuses on the inconsistent case  $X \cap Y = \emptyset$  and reveals a surprising behavior of MAP in this setting. Roughly speaking, we come to the conclusion that infeasibility works in favor of MAP. Quite intuitive when looking at the scenarios displayed in Figure 1, the fact that infeasibility has a strong positive impact on MAP has apparently not been seen anywhere in its extensive literature. Actually, we prove that infeasibility added by an error bound condition provides finite convergence of the very pure MAP. More precisely, by assuming  $X \cap Y = \emptyset$  and a suitable error bound condition, we prove finite convergence of the MAP sequence defined by  $x^{k+1} := P_X P_Y(x^k)$  starting at any point  $x^0 \in \mathbb{R}^n$ . Finite convergence means that for some non-negative integer  $k$ , MAP reaches a best point  $\bar{x} \in X$ , that is,  $x^{\bar{k}} = \bar{x}$  and  $\bar{y} = P_Y(\bar{x})$  form a best approximation pair to  $X$  and  $Y$ . Here and throughout the text,  $P_X$  and  $P_Y$  stand for the orthogonal projections onto  $X$  and  $Y$ , respectively.

\*

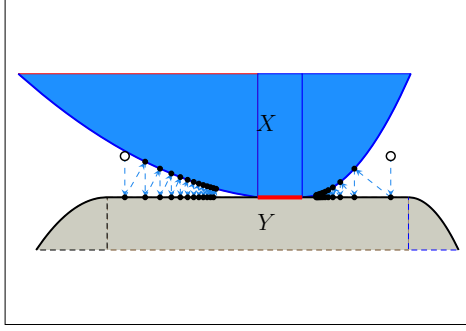
**Funding:** RB was partially supported by the *Brazilian Agency Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq), Grants 304392/2018-9 and 429915/2018-7; YBC was partially supported by the *National Science Foundation* (NSF), Grant DMS — 1816449.

<sup>†</sup>School of Applied Mathematics, Fundação Getúlio Vargas Rio de Janeiro, RJ — 22250-900, Brazil. (roger-behling@gmail.com) ✉

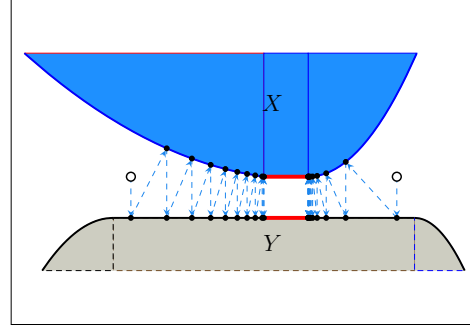
<sup>‡</sup>Department of Mathematical Sciences, Northern Illinois University. DeKalb, IL — 60115-2828, USA. (yunier-bello@niu.edu)

<sup>§</sup>Department of Mathematics, Federal University of Santa Catarina. Blumenau, SC — 88040-900, Brazil. (l.r.santos@ufsc.br)

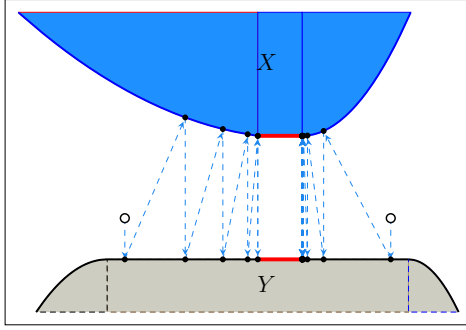
Let us now look at a collection of illustrations that serves as a summary of our results.



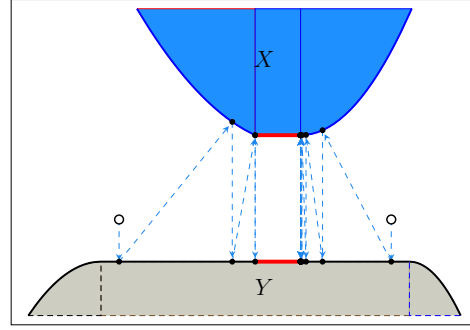
(a) Consistency; linear convergence on the left.



(b) Inconsistency;  $k = 10$  on the left.



(c) Inconsistency; increasing the distance between  $X$  and  $Y$ ;  $k = 5$  on the left.



(d) Inconsistency; improving the error bound on the left;  $k = 3$  on the left.

Fig. 1: Error bound on the left-hand sides.

Figure 1 displays four scenarios of MAP acting on sets  $X$  and  $Y$ . We start with a consistent problem in Figure 1a and analyze two MAP sequences. The one starting from the left converges linearly to a point in  $X \cap Y$ . This is due to the fact that in this region  $X$  and  $Y$  form a non-zero angle. In other words, a Lipschitzian error bound holds. This is not the case on the right-hand side of the picture and therefore, MAP only achieves sublinear convergence over there. By lifting  $X$ , we generate inconsistent intersection problems in 1b to 1d. The result is that MAP responds favorably speed-wise to this translation of  $X$ . More than that, MAP improves when increasing infeasibility and also when the error bound gets better, that is, the left border of  $X$  gets steeper. This can be noticed looking at Figure 1 as a film from Figures 1a to 1d. Lifting  $X$  from Figure 1a to Figure 1b makes MAP's convergence jump from linear to finite on the left. On the right-hand side, MAP leaps its convergence rate from sublinear to linear. After a further lift of  $X$  from Figure 1b to Figure 1c, MAP reaches a best point in 5 iterations instead of 10 on the left-hand side. MAP only needs 3 iterates to get a best approximation pair when improving the error bound in the left part of Figure 1d.

The message of Figure 1 is fairly clear. As for the aforementioned error bound condition, it will be formally introduced later. We anticipate, though, that it regards the target sets and their relation with what we are going to call optimal supporting hyperplanes. We point out that such an error bound is automatically globally fulfilled if  $X$  is a polyhedron and  $Y$  a hyperplane,

providing finite convergence of MAP if, in addition, these particular polyhedral sets have empty intersection. Note that finite convergence of MAP may happen even under a non-polyhedral structure, as depicted in Figure 1.

Before outlining the structure of our paper, we briefly go through some turning points in the history of MAP. This method has fascinated scientists for nearly a century now, and although many results on this simple tool have been derived, open questions remain. In the early 1930s, von Neumann firstly studied MAP for subspaces. Yet, his results were only published in 1950 [51], proving the convergence of MAP to a so-called best approximation solution for the intersection of two subspaces. In 1937, Kaczmarz [43] proposed a MAP related algorithm (also known as Cyclic projections) for finding best approximation solutions of linear systems. In 1959, MAP was studied for the convex case by Cheney and Goldstein [28], covering also inconsistent feasibility problems. Precisely the theme of our paper, projection methods for inconsistent inclusions have a history on their own; see the 2018 review by Censor and Zaknoon [27]. The ability of MAP to find best approximation pairs is a notable characteristic, making MAP (and its variants) one of the most used algorithms in Optimization [1, 8, 9, 17, 24, 30, 33, 35–37, 50]. Although MAP always converges under the existence of best approximation pairs, the convergence rate may be arbitrarily slow [6, 16, 41]. In 1950, Aronszajn [4] found the lower bound for the linear rate of MAP, given by the square of the cosine of the minimal angle (Friedrichs angle) between two subspaces, which turns out to be the sharpest one, as proved by Kayalar and Weinert [44] in 1988. For an in depth related literature on MAP; see, for instance, Bauschke and Borwein [10, 12], and Deutsch [34].

The paper is organized as follows. In Section 2 we collect known facts on MAP and some auxiliary material. Section 3 contains our mayor results concerning the geometry of two disjoint closed convex sets in the presence of an error bound condition. This analysis allows us to derive in Section 4 our main result, namely, finite convergence of MAP under inconsistency and what we call BAP error bound. The BAP error bound is automatically fulfilled for a polyhedron and a hyperplane, and we study its connection with linear regularity and intrinsic transversality. Also in Section 4, we will see that for a MAP sequence to converge in a finite number of steps, it is necessary and sufficient that it satisfies the BAP error bound at its iterates. Section 4 ends by investigating finite termination under inconsistency of Cyclic projections, Cimmino and Douglas-Rachford methods. In Section 5, we present insightful examples and applications. In particular, we connect our results to Linear Programming and convex min-max problems. We consider as well a simple problem giving rise to the question on whether a Hölder type error bound could make MAP's rate of convergence go from sublinear to linear when shifting the target sets apart. Some concluding remarks are presented in Section 6.

**2. Background material.** Let  $X, Y \subset \mathbb{R}^n$  be closed, convex and nonempty. We recall that the orthogonal projection of  $x \in \mathbb{R}^n$  onto  $X$  is given by  $P_X(x) \in X$  if, and only if,  $\langle y - P_X(x), x - P_X(x) \rangle \leq 0$ , for all  $y \in X$ . Throughout the text,  $\langle \cdot, \cdot \rangle$  stands for the Euclidean inner product inducing the norm  $\|\cdot\| := \sqrt{\langle \cdot, \cdot \rangle}$ . The non-negative integer numbers will be denoted by  $\mathbb{N}$ . The open ball centered in  $x$  with radius  $\delta > 0$  is the set  $\mathcal{B}_\delta(x) := \{z \in \mathbb{R}^n \mid \|x - z\| < \delta\}$ .

We define the *distance* between  $X$  and  $Y$  by  $\text{dist}(X, Y) := \inf\{\|x - y\| \mid x \in X, y \in Y\}$ . When one of the sets is a singleton, for instance  $X = \{x\}$ , we use the notation  $\text{dist}(x, Y)$ . A *best approximation pair* (BAP) relative to  $X$  and  $Y$  is a pair  $(\bar{x}, \bar{y}) \in X \times Y$  attaining the distance between  $X$  and  $Y$ , that is,  $\text{dist}(\bar{x}, \bar{y}) = \text{dist}(X, Y)$ . The set of all BAP relative to  $X$  and  $Y$  is denoted by  $\text{bap}(X, Y) \subset X \times Y$  and, accordingly, we define the sets

$$\text{bap}_Y(X) := \{x \in X \mid (x, y) \in \text{bap}(X, Y)\} = \{x \in X \mid \text{dist}(x, Y) = \text{dist}(X, Y)\}$$

and

$$\text{bap}_X(Y) := \{y \in Y \mid (x, y) \in \text{bap}(X, Y)\} = \{y \in Y \mid \text{dist}(y, X) = \text{dist}(X, Y)\}.$$

Note that  $\text{bap}_Y(X)$  (respectively  $\text{bap}_X(Y)$ ) is the subset of points in  $X$  (respectively  $Y$ ) *nearest* to  $Y$  (respectively  $X$ ). In the consistent case, that is, when  $X \cap Y$  is nonempty, we have  $\text{bap}_X(Y) = \text{bap}_Y(X) = X \cap Y$ . As we are interested in the inconsistent case, henceforth, we suppose that  $X \cap Y = \emptyset$ . In this context we define the *displacement vector* as  $d := P_{\text{cl}(X-Y)}(0)$ . So,  $\|d\| = \text{dist}(X, Y)$  and  $\text{dist}(X, Y)$  is attained if, and only, if  $d \in X - Y$ . In particular,  $\text{dist}(X, Y)$  is attained whenever  $X - Y$  is closed.

Given any point  $x \in \mathbb{R}^n$ , define the terms of the sequence  $(x^k)_{k \in \mathbb{N}}$  by

$$(2.1) \quad x^0 = x, \quad x^{k+1} = P_X P_Y(x^k),$$

for every  $k \in \mathbb{N}$ . The sequence  $(x^k)_{k \in \mathbb{N}}$  is the *alternating projection sequence* starting at  $x^0 = x$ . Cheney and Goldstein established in [28] that if one of the sets is compact or if one of the sets is finitely generated, the fixed point set of the operator  $P_X P_Y$  is nonempty and the sequence (2.1) converges to a fixed point of this operator. The general result was summarized and enlarged in [11] as follows.

FACT 2.1 (BAP sets [11, Lemma 2.2]). *Denote  $\text{Fix}(P_X P_Y) = \{x \in \mathbb{R}^n \mid P_X P_Y(x) = x\}$ . Then,*

- (i)  $\text{bap}_Y(X) = \text{Fix}(P_X P_Y)$ .
- (ii)  $\text{bap}_Y(X)$  and  $\text{bap}_X(Y)$  are closed convex sets.
- (iii) If  $\text{bap}_Y(X)$  or  $\text{bap}_X(Y)$  is nonempty then  $\text{dist}(X, Y)$  is attained. Moreover, let  $d$  be the displacement vector. Then

$$P_Y(\bar{x}) = \bar{x} - d, \quad \forall \bar{x} \in \text{bap}_Y(X),$$

$$\text{and } \text{bap}_Y(X) - d = \text{bap}_X(Y), \text{bap}_Y(X) = X \cap (Y + d), \text{bap}_X(Y) = (X - d) \cap Y.$$

In the next fact, we abuse notation and use  $\langle X, y \rangle \leq 0$  to denote that  $\langle x, y \rangle \leq 0, \forall x \in X$ .

FACT 2.2 (BAP pairs). *Let  $x \in \mathbb{R}^n$  be given. Then, if  $\text{dist}(X, Y)$  is attained, with  $d$  being the displacement vector, then*

$$P_{\text{bap}_X(Y)}(x) = P_{\text{bap}_Y(X)}(x) - d,$$

and  $\langle X - \text{bap}_Y(X), d \rangle \geq 0$  and  $\langle Y - \text{bap}_X(Y), d \rangle \leq 0$ .

FACT 2.3 (Convergence of MAP [11, Theorem 4.8]). *Let  $(x^k)_{k \in \mathbb{N}}$  be an alternating projection sequence as given in (2.1). Then,*

$$x^k - P_Y(x^k) \rightarrow d,$$

where  $d$  is the displacement vector. Moreover,

- (i) if  $\text{dist}(X, Y)$  is attained then  $x^k \rightarrow \bar{x} \in \text{bap}_Y(X)$  and  $P_Y(x^k) \rightarrow \bar{y} := \bar{x} - d \in \text{bap}_X(Y)$ ;
- (ii) if  $\text{dist}(X, Y)$  is not attained then  $\|x^k\| \rightarrow +\infty$ .

Next we present some definitions and well-known results concerning polyhedra, useful in Lemma 3.1 and Theorem 4.1.

DEFINITION 2.4 (Polyhedron). *A set  $\Omega \subset \mathbb{R}^n$  is said to be a polyhedron, if it can be expressed as the intersection of a finite family of closed half-spaces, that is,*

$$(2.2) \quad \Omega := \{x \in \mathbb{R}^n \mid \langle a_i, x \rangle \leq \alpha_i, \text{ for } i = 1, \dots, m\},$$

where  $a_i \in \mathbb{R}^n, \alpha_i \in \mathbb{R}$ .

Note that a polyhedron, also referred to as a *polyhedral set*, is always convex.

**DEFINITION 2.5** (Finitely generated cone and conic base). *A set  $K \subset \mathbb{R}^n$  is a cone if it is closed under positive scalar multiplication. A cone  $K$  is finitely generated if there exists a finite set  $S \subset \mathbb{R}^n$  such that  $\text{cone}(S) = K$ , where  $\text{cone}(S)$ , is the set of all conic combinations of elements of  $S$ . A conic base of a finitely generated cone  $K$  is a finite set  $B_K \subset \mathbb{R}^n$  with minimal cardinality such that  $\text{cone}(B_K) = K$ .*

**FACT 2.6** (Polyhedron is finitely generated [20, Proposition B.17]). *A set  $\Omega \subset \mathbb{R}^n$  is polyhedral if, and only if, it is finitely generated, i.e., there exist a nonempty and finite set of vectors  $\{v_1, \dots, v_m\}$ , and a finitely generated cone  $K$  such that*

$$\Omega = \{x \in \mathbb{R}^n \mid x = y + \sum_{j=1}^m \mu_j v_j, y \in K, \sum_{j=1}^m \mu_j = 1, \mu_j \geq 0, j = 1, \dots, m\}.$$

Let us note that finitely generated cones are the same as polyhedral cones, because of the well-known Minkowski-Weyl Theorem [49, Theorem 3.52].

**DEFINITION 2.7** (Tangent and Normal cones). *Let  $X$  be a nonempty closed convex set in  $\mathbb{R}^n$  and  $x \in X$ . The tangent cone of  $X$  at  $x$  is given by*

$$\mathcal{T}_X(x) := \text{cl}(\{\lambda(y - x) \in \mathbb{R}^n \mid y \in X, \lambda \in \mathbb{R}_+\}).$$

The normal cone of  $X$  at  $x$  is the set defined by

$$\mathcal{N}_X(x) := \{y \in \mathbb{R}^n \mid \langle y, z - x \rangle \leq 0, z \in X\}.$$

**FACT 2.8** (Tangent cone of polyhedron [49, Theorem 6.46]). *If  $\Omega \subset \mathbb{R}^n$  is a polyhedron defined as in (2.2), then the tangent cone  $\mathcal{T}_\Omega(x)$ , at any point  $x \in \Omega$ , is a polyhedral cone and can be represented as*

$$\mathcal{T}_\Omega(x) = \{w \in \mathbb{R}^n \mid \langle a_i, w \rangle \leq 0, \text{ for } i \in \mathcal{I}(x)\},$$

where  $a_i \in \mathbb{R}^n$  defines the polyhedron  $\Omega$  and  $\mathcal{I}(x) := \{i \mid \langle a_i, x \rangle = \alpha_i\}$  is the active index set of  $\Omega$  at  $x$ .

**FACT 2.9** (Finitely many tangent cones of a polyhedron [20]). *If  $\Omega$  is a polyhedron, then the set of all tangent cones  $\{\mathcal{T}_\Omega(x) \mid x \in \Omega\}$  has finite cardinality. Moreover, for any  $x \in \Omega$ , there exists a radius  $\delta > 0$  such that  $(\mathcal{T}_\Omega(x) + x) \cap \mathcal{B}_\delta(x) = \Omega \cap \mathcal{B}_\delta(x)$ , that is,  $\Omega$  coincides locally with any shifted tangent cone to it.*

Finally, it is noteworthy that the distance between two non-intersecting polyhedral sets is attained.

**FACT 2.10** (Distance between polyhedra is attained [28, Theorem 5]). *Let  $\Omega_1, \Omega_2 \subset \mathbb{R}^n$  be nonempty polyhedra with  $\Omega_1 \cap \Omega_2 = \emptyset$ . Then,  $\text{dist}(\Omega_1, \Omega_2)$  is attained.*

### 3. On the geometry of two disjoint convex sets under error bound condition.

This section is divided in two subsections, gathering key contributions of our paper. In the first subsection, we investigate the geometry under which a single alternating projection step reaches a best approximation pair when the sets are apart from each other. In the second one, we compare well-known regularity conditions from the literature with our proposed error bound.

**3.1. BAP-EB and alternating projection step.** Here we start stating that a single alternating projection step yields a best approximation pair to a polyhedron  $\Omega$  and a hyperplane  $H$ , if  $\Omega \cap H = \emptyset$  and the alternating projection step is taken from a point sufficiently close to  $\text{bap}_H(\Omega)$ .

LEMMA 3.1 (Alternating projection step for polyhedron versus hyperplane under inconsistency). *Consider two nonempty sets  $\Omega, H \subset \mathbb{R}^n$  such that  $\Omega$  is a polyhedron,  $H$  is a hyperplane and  $\Omega \cap H = \emptyset$ . Then, there exists a radius  $r > 0$  such that*

$$(3.1) \quad P_\Omega P_H(z) \in \text{bap}_H(\Omega),$$

for all  $z \in \mathbb{R}^n$  satisfying  $\text{dist}(z, \text{bap}_H(\Omega)) \leq r$ .

*Proof.* For a point  $x \in \Omega$ , let  $\mathcal{T}_\Omega(x)$  denote the tangent cone of  $\Omega$  at  $x$ . Since  $\Omega$  is a polyhedron, the set of all tangent cones  $\{\mathcal{T}_\Omega(x) \mid x \in \Omega\}$  is finite and each  $\mathcal{T}_\Omega(x)$  is a finitely generated cone (see [Facts 2.6, 2.8, and 2.9](#)). In particular, the cardinality of  $\Gamma := \{\mathcal{T}_\Omega(x) \mid x \in \text{bap}_H(\Omega)\}$  is finite since  $\text{bap}_H(\Omega) \subset \Omega$  and each tangent cone in  $\Gamma$  has a finite number of generators. Consider now the collection of all normalized generators with respect to cones in  $\Gamma$  denoted by  $W := \{w \in \mathbb{R}^n \mid \|w\| = 1, \text{ such that } w \text{ belongs to a conic base of some } \mathcal{T}_\Omega(x), x \in \text{bap}_H(\Omega)\}$ .

Let  $d$  be the displacement vector. [Fact 2.2](#) allows us to conveniently categorize the generators in  $W$ . For the disjoint finite sets  $U := \{u \in W \mid \|u\| = 1, \langle u, d \rangle = 0\}$  and  $V := \{v \in W \mid \|v\| = 1, \langle v, d \rangle > 0\}$ , we have  $W = U \cup V$ . The finiteness and definition of  $V$  provide the existence and positivity of

$$(3.2) \quad r := \min_{v \in V} \{\langle v, d \rangle\}.$$

The facts that  $r > 0$  and  $\|d\| > 0$  are key for the statement [\(3.1\)](#).

Take  $z \in \mathbb{R}^n$  arbitrary, but fixed, such that  $\text{dist}(z, \text{bap}_H(\Omega)) \leq r$ . In order to shorten the notation, set  $\bar{z} := P_{\text{bap}_H(\Omega)}(z)$  and  $z_H := P_H(z)$ . Let  $y$  be an element of  $\text{bap}_H(\Omega)$ . [Fact 2.1\(iii\)](#) implies that  $y - d$  and  $\bar{z} - d$  are contained in  $\text{bap}_\Omega(H) \subseteq H$ . Since  $z_H \in H$  and  $H$  is affine,  $y - \bar{z} + z_H = (y - d) - (\bar{z} - d) + z_H \in H$ . The fact that  $H$  is a hyperplane and the characterization of the best approximation for  $\Omega$  and  $H$  imply

$$\langle z_H - \bar{z}, y - \bar{z} \rangle = \langle z - \bar{z}, y - \bar{z} \rangle + \langle z_H - z, (y - \bar{z} + z_H) - z_H \rangle = \langle z - \bar{z}, y - \bar{z} \rangle \leq 0.$$

Thus,  $\bar{z} = P_{\text{bap}_H(\Omega)}(z_H)$ .

Let us now look at the angle between  $z_H - \bar{z}$  and vectors in  $\mathcal{T}_\Omega(\bar{z})$ . Since  $\bar{z}$  is in  $\text{bap}_H(\Omega)$ , all the generators of the tangent cone  $\mathcal{T}_\Omega(\bar{z})$  must be contained in  $W$ . Recall that  $W$  is split as  $U \cup V$  and, therefore, in order to investigate the sign of  $\langle w, z_H - \bar{z} \rangle$ , consider the two cases below:

- (a)  $w$  is a generator of  $\mathcal{T}_\Omega(\bar{z})$  belonging to  $U$ ;
- (b)  $w$  is a generator of  $\mathcal{T}_\Omega(\bar{z})$  belonging to  $V$ .

Case (a). For  $w$  to be a generator of  $\mathcal{T}_\Omega(\bar{z})$  belonging to  $U$ , it must be a generator of  $\mathcal{T}_{\text{bap}_H(\Omega)}(\bar{z})$ , since  $\text{bap}_H(\Omega) \subset H + d$ . Bear in mind that a polyhedron coincides locally with its tangent cone (See [Fact 2.9](#)). So, for some radius  $\varepsilon > 0$ ,  $\text{bap}_H(\Omega) \cap \mathcal{B}_\varepsilon(\bar{z}) = (\mathcal{T}_{\text{bap}_H(\Omega)}(\bar{z}) + \bar{z}) \cap \mathcal{B}_\varepsilon(\bar{z})$ . Thus, for all  $t > 0$  sufficiently small,  $\bar{z} + tw \in \text{bap}_H(\Omega)$  and since  $P_{\text{bap}_H(\Omega)}(z_H) = \bar{z}$ , by projections onto convex sets, we have  $\langle \bar{z} + tw - \bar{z}, z_H - \bar{z} \rangle \leq 0$ , therefore  $\langle w, z_H - \bar{z} \rangle \leq 0$ .

Case (b). This is the case in which the constant  $r > 0$  defined in [\(3.2\)](#) is going to be employed. We now have  $w \in V$ ,

$$\begin{aligned} \langle w, z_H - \bar{z} \rangle &= \langle w, z_H - \bar{z} + d \rangle - \langle w, d \rangle \\ &\leq \|w\| \|z_H + d - \bar{z}\| - \langle w, d \rangle \\ &\leq \|z - \bar{z}\| - \langle w, d \rangle \\ &\leq r - \langle w, d \rangle \leq 0, \end{aligned}$$

where we used Cauchy-Schwarz in the first inequality, the fact that  $w$  is a unit vector and the Pythagoras argument  $\|z_H + d - \bar{z}\|^2 = \|z - \bar{z}\|^2 - \|z_H + d - z\|^2$  in the second one, and the definition of  $r$  given in [\(3.2\)](#) in the last one.

Hence, cases (a) and (b) have shown that  $\langle z_H - \bar{z}, w \rangle \leq 0$  for any normalized generator  $w$  of the cone  $\mathcal{T}_\Omega(\bar{z})$ . This means that the projection of  $z_H$  onto the shifted cone  $\mathcal{T}_\Omega(\bar{z}) + \bar{z}$  is given by  $\bar{z}$ . Since  $\bar{z} \in \Omega$  and  $\Omega \subset \mathcal{T}_\Omega(\bar{z}) + \bar{z}$ , we get that  $P_\Omega(z_H) = \bar{z}$ . Recalling that we set  $z_H = P_H(z)$ , the proof is finished.  $\square$

We remark that replacing the hyperplane  $H$  by an arbitrary polyhedron in the previous lemma does not, in general, guarantee the statement (3.1); see [Example 5.2](#). The key for reaching a best point in a single alternating step essentially relies on the existence of a suitable error bound between the two non-intersecting sets, which in the case of [Lemma 3.1](#) is automatically fulfilled.

The geometrical appeal of [Lemma 3.1](#) will inspire us to formulate the so-called *BAP error bound*, which depends on the definition of optimal supporting hyperplane, next.

**DEFINITION 3.2** (Optimal supporting hyperplane). *Let  $X, Y \subset \mathbb{R}^n$  be closed convex sets such that  $X \cap Y = \emptyset$  and that  $\text{dist}(X, Y)$  is attained. We say that*

$$\mathbb{H}_Y(X) := \{z \in \mathbb{R}^n \mid \langle z - \bar{x}, d \rangle = 0, \bar{x} \in \text{bap}_Y(X)\}$$

*is the optimal supporting hyperplane to  $X$  regarding  $Y$ , where  $d$  is the displacement vector.*

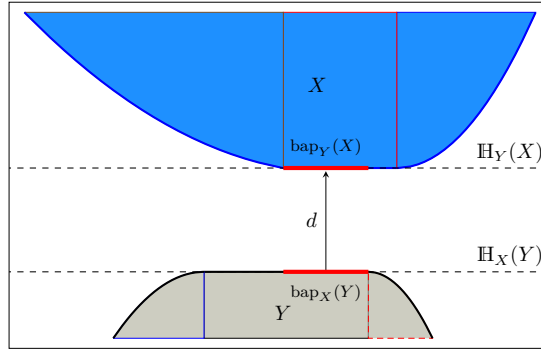


Fig. 2: Disjoint convex sets  $X$  and  $Y$ , BAPs and optimal hyperplanes.

Note that  $\mathbb{H}_Y(X)$  is well-defined since any points  $\hat{x}, \bar{x} \in \text{bap}_Y(X)$  satisfy  $\langle \hat{x} - \bar{x}, d \rangle = 0$ . [Definition 3.2](#) together with [Fact 2.1\(iii\)](#) imply that, for two disjoint closed convex sets  $X, Y \subset \mathbb{R}^n$  with attainable distance,  $\text{bap}_Y(X) = X \cap (Y + d) \subset \mathbb{H}_Y(X)$  and  $\text{bap}_X(Y) = (X - d) \cap Y \subset \mathbb{H}_X(Y)$ . [Figure 2](#) illustrates all these sets related to  $X$  and  $Y$  as well as the displacement vector  $d$ .

We now formally present the BAP error bound in a non-symmetrical version.

**DEFINITION 3.3** (Unilateral BAP error bound). *Let  $X, Y \subset \mathbb{R}^n$  be non-intersecting closed convex sets and assume that the distance between them is attained. We say that  $X$  and  $Y$  satisfy the unilateral BAP error bound (unilateral BAP-EB) at  $x^* \in \text{bap}_Y(X)$  if there exist a bound  $\omega > 0$  and a radius  $\delta > 0$  such that the following inequality holds*

$$(3.3) \quad \omega \text{dist}(x, \text{bap}_Y(X)) \leq \text{dist}(x, \mathbb{H}_Y(X)), \text{ for all } x \in \mathcal{B}_\delta(x^*) \cap X.$$

[Definition 3.3](#) is a condition where the error bound is concentrated in only one of the two sets, therefore, the term *unilateral*. Nevertheless, the responsibility of carrying an error bound can be distributed between the sets  $X$  and  $Y$ . In this regard, we will introduce in [Definition 3.8](#) a more general error bound condition, in which a bilateral concern about the error bound is embedded.

Error bounds are, in general, regularity conditions that allow one to deal with non-isolated solutions; see, for example, [\[18, 19\]](#). The BAP-EB resembles the well-known concept of local



linear regularity [10, 12], formally presented in Definition 3.6. Linear regularity is also known as subtransversality [45]. BAP-EB requires, depending on the context, a further geometrical feature between the sets, in comparison to linear regularity. In Subsection 3.2, we present a detailed discussion on BAP-EB in which we compare it with the standard linear regularity (subtransversality) and with another regularity condition known as intrinsic transversality [39, 40].

Along this subsection, we are going to derive two important lemmas in which we consider the unilateral BAP-EB. They will be valuable tools towards establishing finite convergence of MAP in Section 4 under the more general error bound condition BAP-EB that we introduce later in Definition 3.8.

We present next a result encompassing a broader class of instances than the one in Lemma 3.1, since we are going to consider a closed convex set  $X$  and a hyperplane  $H$  with empty intersection and attainable distance. Here, the optimal supporting hyperplane  $\mathbb{H}_H(X)$  to  $X$  regarding  $H$  coincides precisely with  $H + d$ , that is, the hyperplane obtained by shifting  $H$  by the displacement vector  $d$ . In this context, unilateral BAP-EB from Definition 3.3 is equivalent to local linear regularity. We prove this equivalence in Proposition 3.7. Taking this equivalence into account, the next result states that the standard linear regularity condition leads to finite convergence of alternating projections when the target convex sets are disjoint and one of them is a hyperplane.

**LEMMA 3.4** (Alternating projection step for a convex set versus hyperplane under inconsistency). *Let  $X, H \subset \mathbb{R}^n$  be a closed convex set and a hyperplane, respectively, and suppose that  $X$  and  $H$  are disjoint with attainable distance. Assume that  $X$  and  $H$  satisfy the unilateral BAP error bound (3.3) from Definition 3.3 at  $x^* \in \text{bap}_H(X)$ , that is,*

$$(3.4) \quad \omega \text{dist}(x, \text{bap}_H(X)) \leq \text{dist}(x, \mathbb{H}_H(X)), \text{ for all } x \in \mathcal{B}_\delta(x^*) \cap X,$$

with bound  $\omega > 0$  and radius  $\delta > 0$ . By setting  $r := \min \{ \omega \text{dist}(X, H), \frac{\delta}{2} \}$ , we have that

$$P_X P_H(z) \in \text{bap}_H(X), \text{ for all } z \in \mathcal{B}_r(x^*).$$

*Proof.* Take  $z \in \mathcal{B}_r(x^*)$ , arbitrary, but fixed, with  $x^*$  and  $r$  as enunciated in the hypothesis. Set  $\bar{z} := P_{\text{bap}_H(X)}(z)$  and  $z_H := P_H(z)$  and define

$$S := \{s \in \mathbb{R}^n \mid \langle s - \bar{z}, z_H - \bar{z} \rangle \leq 0\}.$$

Since,  $X \cap H = \emptyset$ , we have that  $z_H - \bar{z} \neq 0$  and thus,  $S$  is an affine half-space. The keystone of the proof is to show that

$$(3.5) \quad X \subset S.$$

If this claim is proved, we get, directly from the characterization of a projection onto a closed convex set, that  $\bar{z} = P_X(z_H) = P_X P_H(z)$ . Therefore, let us draw our attention to proving that (3.5) holds. Assume the contrary, that is, there exists a point  $w \in X$  which does not lie in  $S$ . Then, for a sufficiently small  $t \in (0, 1]$ , we have  $x := tw + (1-t)\bar{z} \in \mathcal{B}_\delta(x^*) \cap X$ . In fact,  $x \in X$ , by convexity. Now, we have that

$$\begin{aligned} \|x - x^*\| &= \|tw + (1-t)\bar{z} - x^*\| = \|t(w - x^*) + (1-t)(\bar{z} - x^*)\| \\ &\leq t\|w - x^*\| + (1-t)\|\bar{z} - x^*\| \\ &= t\|w - x^*\| + (1-t)\|P_{\text{bap}_H(X)}(z) - P_{\text{bap}_H(X)}(x^*)\| \\ &\leq t\|w - x^*\| + (1-t)\|z - x^*\| \\ &< t\|w - x^*\| + (1-t)r \\ &\leq t\|w - x^*\| + (1-t)\frac{\delta}{2} \end{aligned}$$



$$< t \|w - x^*\| + \frac{\delta}{2},$$

where we used the definition of  $x$ , the convexity of the norm, the definition of  $\bar{z}$  and the fact that  $x^* \in \text{bap}_H(X)$ , the nonexpansiveness of the projection, the assumption  $z \in \mathcal{B}_r(x^*)$ , the definition of  $r$ , and the fact that  $1 - t < 1$ , respectively. Let us take a fixed  $t \in \left(0, \min \left\{1, \frac{\delta}{2\|w - x^*\|}\right\}\right]$ . Thus, the correspondent  $x = tw + (1 - t)\bar{z} \in \mathcal{B}_\delta(x^*) \cap X$ . Note also that

$$\begin{aligned} \langle x - \bar{z}, z_H - \bar{z} \rangle &= \langle tw + (1 - t)\bar{z} - \bar{z}, z_H - \bar{z} \rangle \\ &= t \langle w - \bar{z}, z_H - \bar{z} \rangle \\ &> 0, \end{aligned}$$

since  $t > 0$  and  $w \notin S$ . Hence,  $x \notin S$ .

We proceed by showing that this point  $x$  does not comply with the error bound condition (3.3), leading to a contradiction.

Simple Pythagoras arguments imply that  $P_{\text{bap}_H(X)}(z_H) = \bar{z} := P_{\text{bap}_H(X)}(z)$ , as we will see next. Let  $z_{\mathbb{H}} := P_{\mathbb{H}_H(X)}(z)$ , where  $\mathbb{H}_H(X)$  is the optimal supporting hyperplane to  $X$  regarding  $H$  introduced in Definition 3.2. Reminding that, in this case,  $\mathbb{H}_H(X) = H + d$ , we get  $z_{\mathbb{H}} = z_H + d$ . Then,

$$\begin{aligned} \|z - \bar{z}\|^2 - \|\bar{z} - z_{\mathbb{H}}\|^2 &= \|z - z_{\mathbb{H}}\|^2 = \|z - P_{\text{bap}_H(X)}(z_{\mathbb{H}})\|^2 - \|P_{\text{bap}_H(X)}(z_{\mathbb{H}}) - z_{\mathbb{H}}\|^2 \\ &\geq \|z - \bar{z}\|^2 - \|P_{\text{bap}_H(X)}(z_{\mathbb{H}}) - z_{\mathbb{H}}\|^2. \end{aligned}$$

Crossing out  $\|P_{\text{bap}_H(X)}(z_{\mathbb{H}}) - z_{\mathbb{H}}\|^2$  yields  $\|\bar{z} - z_{\mathbb{H}}\| \leq \|P_{\text{bap}_H(X)}(z_{\mathbb{H}}) - z_{\mathbb{H}}\|$ , which gives us  $P_{\text{bap}_H(X)}(z_{\mathbb{H}}) = \bar{z}$ . Similarly, considering the Pythagoras relations, we have

$$\begin{aligned} \|z_H - \bar{z}\|^2 - \|\bar{z} - z_{\mathbb{H}}\|^2 &= \|z_H - z_{\mathbb{H}}\|^2 = \|z_H - P_{\text{bap}_H(X)}(z_H)\|^2 - \|P_{\text{bap}_H(X)}(z_H) - z_{\mathbb{H}}\|^2 \\ &\leq \|z_H - P_{\text{bap}_H(X)}(z_H)\|^2 - \|\bar{z} - z_{\mathbb{H}}\|^2, \end{aligned}$$

because  $\|\bar{z} - z_{\mathbb{H}}\| \leq \|P_{\text{bap}_H(X)}(z_H) - z_{\mathbb{H}}\|$ . After a cancellation, we get

$$\|z_H - \bar{z}\| \leq \|z_H - P_{\text{bap}_H(X)}(z_H)\|,$$

providing  $P_{\text{bap}_H(X)}(z_H) = \bar{z}$ .

The fact that  $P_{\text{bap}_H(X)}(z_H) = \bar{z}$  implies that  $\langle s - \bar{z}, z_H - \bar{z} \rangle \leq 0$ , for all  $s \in \text{bap}_H(X)$ , and we conclude by definition of  $S$  that  $\text{bap}_H(X) \subset S$ . On the other hand, recall that  $\text{bap}_H(X) \subset \mathbb{H}_H(X)$  (see Definition 3.2). In particular,

$$(3.6) \quad \text{bap}_H(X) \subset S \cap \mathbb{H}_H(X).$$

Now, let us define  $\hat{x} := P_{S \cap \mathbb{H}_H(X)}(x)$ . Since  $x \notin S$ , we have  $\langle x - \bar{z}, z_H - \bar{z} \rangle > 0$  and because  $\hat{x}$  lies on the affine half-space  $S$ , it holds that  $\langle \hat{x} - \bar{z}, z_H - \bar{z} \rangle \leq 0$ . Hence,

$$\langle x - \hat{x}, z_H - \bar{z} \rangle = \langle x - \bar{z}, z_H - \bar{z} \rangle + \langle \bar{z} - \hat{x}, z_H - \bar{z} \rangle > 0.$$

Moreover,

$$\begin{aligned} 0 &< \langle x - \hat{x}, z_H - \bar{z} \rangle = \langle x - \hat{x}, z_H - (\bar{z} - d) - d \rangle \\ &= \langle x - \hat{x}, z_H - (\bar{z} - d) \rangle - \langle x - \hat{x}, d \rangle \\ &\leq \|x - \hat{x}\| \|z_H + d - \bar{z}\| - \langle x - \hat{x}, d \rangle \end{aligned}$$

$$\begin{aligned}
&= \|x - \hat{x}\| \|P_{\mathbb{H}_H(X)}(z) - P_{\mathbb{H}_H(X)}(\bar{z})\| - \langle x - \hat{x}, d \rangle \\
&\leq \|x - \hat{x}\| \|z - \bar{z}\| - \langle x - \hat{x}, d \rangle \\
(3.7) \quad &< \|x - \hat{x}\| r - \langle x - \hat{x}, d \rangle,
\end{aligned}$$

where the second inequality is by Cauchy-Schwarz, the third is by the nonexpansiveness of projection and the last one follows from  $\|z - \bar{z}\| \leq \|z - x^*\| < r$ . Therefore,

$$(3.8) \quad \langle x - \hat{x}, d \rangle < r \|x - \hat{x}\| \leq \omega \|d\| \|x - \hat{x}\|,$$

by (3.7) and the definition of  $r$ .

Now, let  $x_H := P_H(x)$  and  $x_{\mathbb{H}} := P_{\mathbb{H}_H(X)}(x) = x_H + d$ . Then,

$$(3.9) \quad \langle x - \hat{x}, d \rangle = \langle x - x_H + x_H - \hat{x}, d \rangle = \langle x - x_H, d \rangle + \langle x_H - \hat{x}, d \rangle = \langle x - x_H, d \rangle,$$

as  $\hat{x}, x_H \in \mathbb{H}_H(X)$  and  $d \perp \mathbb{H}_H(X)$ . Due to the fact that  $\mathbb{H}_H(X)$  is a hyperplane,  $x - x_H$  is collinear to  $d$ , so Cauchy-Schwarz holds sharply, that is,  $|\langle x - x_H, d \rangle| = \|x - x_H\| \|d\|$ . Moreover, since  $x \in X$ , the inner product  $\langle x - x_H, d \rangle$  is nonnegative and thus  $\langle x - x_H, d \rangle = \|x - x_H\| \|d\|$ , which combined with (3.8) and (3.9) provides

$$\|x - x_H\| \|d\| < \omega \|d\| \|x - \hat{x}\|.$$

This inequality, together with  $\text{bap}_H(X) \subset S \cap \mathbb{H}_H(X)$ , as proved in (3.6), yields

$$\begin{aligned}
\text{dist}(x, \mathbb{H}_H(X)) &= \|x - x_H\| \\
&< \omega \|x - \hat{x}\| = \omega \text{dist}(x, S \cap \mathbb{H}_H(X)) \\
&\leq \omega \text{dist}(x, \text{bap}_H(X)),
\end{aligned}$$

which contradicts the error bound assumption (3.4), because  $x \in \mathcal{B}_\delta(x^*) \cap X$  and  $\text{bap}_H(X) = X \cap \mathbb{H}_H(X)$ .  $\square$

The previous result leads to another contribution of this paper. We show that the two ingredients, infeasibility and unilateral BAP-EB, imply that a single alternating projection step can locally reach a best approximation pair.

**LEMMA 3.5** (Alternating projection step under infeasibility and unilateral BAP-EB). *Let  $X, Y \subset \mathbb{R}^n$  be closed convex sets such that  $X \cap Y = \emptyset$  and assume that the distance between them is attained. Assume that  $X$  and  $Y$  satisfy the unilateral BAP error bound (3.3) from Definition 3.3 at  $x^* \in \text{bap}_Y(X)$ , that is,*

$$(3.10) \quad \omega \text{dist}(x, \text{bap}_Y(X)) \leq \text{dist}(x, \mathbb{H}_Y(X)), \text{ for all } x \in \mathcal{B}_\delta(x^*) \cap X,$$

*with bound  $\omega > 0$  and radius  $\delta > 0$ . By setting  $r := \min\{\omega \text{dist}(X, Y), \frac{\delta}{2}\}$  we have that for all  $z \in \mathcal{B}_r(x^*)$ ,*

$$P_X P_Y(z) \in \text{bap}_Y(X).$$

*Proof.* Consider  $x^*$  and  $r$  as stated in the assumptions and let  $z \in \mathcal{B}_r(x^*)$  be arbitrary, but fixed. Now, set  $z_Y := P_Y(z)$  and  $z^\diamond := P_{\mathbb{H}_X(Y)}(z_Y)$ , where  $\mathbb{H}_X(Y)$  is the optimal supporting hyperplane to  $Y$  regarding  $X$ , i.e.,  $\mathbb{H}_Y(X) = \mathbb{H}_X(Y) + d$ , where  $d$  is the displacement vector. Moreover, using the nonexpansiveness of projection operators onto convex sets and that  $x^* - d$  lies in both  $Y$  and  $\mathbb{H}_X(Y)$ , we obtain

$$\|(z^\diamond + d) - x^*\| = \|z^\diamond - (x^* - d)\|$$

$$\begin{aligned}
&= \|P_{\mathbb{H}_X(Y)}(z_Y) - P_{\mathbb{H}_X(Y)}(x^* - d)\| \\
&\leq \|z_Y - (x^* - d)\| \\
&= \|P_Y(z) - P_Y(x^*)\| \\
(3.11) \quad &\leq \|z - x^*\| \leq r,
\end{aligned}$$

that is,  $z^\diamond + d \in \mathcal{B}_r(x^*)$ . Taking into account that  $\text{dist}(X, \mathbb{H}_X(Y)) = \text{dist}(X, Y) = \|d\|$ , [Lemma 3.4](#) can be applied to  $z^\diamond + d$ , with  $\mathbb{H}_X(Y)$  playing the role of  $H$ , yielding  $P_X P_{\mathbb{H}_X(Y)}(z^\diamond + d) \in \text{bap}_{\mathbb{H}_X(Y)}(X)$ . Note that  $\text{bap}_{\mathbb{H}_X(Y)}(X) \subset X \cap \mathbb{H}_Y(X)$  and clearly, the fact that the error bound constant  $\omega$  in [\(3.10\)](#) is strictly positive implies that  $\text{bap}_Y(X) \cap \mathcal{B}_r(x^*) = (X \cap \mathbb{H}_Y(X)) \cap \mathcal{B}_r(x^*)$ . Thus,

$$(3.12) \quad P_X(z^\diamond) \in \text{bap}_Y(X),$$

since  $P_X(z^\diamond) \in \mathcal{B}_r(x^*)$ . This occurs because  $P_X(z^\diamond) = P_X P_{\mathbb{H}_X(Y)}(z^\diamond + d)$ ,  $x^* \in X \cap \mathbb{H}_Y(X)$ , the nonexpansiveness of projections and [\(3.11\)](#) as

$$\begin{aligned}
\|P_X(z^\diamond) - x^*\| &= \|P_X P_{\mathbb{H}_X(Y)}(z^\diamond + d) - P_X P_{\mathbb{H}_X(Y)}(x^*)\| \\
&\leq \|P_{\mathbb{H}_X(Y)}(z^\diamond + d) - P_{\mathbb{H}_X(Y)}(x^*)\| \\
&\leq \|(z^\diamond + d) - x^*\| \\
&\leq r.
\end{aligned}$$

Bearing in mind the definition of  $\mathbb{H}_X(Y)$  and that  $z_Y \in Y$ , observe that there exists  $t \leq 0$  such that  $z_Y - z^\diamond = td$ . For all  $x \in X$ , we have

$$\begin{aligned}
\langle z_Y - P_X(z^\diamond), x - P_X(z^\diamond) \rangle &= \langle z_Y - z^\diamond, x - P_X(z^\diamond) \rangle + \langle z^\diamond - P_X(z^\diamond), x - P_X(z^\diamond) \rangle \\
&= \underbrace{t}_{\leq 0} \underbrace{\langle d, x - P_X(z^\diamond) \rangle}_{\geq 0} + \underbrace{\langle z^\diamond - P_X(z^\diamond), x - P_X(z^\diamond) \rangle}_{\leq 0} \\
(3.13) \quad &\leq 0,
\end{aligned}$$

where the second under-brace remark is by [Fact 2.2](#) and the third is by characterization of the orthogonal projection of  $z^\diamond$  onto  $X$ . Hence, [\(3.13\)](#) implies that  $P_X(z_Y) = P_X(z^\diamond)$ . Therefore, due to [\(3.12\)](#) and that  $P_X P_Y(z) = P_X(z_Y)$ , it holds that

$$P_X P_Y(z) \in \text{bap}_Y(X),$$

proving the theorem.  $\square$

**3.2. BAP-EB and other regularity conditions.** We now proceed to discuss the connection of BAP-EB with two other regularity conditions: local linear regularity, which is the keystone for providing linear convergence of MAP [\[10, Corollary 3.14\]](#), and intrinsic transversality.

First, we show that unilateral BAP-EB implies the standard local linear regularity and moreover, both coincide in the context of [Lemma 3.4](#), that is, when one of the sets is a hyperplane. This is proved in [Proposition 3.7](#) below. This subsection ends with a proof that BAP-EB is more general than intrinsic transversality. For clarity, we present the definitions of linear regularity and intrinsic transversality.

**DEFINITION 3.6** (Local linear regularity [\[10, Definition 3.11\]](#)). *Let  $X, Y \subset \mathbb{R}^n$  be closed convex sets such that  $X \cap Y = \emptyset$ , assume that the distance between them is attained and suppose that  $d$  is the displacement vector. We say that  $X$  and  $Y$  are locally linearly regular if for some point  $x^* \in \text{bap}_Y(X)$  there exist a bound  $\kappa > 0$  and a radius  $\rho > 0$  and such that, for all  $z \in \mathcal{B}_\rho(x^*)$ ,*

$$\text{dist}(z, \text{bap}_Y(X)) \leq \kappa \max\{\text{dist}(z, X), \text{dist}(z, Y + d)\}.$$

PROPOSITION 3.7. *Let  $X, Y \subset \mathbb{R}^n$  be closed convex sets, suppose that  $X$  and  $Y$  are disjoint with attainable distance and let  $d$  be the displacement vector. Then, the unilateral BAP-EB condition from Definition 3.3 implies local linear regularity. If, in addition,  $Y$  is a hyperplane, then local linear regularity is equivalent to the unilateral BAP-EB.*

*Proof.* Assume that the unilateral BAP error bound condition from Definition 3.3 holds, that is, for some point  $x^* \in \text{bap}_Y(X)$  there exist  $\omega > 0$  and  $\delta > 0$  such that, for all  $x \in \mathcal{B}_\delta(x^*) \cap X$ ,

$$\omega \text{dist}(x, \text{bap}_Y(X)) \leq \text{dist}(x, \mathbb{H}_Y(X)).$$

Note that, for all  $x \in X$ , we have  $\text{dist}(x, \mathbb{H}_Y(X)) \leq \text{dist}(x, Y + d)$ . From the fact that  $\text{bap}_Y(X) = X \cap (Y + d)$ , it follows that

$$\text{dist}(x, X \cap (Y + d)) \leq \frac{1}{\omega} \text{dist}(x, Y + d),$$

for all  $x \in \mathcal{B}_\delta(x^*) \cap X$ . Using [10, Lemma 4.1] in the last inequality, with  $X$  playing the role of  $M$  and  $Y + d$  playing the role of  $N$ , we get

$$\text{dist}(z, X \cap (Y + d)) \leq \left( \frac{2}{\omega} + 1 \right) \max\{\text{dist}(z, X), \text{dist}(z, Y + d)\},$$

for all  $z \in \mathcal{B}_{\frac{\delta}{2}}(x^*)$ . Thus, the local linear regularity holds with  $\kappa := \frac{2}{\omega} + 1$  and  $\rho := \frac{\delta}{2}$ .

To complete the proof, we just have to show that, if  $Y$  is a hyperplane, local linear regularity implies (3.3), as in this case  $Y + d$  coincides with  $\mathbb{H}_Y(X)$ , the optimal supporting hyperplane to  $X$  regarding  $Y$ . Therefore, for all points  $z \in X$ ,  $\max\{\text{dist}(z, X), \text{dist}(z, Y + d)\} = \text{dist}(z, Y + d) = \text{dist}(z, \mathbb{H}_Y(X))$ . Hence, the result follows, with  $\omega := \frac{1}{\kappa}$  and  $\delta := \rho$ .  $\square$

Later, in Example 5.2, we will see that if none of the sets is a hyperplane, local linear regularity might not coincide with the unilateral BAP error bound from Definition 3.3.

Next, we introduce a bilateral error bound condition (referred to as BAP-EB), which is more general than the unilateral BAP-EB from Definition 3.3 and also implies finite convergence of MAP under inconsistency; see Theorem 4.5. Furthermore, we will see in Theorem 4.6 that, under inconsistency, a given MAP sequence converges in a finite number of steps if, and only if, the BAP-EB is satisfied along it.

DEFINITION 3.8 (BAP error bound). *Let  $X, Y \subset \mathbb{R}^n$  be non-intersecting closed convex sets and assume that the distance between them is attained. We say that  $X$  and  $Y$  satisfy the BAP error bound (BAP-EB) at  $x^* \in \text{bap}_Y(X)$  if there exist a bound  $\omega > 0$  and a radius  $\delta > 0$  such that, for all  $x \in \mathcal{B}_\delta(x^*) \cap X$ , at least one of the following inequalities holds*

$$(3.14) \quad \begin{aligned} \omega \text{dist}(P_Y(x), \text{bap}_X(Y)) &\leq \text{dist}(P_Y(x), \mathbb{H}_X(Y)), \\ \omega \text{dist}(P_X P_Y(x), \text{bap}_Y(X)) &\leq \text{dist}(P_X P_Y(x), \mathbb{H}_Y(X)). \end{aligned}$$

Before starting to state theorems on finite convergence of MAP and variants, we will compare BAP-EB with intrinsic transversality. Drusvyatskiy presents intrinsic transversality in [39, Definition 3.1] and comments on its importance. The definition originally considers intersecting sets, and says that two closed sets  $X, Y \subset \mathbb{R}^n$  are intrinsically transversal at a common point if there exists an angle  $\alpha \in (0, \frac{\pi}{2}]$  such that, near this common point, any two points  $x \in X \setminus Y$  and  $y \in Y \setminus X$  cannot have difference  $x - y$  simultaneously making an angle strictly less than  $\alpha$  with both the normal cones  $-\mathcal{N}_X(x)$  and  $\mathcal{N}_Y(y)$ .

The concept of intrinsic transversality was adapted for nonintersecting sets in [23]. This nice manuscript first appeared in the form of a preprint on arXiv a few days after the submission of the present paper. The authors of [23] have similar results to ours as they derived finite

convergence of MAP under inconsistency and intrinsic transversality. We will actually prove that intrinsic transversality for nonintersecting convex sets implies BAP-EB. However, the converse is not true. [Example 5.3](#) shows that BAP-EB is more general. More than that, BAP-EB along a MAP sequence is a necessary and sufficient condition for its finite convergence under infeasibility; see [Theorem 4.6](#). Next, we provide the definition of intrinsic transversality for two nonintersecting closed convex sets introduced in [\[23, Condition 1'\]](#).

**DEFINITION 3.9** (Intrinsic transversality for disjoint sets). *Two closed convex sets  $X, Y \subset \mathbb{R}^n$  with empty intersection and attainable distance are intrinsically transversal at  $(x^*, y^*) \in \text{bap}(X, Y)$  if there exist  $\kappa \in (0, 1)$  and  $\delta > 0$  such that*

$$(3.15) \quad \max \left\{ \text{dist} \left( \frac{x - y}{\|x - y\|}, \mathcal{N}_Y(y) \right), \text{dist} \left( \frac{x - y}{\|x - y\|}, -\mathcal{N}_X(x) \right) \right\} \geq \kappa,$$

for all  $x \in X \cap \mathcal{B}_\delta(x^*)$  and  $y \in (Y \setminus \text{bap}_X(Y)) \cap \mathcal{B}_\delta(y^*)$ .

We point out that the previous definition is not symmetrical. It is unilateral, since one is considering local points  $y \in Y$  such that  $\text{dist}(X, y) > \text{dist}(X, Y)$ , but this condition is not required for points in  $X$ .

We now establish that intrinsic transversality is a sufficient condition for BAP-EB to be fulfilled.

**PROPOSITION 3.10** (Intrinsic transversality implies BAP-EB). *If intrinsic transversality from [Definition 3.9](#) is satisfied, then BAP-EB from [Definition 3.8](#) holds.*

*Proof.* We are going to prove the statement by showing that the absence of BAP-EB prevents intrinsic transversality to hold.

Consider two closed and convex sets  $X, Y \in \mathbb{R}^n$  with empty intersection and attainable distance. Let  $(x^*, y^*) \in \text{bap}(X, Y)$  and suppose that BAP-EB does not hold at  $x^*$ . The lack of BAP-EB guarantees the existence of a sequence  $(x^k)_{k \in \mathbb{N}} \subset X$  converging to  $x^*$  such that the sequence defined by  $y^k := P_Y(x^k)$  has no term in  $\text{bap}_X(Y)$  and converges to  $y^*$ , and the sequence determined by  $w^k := P_X P_Y(x^k)$  does not have any term in  $\text{bap}_Y(X)$  and converges to  $x^*$ . The existence of such sequence  $(x^k)_{k \in \mathbb{N}}$  is indeed easy to verify when denying both inequalities [\(3.14\)](#) in BAP-EB.

Obviously,  $x^k - y^k \in \mathcal{N}_Y(y^k)$  and  $w^k - y^k \in -\mathcal{N}_X(w^k)$ . Taking into account the nonexpansiveness of projections, we have  $x^k - y^k \rightarrow x^* - y^*$  and  $w^k - y^k \rightarrow x^* - y^*$ .

Hence,

$$\begin{aligned} \lim_{k \rightarrow \infty} \text{dist} \left( \frac{w^k - y^k}{\|w^k - y^k\|}, \mathcal{N}_Y(y^k) \right) &\leq \lim_{k \rightarrow \infty} \left\| \frac{w^k - y^k}{\|w^k - y^k\|} - \frac{x^k - y^k}{\|x^k - y^k\|} \right\| \\ &= \left\| \frac{x^* - y^*}{\|x^* - y^*\|} - \frac{x^* - y^*}{\|x^* - y^*\|} \right\| = 0, \end{aligned}$$

and

$$\lim_{k \rightarrow \infty} \text{dist} \left( \frac{w^k - y^k}{\|w^k - y^k\|}, -\mathcal{N}_X(w^k) \right) \leq \lim_{k \rightarrow \infty} \left\| \frac{w^k - y^k}{\|w^k - y^k\|} - \frac{w^k - y^k}{\|w^k - y^k\|} \right\| = 0,$$

which invalids [\(3.15\)](#), i.e., intrinsic transversality fails.  $\square$

#### 4. Finite convergence of projection-based methods.

**4.1. Finite convergence of MAP.** In this subsection we present straightforward consequences of [Lemmas 3.1, 3.4](#), and [3.5](#) on MAP. The first theorem establishes finite convergence of MAP for a polyhedron and a disjoint hyperplane. We then show that under local linear regularity,

MAP for a closed convex set and a disjoint hyperplane also converges in a finite number of steps, as for a hyperplane versus a closed convex set, local linear regularity coincides with BAP-EB; recall [Proposition 3.7](#). Then, we see in [Theorem 4.3](#) that infeasibility combined with unilateral BAP-EB implies finite convergence of alternating projections for two closed convex sets. Finally, in [Theorem 4.5](#), we prove the most important result of this work, namely, that BAP-EB from [Definition 3.8](#) under inconsistency also yields finite convergence of MAP.

**THEOREM 4.1** (Finite convergence of MAP for polyhedron versus hyperplane under inconsistency). *Consider two nonempty sets  $\Omega, H \subset \mathbb{R}^n$  such that  $\Omega$  is a polyhedron,  $H$  is a hyperplane and  $\Omega \cap H = \emptyset$ . Let  $x^0 \in \mathbb{R}^n$  be given and  $(x^k)_{k \in \mathbb{N}}$  be the MAP sequence defined by  $x^{k+1} := P_\Omega P_H(x^k)$ . Then,  $(x^k)_{k \in \mathbb{N}}$  converges in finitely many steps to  $\bar{x} \in \text{bap}_H(\Omega)$ .*

*Proof.* [Fact 2.10](#) provides that  $\text{dist}(\Omega, H)$  is attained, as  $H$  is also a polyhedron. Hence, the sequence  $(x^k)_{k \in \mathbb{N}}$  converges to some  $\bar{x} \in \text{bap}_H(\Omega)$  ([Fact 2.3\(i\)](#)). Hence, for a sufficiently large  $\bar{k}$ ,  $\text{dist}(x^{\bar{k}}, \text{bap}_H(\Omega)) \leq r$ , where  $r$  set as in [Lemma 3.1](#). Therefore, by [Lemma 3.1](#),  $x^{\bar{k}+1} = P_\Omega P_H(x^{\bar{k}}) = \bar{x}$ , hence proved.  $\square$

[Theorem 4.1](#) alone encompasses a very relevant setting featuring non-isolated solutions as it concerns a hyperplane versus a polyhedron (see [Example 5.1](#) for an application on linear programming). Below, we will state results for non-intersecting convex sets under BAP-EB conditions beyond the polyhedral-affine setting.

**THEOREM 4.2** (Finite convergence of MAP for a convex set versus hyperplane under inconsistency). *Let  $X, H \subset \mathbb{R}^n$  be a closed convex set and a hyperplane, respectively, and suppose that  $X$  and  $H$  are disjoint with attainable distance. Assume that  $X$  and  $H$  satisfy the BAP error bound [\(3.3\)](#) from [Definition 3.3](#) at  $x^* \in \text{bap}_H(X)$  with bounds  $\omega > 0$  and radius  $\delta > 0$ . By setting  $r := \min\{\omega \text{dist}(X, H), \frac{\delta}{2}\}$ , we have that for any given  $x^0 \in \mathbb{R}^n$ , the MAP sequence  $(x^k)_{k \in \mathbb{N}}$  defined by  $x^{k+1} := P_X P_H(x^k)$  converges to a point  $\bar{x} \in \text{bap}_H(X)$ . Moreover, if there exists an index  $\bar{k} \geq 0$ , such that  $x^{\bar{k}} \in \mathcal{B}_r(x^*)$ , then  $x^{\bar{k}+1} = \bar{x}$ , that is, in this case, MAP converges in at most  $\bar{k} + 1$  steps.*

*Proof.* It is well-known that MAP converges globally for two closed convex sets with attainable distance (see [Fact 2.3\(i\)](#)) and thus, the first part of item (ii) follows. Its second part is a straightforward consequence of [Lemma 3.4](#).  $\square$

**THEOREM 4.3** (Finite convergence of MAP under infeasibility and unilateral BAP-EB). *Let  $X, Y \subset \mathbb{R}^n$  be closed convex sets such that  $X \cap Y = \emptyset$  and assume that the distance between them is attained. Assume that  $X$  and  $Y$  satisfy the unilateral BAP error bound of [Definition 3.3](#) at  $x^* \in \text{bap}_Y(X)$  with bound  $\omega > 0$  and radius  $\delta > 0$ . Set  $r := \min\{\omega \text{dist}(X, Y), \frac{\delta}{2}\}$ , and then we have that for any given  $x^0 \in \mathbb{R}^n$ , the MAP sequence  $(x^k)_{k \in \mathbb{N}}$  defined by  $x^{k+1} := P_X P_Y(x^k)$  converges to a point  $\bar{x} \in \text{bap}_Y(X)$ . Moreover, if there exists an index  $\bar{k} \geq 0$ , such that  $x^{\bar{k}} \in \mathcal{B}_r(x^*)$ , then  $x^{\bar{k}+1} = \bar{x}$ , that is, in this case, MAP converges in at most  $\bar{k} + 1$  steps.*

*Proof.* The result follows directly from [Fact 2.3\(i\)](#) and [Lemma 3.5](#).  $\square$

In order to prove finite convergence of MAP under the bilateral BAP-EB from [Definition 3.8](#), we need the following auxiliary proposition.

**PROPOSITION 4.4.** *Let  $X, Y \subset \mathbb{R}^n$  be closed convex sets such that  $X \cap Y = \emptyset$ , assume that the distance between them is attained. Consider a sequence  $(z^k)_{k \in \mathbb{N}} \subset X$  converging to a point  $z^* \in \text{bap}_Y(X)$ . Assume there exists  $\omega > 0$  such that, for all  $k$ ,*

$$(4.1) \quad \omega \text{dist}(z^k, \text{bap}_Y(X)) \leq \text{dist}(z^k, \mathbb{H}_Y(X)).$$

*Then,*

$$(4.2) \quad \omega \text{dist}(z, \text{bap}_Y(X)) \leq \text{dist}(z, \mathbb{H}_Y(X))$$

holds for all  $z \in \bar{S} := \text{cl}(\text{conv}((z^k)_{k \in \mathbb{N}} \cup \text{bap}_Y(X)))$ . Moreover,  $\text{dist}(\bar{S}, Y) = \text{dist}(X, Y)$ ,  $\text{bap}_Y(\bar{S}) = \text{bap}_Y(X)$  and  $\mathbb{H}_Y(\bar{S}) = \mathbb{H}_Y(X)$ .

*Proof.* We start proving that the convex hull  $S := \text{conv}((z^k)_{k \in \mathbb{N}} \cup \text{bap}_Y(X))$  satisfies the announced error bound (4.2). This will be done by induction over the sets

$$S_k := \text{conv}(\{z^0, z^1, \dots, z^k\} \cup \text{bap}_Y(X)).$$

Note that  $S_0 = \{z_\lambda \in X \mid z_\lambda = \lambda z^0 + (1-\lambda)v, \lambda \in [0, 1], v \in \text{bap}_Y(X)\}$ . Taking arbitrary but fixed  $v \in \text{bap}_Y(X)$  and  $\lambda \in [0, 1]$ , we define accordingly  $z_\lambda \in S_0$  and denote by  $\bar{z}_\lambda := P_{\text{bap}_Y(X)}(z_\lambda)$  and by  $\bar{z}^0 := P_{\text{bap}_Y(X)}(z^0)$ . Note that  $\hat{z}_\lambda := \lambda \bar{z}^0 + (1-\lambda)v \in \text{bap}_Y(X)$ , so we get

$$\begin{aligned} \omega \text{dist}(z_\lambda, \text{bap}_Y(X)) &= \omega \|z_\lambda - \bar{z}_\lambda\| \\ &\leq \omega \|z_\lambda - \hat{z}_\lambda\| \\ &= \omega \|\lambda z^0 + (1-\lambda)v - (\lambda \bar{z}^0 + (1-\lambda)v)\| \\ &= \omega \|\lambda(z^0 - \bar{z}^0)\| \\ &= \lambda \omega \text{dist}(z^0, \text{bap}_Y(X)). \end{aligned} \tag{4.3}$$

Define now  $\tilde{z}_\lambda := P_{\mathbb{H}_Y(X)}(z_\lambda)$ . Since  $\mathbb{H}_Y(X)$  is a hyperplane we get

$$\tilde{z}_\lambda = P_{\mathbb{H}_Y(X)}(\lambda z^0 + (1-\lambda)v) = \lambda P_{\mathbb{H}_Y(X)}(z^0) + (1-\lambda)P_{\mathbb{H}_Y(X)}(v) = \lambda \tilde{z}^0 + (1-\lambda)v,$$

using in the last equality that  $\tilde{z}^0 := P_{\mathbb{H}_Y(X)}(z^0)$  and  $v \in \text{bap}_Y(X) \subset \mathbb{H}_Y(X)$ . Then,

$$\begin{aligned} \text{dist}(z_\lambda, \mathbb{H}_Y(X)) &= \|z_\lambda - \tilde{z}_\lambda\| \\ &= \|(\lambda z^0 + (1-\lambda)v) - (\lambda \tilde{z}^0 + (1-\lambda)v)\| \\ &= \|\lambda(z^0 - \tilde{z}^0)\| \\ &= \lambda \text{dist}(z^0, \mathbb{H}_Y(X)). \end{aligned} \tag{4.4}$$

The hypothesis (4.1) for  $k = 0$  reads as

$$\omega \text{dist}(z^0, \text{bap}_Y(X)) \leq \text{dist}(z^0, \mathbb{H}_Y(X)).$$

Multiplying this inequality by  $\lambda \in [0, 1]$  and combining it with equalities (4.3) and (4.4), gives us the first step of the induction.

Now, note that

$$S_{k+1} = \text{conv}(\{z^0, z^1, \dots, z^k, z^{k+1}\} \cup \text{bap}_Y(X)) = \text{conv}(S_k \cup \{z^{k+1}\}),$$

our induction hypothesis says that all  $u \in S_k$  satisfies

$$\omega \text{dist}(u, \text{bap}_Y(X)) \leq \text{dist}(u, \mathbb{H}_Y(X)) \tag{4.5}$$

and inequality (4.1) specialized for  $k+1$  gives us

$$\omega \text{dist}(z^{k+1}, \text{bap}_Y(X)) \leq \text{dist}(z^{k+1}, \mathbb{H}_Y(X)). \tag{4.6}$$

Consider an arbitrary point in  $S_{k+1}$ . It can be written as  $z_\lambda = \lambda u + (1-\lambda)z^{k+1}$ , for some  $u \in S_k$  and  $\lambda \in [0, 1]$ . Let  $\bar{u}, \bar{z}^{k+1}$  be the distance realizers of  $u, z^{k+1}$  with respect to  $\text{bap}_Y(X)$ , respectively and  $\tilde{u}, \tilde{z}^{k+1}$  be the distance realizers of  $u, z^{k+1}$  regarding  $\mathbb{H}_Y(X)$ , respectively. Note that  $\tilde{z}_\lambda := P_{\mathbb{H}_Y(X)}(z_\lambda) = \lambda \tilde{u} + (1-\lambda)\tilde{z}^{k+1}$ , since  $\mathbb{H}_Y(X)$  is a hyperplane. Moreover, by the same



token,  $u - \tilde{u} = \alpha d$  and  $z^{k+1} - \tilde{z}^{k+1} = \beta d$ , for some  $\alpha, \beta \geq 0$  and where  $d$  is the displacement vector (pointing from  $Y$  to  $X$ ). Thus,

$$\begin{aligned}
 \text{dist}(z_\lambda, \mathbb{H}_Y(X)) &= \|z_\lambda - \tilde{z}_\lambda\| = \|\lambda(u - \tilde{u}) + (1 - \lambda)(z^{k+1} - \tilde{z}^{k+1})\| \\
 &= \|\lambda(\alpha d) + (1 - \lambda)(\beta d)\| = \|(\lambda\alpha + (1 - \lambda)\beta)d\| \\
 &= (\lambda\alpha + (1 - \lambda)\beta) \|d\| = \lambda\alpha \|d\| + (1 - \lambda)\beta \|d\| \\
 &= \lambda\|\alpha d\| + (1 - \lambda)\|\beta d\| = \lambda\|u - \tilde{u}\| + (1 - \lambda)\|z^{k+1} - \tilde{z}^{k+1}\| \\
 (4.7) \quad &= \lambda \text{dist}(u, \mathbb{H}_Y(X)) + (1 - \lambda) \text{dist}(z^{k+1}, \mathbb{H}_Y(X)).
 \end{aligned}$$

Now, the point  $\hat{z}_\lambda := \lambda\bar{u} + (1 - \lambda)\bar{z}^{k+1}$  belongs to  $\text{bap}_Y(X)$  but need not coincide with  $\bar{z}_\lambda$ , the distance realizer of  $z_\lambda$  to  $\text{bap}_Y(X)$ . Nevertheless, we can write

$$\begin{aligned}
 \omega \text{dist}(z_\lambda, \text{bap}_Y(X)) &= \omega \|z_\lambda - \bar{z}_\lambda\| \leq \omega \|z_\lambda - \hat{z}_\lambda\| \\
 &= \omega \|(\lambda u + (1 - \lambda)z^{k+1}) - (\lambda\bar{u} + (1 - \lambda)\bar{z}^{k+1})\| \\
 &\leq \lambda\omega \|u - \bar{u}\| + (1 - \lambda)\omega \|z^{k+1} - \bar{z}^{k+1}\| \\
 &= \lambda\omega \text{dist}(u, \text{bap}_Y(X)) + (1 - \lambda)\omega \text{dist}(z^{k+1}, \text{bap}_Y(X)) \\
 &\leq \lambda \text{dist}(u, \mathbb{H}_Y(X)) + (1 - \lambda) \text{dist}(z^{k+1}, \mathbb{H}_Y(X)) \\
 &= \text{dist}(z_\lambda, \mathbb{H}_Y(X)),
 \end{aligned}$$

where the last inequality is by (4.5) and (4.6) and the last equality is due to (4.7). Therefore, the induction argument is completed and the error bound (4.2) holds for all points in  $S$ .

Next we are going to prove that the error bound (4.2) extends to  $\bar{S}$ , the closure of  $S$ . Take  $s \in \bar{S}$ . Then, there exists a sequence  $(s^\ell)_{\ell \in \mathbb{N}} \subset S$  so that  $s^\ell \rightarrow s$ . Note that we have just proven the error bound for all points in  $S$ , thus all  $s^\ell$  satisfies

$$\omega \text{dist}(s^\ell, \text{bap}_Y(X)) \leq \text{dist}(s^\ell, \mathbb{H}_Y(X)).$$

Taking into account that the distance functions to both sets  $\text{bap}_Y(X)$  and  $\mathbb{H}_Y(X)$  are continuous, we can take limits as  $\ell$  goes to infinity in both sides of the previous inequality, getting the result.

Finally, since  $\text{bap}_Y(X) \subset \bar{S} \subset X$ , we have  $\text{dist}(X, Y) \leq \text{dist}(\bar{S}, Y) \leq \text{dist}(\text{bap}_Y(X), Y) = \text{dist}(X, Y)$ , that is,  $\text{dist}(\bar{S}, Y) = \text{dist}(X, Y)$ . Thus,  $\text{bap}_Y(\bar{S}) = \text{bap}_Y(X)$  and  $\mathbb{H}_Y(\bar{S}) = \mathbb{H}_Y(X)$   $\square$

We are now ready to establish finite convergence of MAP under inconsistency added by BAP-EB.

**THEOREM 4.5** (Finite convergence of MAP under infeasibility and BAP-EB). *Let  $X, Y \subset \mathbb{R}^n$  be closed convex sets such that  $X \cap Y = \emptyset$ , assume that the distance between them is attained and let  $x^0 \in \mathbb{R}^n$  be given. Then, the MAP sequence  $(x^k)_{k \in \mathbb{N}}$ , defined by  $x^{k+1} := P_X P_Y(x^k)$ , converges to some  $x^* \in \text{bap}_Y(X)$ . If BAP-EB from Definition 3.8 is satisfied by  $X$  and  $Y$  at  $x^*$ , then  $(x^k)_{k \in \mathbb{N}}$  converges to  $x^*$  in a finite number of steps.*

*Proof.* The convergence of the MAP sequence to some  $x^* \in \text{bap}_Y(X)$  for two disjoint closed convex sets  $X$  and  $Y$ , assuming its distance is attainable, is due to Fact 2.3(i). Assume BAP-EB as in the statement of the theorem and that the MAP sequence  $(x^k)_{k \in \mathbb{N}}$  does not converge in a finite number of steps. Thus, since MAP iterates satisfy  $\|x^{k+1} - x^*\| \leq \|x^k - x^*\|$ , all the terms  $x^k$  lie outside  $\text{bap}_Y(X)$  and, in particular, they are all different from  $x^*$ .

Note that BAP-EB consists of all points  $x \in X$  near  $x^*$  fulfilling at least one of the two inequalities (3.14) given in Definition 3.8. In the following we divide our proof in two cases.

First, let us assume that there is a subsequence of the MAP sequence satisfying the second inequality in (3.14). So, we have an infinite set of indexes  $J \subset \mathbb{N}$ , such that

$$\omega \text{dist}(P_X P_Y(x^{k-1}), \text{bap}_Y(X)) \leq \text{dist}(P_X P_Y(x^{k-1}), \mathbb{H}_Y(X)), \text{ for all } k \in J,$$

which is the same as

$$(4.8) \quad \omega \operatorname{dist}(x^k, \operatorname{bap}_Y(X)) \leq \operatorname{dist}(x^k, \mathbb{H}_Y(X)), \text{ for all } k \in J.$$

Let us define the closed convex set  $\bar{S} := \operatorname{cl}(\operatorname{conv}((x^k)_{k \in J} \cup \operatorname{bap}_Y(X))) \subset X$ . Clearly,  $x^* \in \bar{S}$  and, since (4.8) holds and  $(x^k)_{k \in J}$  converges to  $x^*$ , Proposition 4.4 can be used. Then, for all  $z \in \bar{S}$ ,

$$\omega \operatorname{dist}(z, \operatorname{bap}_Y(\bar{S})) = \omega \operatorname{dist}(z, \operatorname{bap}_Y(X)) \leq \operatorname{dist}(z, \mathbb{H}_Y(X)) = \operatorname{dist}(z, \mathbb{H}_Y(\bar{S})).$$

This inequality means that the unilateral BAP-EB condition from Definition 3.3 is satisfied for  $\bar{S}$  and  $Y$  at  $x^*$ . Therefore, Lemma 3.5 applies because, for all  $k \in \mathbb{N}$  sufficiently large,  $x^{k-1}$  is near enough to  $x^*$ . Hence, for all large  $k$ , we have  $P_{\bar{S}}P_Y(x^{k-1}) \in \operatorname{bap}_Y(\bar{S}) = \operatorname{bap}_Y(X)$ , where the equality of the BAP sets follows from Proposition 4.4.

Recall that, for all  $k \in J$ ,  $x^k$  belongs to  $\bar{S} \subset X$ . Moreover, for each  $k \in J$  there exists  $j_k \in \mathbb{N}$  such that  $k + j_k \in J$  and thus  $x^{k+j_k} \in \bar{S}$ , and, in addition,  $\|x^{k+j_k} - x^*\| \leq \|x^k - x^*\|$ .

Note that  $x^{k+j_k} = P_X P_Y(x^{k+j_k-1})$  and that  $P_Y(x^{k+j_k-1})$  approaches  $P_Y(x^*)$ , when  $k$  goes to infinity. Thus, for all large  $k \in J$ ,  $P_{\bar{S}}P_Y(x^{k+j_k-1}) \in \operatorname{bap}_Y(X)$ . Then, by the characterization of projection onto convex sets, we get

$$\langle P_Y(x^{k+j_k-1}) - P_{\bar{S}}P_Y(x^{k+j_k-1}), x^{k+j_k} - P_{\bar{S}}P_Y(x^{k+j_k-1}) \rangle \leq 0.$$

Note that both vectors in the above inner product are different from zero, because  $P_{\bar{S}}P_Y(x^{k+j_k-1})$  is in  $\operatorname{bap}_Y(X)$  but not in  $Y$  and  $x^{k+j_k}$  is in  $\bar{S}$  but not in  $\operatorname{bap}_Y(X)$ . Then,

$$\|P_Y(x^{k+j_k-1}) - x^{k+j_k}\| > \|P_Y(x^{k+j_k-1}) - P_{\bar{S}}P_Y(x^{k+j_k-1})\|.$$

On the other hand, since  $\bar{S} \subset X$ , we have

$$(4.9) \quad \begin{aligned} \|P_Y(x^{k+j_k-1}) - P_{\bar{S}}P_Y(x^{k+j_k-1})\| &\geq \|P_Y(x^{k+j_k-1}) - P_X P_Y(x^{k+j_k-1})\| \\ &= \|P_Y(x^{k+j_k-1}) - x^{k+j_k}\| \\ &> \|P_Y(x^{k+j_k-1}) - P_{\bar{S}}P_Y(x^{k+j_k-1})\|, \end{aligned}$$

which is a contradiction.

For second case, suppose now that there exists a subsequence  $(x^k)_{k \in J'}$  of the MAP sequence  $(x^k)_{k \in \mathbb{N}}$  satisfying the first inequality in (3.14), that is,

$$\omega \operatorname{dist}(P_Y(x^k), \operatorname{bap}_X(Y)) \leq \operatorname{dist}(P_Y(x^k), \mathbb{H}_X(Y)), \text{ for all } k \in J'.$$

By defining the sequence  $(y^k)_{k \in J'}$ , where  $y^k := P_Y(x^k)$ , we have

$$(4.10) \quad \omega \operatorname{dist}(y^k, \operatorname{bap}_X(Y)) \leq \operatorname{dist}(y^k, \mathbb{H}_X(Y)), \text{ for all } k \in J'.$$

Then, the proof can be carried out analogously to the one in the first case by changing the roles of  $X$  and  $Y$ . Defining the closed convex set  $\bar{S}' := \operatorname{cl}(\operatorname{conv}((y^k)_{k \in J'} \cup \operatorname{bap}_X(Y))) \subset Y$ , we can use Proposition 4.4, since  $(y^k)_{k \in J'}$  converges to  $y^*$  and (4.10) holds, to get the unilateral BAP-EB now for  $\bar{S}'$  and  $X$  at  $y^*$ . Therefore, Lemma 3.5 applies to prove that  $P_{\bar{S}'}P_X(y^{k-1}) \in \operatorname{bap}_X(\bar{S}') = \operatorname{bap}_X(Y)$ , for all  $k$  sufficiently large. So, we also derive a contradiction similar to (4.9).

Thus, the MAP sequence  $(x^k)_{k \in \mathbb{N}}$  converges to  $x^*$  in a finite number of steps.  $\square$

We finalize this subsection with a necessary and sufficient condition for a MAP sequence to converge in a finite number of steps.

**THEOREM 4.6** (Finite convergence of a MAP sequence under infeasibility and BAP-EB). *Let  $X, Y \subset \mathbb{R}^n$  be closed convex sets such that  $X \cap Y = \emptyset$ , assume that the distance between them is attained. Let  $x^0 \in \mathbb{R}^n$  be given and consider the MAP sequence  $(x^k)_{k \in \mathbb{N}}$ , defined by  $x^{k+1} := P_X P_Y(x^k)$  with limit point  $x^* \in \text{bap}_Y(X)$ . The sequence  $(x^k)_{k \in \mathbb{N}}$  converges in a finite number of steps if, and only if, there exists  $\omega > 0$  such that, for all  $k \in \mathbb{N} \setminus \{0\}$ , at least one of the two inequalities holds*

$$(4.11) \quad \begin{aligned} \omega \text{dist}(P_Y(x^k), \text{bap}_X(Y)) &\leq \text{dist}(P_Y(x^k), \mathbb{H}_X(Y)), \\ \omega \text{dist}(x^{k+1}, \text{bap}_Y(X)) &\leq \text{dist}(x^{k+1}, \mathbb{H}_Y(X)). \end{aligned}$$

*Proof.* Assume that the MAP sequence  $(x^k)_{k \in \mathbb{N}}$  converges in a finite number of steps to  $x^*$  and reaches this limit point for the first time at iteration  $\bar{k} \in \mathbb{N} \setminus \{0\}$ . Thus, for any positive  $\omega$  both inequalities in (4.11) are trivially satisfied for all  $k \geq \bar{k}$ . That said, we are done if  $\bar{k} = 1$ . Suppose  $\bar{k} \geq 2$ . Then, the positive  $\omega$  defined by

$$\omega := \min_{0 \leq k < \bar{k}-1} \left\{ \frac{\text{dist}(x^{k+1}, \mathbb{H}_Y(X))}{\text{dist}(x^{k+1}, \text{bap}_Y(X))} \right\}$$

yields the result, since for all  $1 \leq k < \bar{k}$ , we have  $x^k \notin \text{bap}_Y(X)$  and  $x^k \notin \mathbb{H}_Y(X)$ .

Reciprocally, assume the existence of  $\omega > 0$  as in the statement of the corollary. Then, we can follow the exact same lines of the proof of [Theorem 4.5](#) just replacing the sequence at the beginning of that proof by the MAP sequence  $(x^k)_{k \in \mathbb{N}}$  and the result follows.  $\square$

**4.2. Finite termination of other projection methods.** In this subsection, we present extensions of some results of [Subsection 4.1](#). We address two families of convex feasibility problems. The first deals with two closed convex sets with empty intersection, while the second concerns finite number of closed convex sets having no point in common.

**4.2.1. Best-pair tracking.** Here we discuss how our results apply to a broader class of methods tracking best approximation pairs. Interestingly, one may have a method generating a sequence of infinitely many distinct points for which a MAP step intrinsically provides a finite termination criterion. This is formally presented in the next theorem which, in turn, will serve to identify finite termination of the well-known Cimmino method and the famous Douglas-Rachford method.

For this subsection, let  $X, Y \subset \mathbb{R}^n$  be closed convex sets such that  $X \cap Y = \emptyset$  and assume that the distance between  $X$  and  $Y$  is attained.

**LEMMA 4.7** (Intrinsic finite termination). *For a given  $z^0 \in \mathbb{R}^n$ , consider a method generating a sequence  $(z^k)_{k \in \mathbb{N}}$  such that the shadow sequence onto  $X$ ,  $(P_X(z^k))_{k \in \mathbb{N}}$ , converges to a point  $\hat{x} \in \text{bap}_Y(X)$ . If the unilateral BAP-EB from [Definition 3.3](#) is satisfied at  $\hat{x}$ , then there exists an index  $\hat{k} \geq 0$ , such that for all  $k \geq \hat{k}$ , we have  $P_X P_Y(P_X(z^k)) = P_X P_Y(P_X(\hat{x})) \in \text{bap}_Y(X)$ .*

*Proof.* The result is a consequence of [Lemma 3.5](#). Indeed, let  $\hat{x}$  play the role of  $x^*$  in [Lemma 3.5](#). Since the shadow sequence  $(P_X(z^k))_{k \in \mathbb{N}}$  converges to  $\hat{x}$ , there exists  $\hat{k}$  such that  $P_X(z^k) \in \mathcal{B}_r(\hat{x})$  for all  $k \geq \hat{k}$  where  $r = \min\{\omega \text{dist}(X, Y), \frac{\delta}{2}\}$ . Hence, [Lemma 3.5](#) can be employed and the result follows.  $\square$

We proceed by enforcing [Lemma 4.7](#) for the Cimmino method [\[29\]](#).

**COROLLARY 4.8** (Cimmino's intrinsic finite termination). *For a given  $z^0 \in \mathbb{R}^n$ , consider the Cimmino sequence  $(z^k)_{k \in \mathbb{N}}$  given by  $z^{k+1} := \frac{1}{2}(P_X(z^k) + P_Y(z^k))$ . Then, there exists a best approximation pair  $(\bar{x}, \bar{y}) \in \text{bap}(X, Y)$  such that  $z_k \rightarrow \frac{1}{2}(\bar{x} + \bar{y})$ . If, in addition, the unilateral BAP-EB from [Definition 3.3](#) is satisfied at  $\bar{x}$ , then there exists an index  $\hat{k} \geq 0$ , such that for all  $k \geq \hat{k}$ , we have  $P_X P_Y(P_X(z^k)) = P_X P_Y(P_X(\bar{x})) \in \text{bap}_Y(X)$ .*

*Proof.* The convergence of the sequence  $(z^k)_{k \in \mathbb{N}}$  to the midpoint of a best approximation pair can be found in [11, Theorem 6.3]. Then, of course, both shadow sequences  $(P_X(z^k))_{k \in \mathbb{N}}$  and  $(P_Y(z^k))_{k \in \mathbb{N}}$  converge to best points  $\bar{x}$  and  $\bar{y}$ , respectively. Thus, the hypotheses of Lemma 4.7 are fulfilled, providing the corollary.  $\square$

The last result of this section shows that we can as well suit Lemma 4.7 for the Douglas-Rachford method [38]. Although this method always diverges under inconsistency [15, Theorem 3.13(ii)], one of its shadows detects a best point. This enables us to apply Lemma 4.7.

**COROLLARY 4.9** (Douglas-Rachford's intrinsic finite termination). *For a given  $z^0 \in \mathbb{R}^n$ , consider the Douglas-Rachford sequence  $(z^k)_{k \in \mathbb{N}}$  given by  $z^{k+1} := \frac{1}{2}(z^k + R_Y R_X(z^k))$ , where  $R_X := 2P_X - \text{Id}$  and  $R_Y := 2P_Y - \text{Id}$  are the reflectors through  $X$  and  $Y$ , respectively. Then, there exists  $\bar{x} \in \text{bap}_Y(X)$  such that the shadow sequence  $(P_X(z^k))_{k \in \mathbb{N}}$  converges to  $\bar{x}$ . If, in addition, the unilateral BAP-EB from Definition 3.3 is satisfied at  $\bar{x}$ , then there exists an index  $\hat{k} \geq 0$ , such that for all  $k \geq \hat{k}$ , we have  $P_X P_Y(P_X(z^k)) = P_X P_Y(P_X(z^{\hat{k}})) \in \text{bap}_Y(X)$ .*

*Proof.* The convergence of the shadow sequence  $(P_X(z^k))_{k \in \mathbb{N}}$  to a point  $\bar{x} \in \text{bap}_Y(X)$  is ensured by [15, Theorem 3.13 and Remark 3.14(ii)]. So, Lemma 4.7 applies and the result follows.  $\square$

Finally, it is worth commenting that the famous Dykstra's method is also contemplated by Lemma 4.7, as it generates a sequence converging to a best approximation pair; see [11, Theorem 3.7].

**4.2.2. Inconsistent multi-set intersection.** In this section, we consider a finite number of closed convex sets  $X_1, X_2, \dots, X_m \subset \mathbb{R}^n$ . If these sets have nonempty intersection, a point common to them is found, for instance, by the method of Cyclic projections [14, Corollary 5.26] (which is an extension of MAP) or by the Cimmino method [29] applied to multi-set intersection problems (also called simultaneous projection method or method of barycenters); for further discussions see [25, 26]. When  $X_1 \cap X_2 \cap \dots \cap X_m = \emptyset$  and one of the target sets is bounded, the method of Cyclic projections converge to a point that provides what we call a cycle (see [42, Theorem 2], [13, Theorem 5.4.1], [14, Corollary 5.24]) and the Cimmino method converges to a point minimizing the sum of the squares of the distances to the target sets [31, Theorem 4].

A point  $\bar{y} \in \mathbb{R}^n$  is said to provide a *cycle* with respect to the index order  $\{1, 2, \dots, m\}$  if  $P_{X_1}(\bar{y})$  is a fixed point of the Cyclic projection operator  $P_{X_1} P_{X_2} \dots P_{X_m}$ . The correspondent cycle is the tuple  $(\bar{y}_1, \bar{y}_2, \dots, \bar{y}_m) \in X_1 \times X_2 \times \dots \times X_m$ , such that  $\bar{y}_1 = P_{X_1}(\bar{y})$  and

$$\bar{y}_1 = P_{X_1}(\bar{y}_2), \dots, \bar{y}_{m-1} = P_{X_{m-1}}(\bar{y}_m), \bar{y}_m = P_{X_m}(\bar{y}_1).$$

Note that in the two-set case, cycles reduce to best approximation pairs. Also, any best approximation pair is a cycle in both possible index orders.

Our aim is to establish sufficient conditions for finite convergence of Cyclic projections and the Cimmino method when the underlying sets have empty intersection. This is done in two statements, both connected to Lemma 3.5 and Theorem 4.3. In the first one, we look at Cyclic projections. The second theorem explores the bond between the Cimmino method for the multi-set intersection problem and MAP applied to Pierra's product space reformulation.

It is worth reemphasizing that a Cimmino limit point has the nice variational characterization of minimizing the sum of the squares of the distances to the target sets. We point out that cycles coming from Cyclic projections do not have a variational characterization; this was proven by Baillon, Combettes, and Cominetti in [7]. Recent advances in the study of Cyclic projections were derived in [2].

Next, we present a result tracking limit points of Cyclic projections in a finite number of iterations.

**COROLLARY 4.10** (Finite convergence of Cyclic projections). *Let  $X_1, X_2, \dots, X_m \subset \mathbb{R}^n$  be*

closed convex sets with empty intersection. Assume also that a cycle exists, that is, the operator  $P_{X_1}P_{X_2}\cdots P_{X_m}$  has a fixed point. Let  $x^0 \in \mathbb{R}^n$  be given and consider the Cyclic projection sequence  $(x^k)_{k \in \mathbb{N}}$  defined by

$$x^{k+1} := P_{X_1}P_{X_2}\cdots P_{X_m}(x^k).$$

- (i) Then,  $(x^k)_{k \in \mathbb{N}}$  converges to a point  $\bar{y}_1 \in X_1$ , which provides the cycle  $(\bar{y}_1, \bar{y}_2, \dots, \bar{y}_m)$ , where  $\bar{y}_1 = P_{X_1}(\bar{y}_2)$ ,  $\dots$ ,  $\bar{y}_{m-1} = P_{X_{m-1}}(\bar{y}_m)$ ,  $\bar{y}_m = P_{X_m}(\bar{y}_1)$ .
- (ii) Suppose there exists an index  $\bar{i} \in \{1, \dots, m\}$  such that the cycle points  $\bar{y}_{\bar{i}}$  and  $\bar{y}_{\bar{i}+1}$  form a best approximation pair to  $X_{\bar{i}}$  and  $X_{\bar{i}+1}$ , that is,  $(\bar{y}_{\bar{i}}, \bar{y}_{\bar{i}+1}) \in \text{bap}(X_{\bar{i}}, X_{\bar{i}+1})$  and assume that  $X_{\bar{i}}$  and  $X_{\bar{i}+1}$  satisfy the BAP error bound (3.3) from Definition 3.3 at  $\bar{y}_{\bar{i}} \in \text{bap}_{X_{\bar{i}+1}}(X_{\bar{i}})$ , that is,

$$(4.12) \quad \omega \text{dist}(x, \text{bap}_{X_{\bar{i}+1}}(X_{\bar{i}})) \leq \text{dist}(x, \mathbb{H}_{X_{\bar{i}+1}}(X_{\bar{i}})), \text{ for all } y \in \mathcal{B}_\delta(\bar{y}_{\bar{i}}) \cap X_{\bar{i}},$$

with bound  $\omega > 0$  and radius  $\delta > 0$ , and if  $\bar{i} = m$ ,  $\bar{y}_{\bar{i}+1} = \bar{y}_1$ . Then, we have the existence of an index  $\bar{k} \geq 0$ , such that  $x^k = \bar{y}_1$  for all  $k \geq \bar{k}$ , that is, in this case, the Cyclic projection sequence  $(x^k)_{k \in \mathbb{N}}$  converges in at most  $\bar{k}$  steps.

*Proof.* The convergence of the method of Cyclic projections, under the existence of a cycle, to a point  $\bar{y}_1 \in X_1$  that lies in  $\text{Fix } P_{X_1}P_{X_2}\cdots P_{X_m}$  can be found in [42]. It is straightforward to see that  $\bar{y}_1 = P_{X_1}(\bar{y}_2)$ ,  $\dots$ ,  $\bar{y}_{m-1} = P_{X_{m-1}}(\bar{y}_m)$ ,  $\bar{y}_m = P_{X_m}(\bar{y}_1)$ . Thus, item (i) follows.

Now, the proof of item (ii) relies entirely on Lemma 3.5, except for a few details we are going to check. Assume, without loss of generality, that  $\bar{i} = 1$ .

We will show the existence of a radius  $r > 0$  such that for all  $z \in \mathcal{B}_r(\bar{y}_2) \cap X_2$  we have  $P_{X_1}(z) = \bar{y}_1$ , which suffices to establish item (ii). This will indeed complete the proof, because  $(x^k)_{k \in \mathbb{N}}$  converging to  $\bar{y}_1$  implies that the sequence  $(P_{X_2}\cdots P_{X_m}(x^k))_{k \in \mathbb{N}}$  converges to  $\bar{y}_2$ , minding the continuity of projections. Hence, there exists an index  $\bar{k} \geq 0$  such that  $P_{X_2}\cdots P_{X_m}(x^{\bar{k}-1}) \in \mathcal{B}_r(\bar{y}_2)$ . Thus,

$$x^{\bar{k}} = P_{X_1}P_{X_2}\cdots P_{X_m}(x^{\bar{k}-1}) = \bar{y}_1$$

and then for all  $k \geq \bar{k}$ , we have  $x^k = \bar{y}_1$  because  $\bar{y}_1$  is a fixed point of the Cyclic projection operator  $P_{X_1}P_{X_2}\cdots P_{X_m}$ .

Lemma 3.5 guarantees the existence of a radius  $r > 0$  such that for all  $v \in \mathcal{B}_r(\bar{y}_1)$ ,  $P_{X_1}P_{\mathbb{H}_{X_1}(X_2)}(v) \in \text{bap}_{\mathbb{H}_{X_1}(X_2)}(X_1) = \text{bap}_{X_2}(X_1)$ . Take an arbitrary but fixed  $v \in \mathcal{B}_r(\bar{y}_1)$  and define  $z = P_{X_2}(v)$ . By the nonexpansiveness of projections,  $z \in \mathcal{B}_r(\bar{y}_2)$ . Now when projecting this point  $z$  onto  $X_1$  we cross the optimal supporting hyperplane  $\mathbb{H}_{X_1}(X_2)$  at  $z^\circ$ .

Note that  $z^\circ + d$  belongs to  $\mathcal{B}_r(\bar{y}_1)$ , where  $d = \bar{y}_1 - \bar{y}_2$  is the displacement vector. Furthermore,  $P_{X_1}(z) = P_{X_1}(z^\circ) = P_{X_1}P_{\mathbb{H}_{X_1}(X_2)}(z^\circ + d) \in \text{bap}_{\mathbb{H}_{X_1}(X_2)}(X_1) = \text{bap}_{X_2}(X_1)$ . Therefore,  $P_{X_1}(z) \in \text{bap}_{X_2}(X_1)$  must be  $\bar{y}_1$  because of the convergence of the sequence  $(x_k)_{k \in \mathbb{N}}$  to  $\bar{y}_1$ .  $\square$

We have just investigated the obtainment of a cycle by Cyclic projections in a finite number of iterations. It is easy to see that our hypotheses are sufficient but not necessary. Indeed, if one takes three non-intersecting parallel lines in  $\mathbb{R}^2$ , the method of Cyclic projections always captures a cycle in a finite number of iterations, although some hypotheses of Corollary 4.10 are not fulfilled, including the error bound condition (4.12). Note that this error bound holds, for instance, when the normal cone regarding one of the targets sets, say  $X_i$ , at a respective cycle point  $\bar{y}_i$  contains a neighborhood of the previous cycle point  $\bar{y}_{i+1} \in X_{i+1}$ . However, the result in Corollary 4.10 has to be taken with moderate enthusiasm, as even for a very simple problem with plenty of structure, Cyclic projections may not converge finitely when we have more than two sets. Consider, for instance, three segments as target sets forming an equilateral triangle in  $\mathbb{R}^2$  and fix an order for the sets. The associated cycle is not archived by cyclic projections in finite number of steps for every point near the cycle.

We now move towards the investigation of finite convergence of the Cimmino method. The Cimmino iteration regarding the sets  $X_1, X_2, \dots, X_m$  computed at a point  $z^k \in \mathbb{R}^n$  is given by  $C(z^k) := \frac{1}{m}(P_{X_1} + P_{X_2} + \dots + P_{X_m})(z^k)$ . We remark that this iteration does not depend on any index order, as it works in parallel. Moreover, it can be seen as a MAP step under Pierra's product space reformulation [48]. Pierra's approach considers the Cartesian product  $\mathbf{X} := X_1 \times X_2 \times \dots \times X_m$  and the diagonal subspace  $\mathbf{D} := \{(z, \dots, z) \in \mathbb{R}^{mn} \mid z \in \mathbb{R}^n\}$ . Clearly, the intersection of  $X_1, X_2, \dots, X_m$  is related to the intersection of  $\mathbf{X}$  and  $\mathbf{D}$ , that is,  $z \in \bigcap_{i=1}^m X_i$  if, and only if, the diagonal point  $(z, \dots, z) \in \mathbf{D}$  lies in  $\mathbf{X}$ .

It is well-known that, for any  $z^k \in \mathbb{R}^n$ ,

$$P_{\mathbf{D}}P_{\mathbf{X}}(z^k, z^k, \dots, z^k) = (C(z^k), C(z^k), \dots, C(z^k));$$

see [48, Theorem 1.1]. Hence, Cimmino iterations can indeed always be regarded as MAP steps. This allows us to connect Lemma 3.5 with the sought of a Cimmino limit point in a finite number of iterations. Formally, we have the following result.

**COROLLARY 4.11** (Finite termination of multi-set Cimmino (MAP in product space)). *Let  $X_1, X_2, \dots, X_m \subset \mathbb{R}^n$  be closed convex sets with no point in common, assume that at least one the sets is bounded and define  $\mathbf{D} := \{(z, \dots, z) \in \mathbb{R}^{mn} \mid z \in \mathbb{R}^n\}$  and  $\mathbf{X} := X_1 \times X_2 \times \dots \times X_m$ . For a given  $x^0 \in \mathbb{R}^n$ , set  $\mathbf{x}^0 := (x^0, \dots, x^0) \in \mathbf{D}$  and consider the MAP sequence  $(\mathbf{x}^k)_{k \in \mathbb{N}}$  such that  $\mathbf{x}^{k+1} := P_{\mathbf{D}}P_{\mathbf{X}}(\mathbf{x}^k)$ .*

- (i) *Then, the MAP iterates  $\mathbf{x}^k = (x^k, \dots, x^k)$  converge to a point  $\hat{\mathbf{x}} := (\hat{x}, \dots, \hat{x}) \in \text{bap}_{\mathbf{X}}(\mathbf{D})$ , with  $\hat{x} \in \mathbb{R}^n$  being the limit point of the Cimmino sequence  $(x^k)_{k \in \mathbb{N}}$  ruled by  $x^{k+1} := C(x^k) = \frac{1}{m} \sum_{i=1}^m P_{X_i}(x^k)$ .*
- (ii) *The Cimmino limit point  $\hat{x}$  is a solution of the least squares problem of minimizing the sum of the squares of the distances to the target sets  $X_1, \dots, X_m$ .*
- (iii) *If, in addition, there exists an error bound  $\omega > 0$  and a radius  $\delta > 0$  such that*

$$\omega \text{dist}(x, \text{bap}_{\mathbf{X}}(\mathbf{D})) \leq \text{dist}(x, \mathbf{H}_{\mathbf{X}}(\mathbf{D})),$$

*for all  $\mathbf{x} \in \mathcal{B}_{\delta}(\hat{\mathbf{x}}) \cap \mathbf{D}$ , where  $\mathbf{H}_{\mathbf{X}}(\mathbf{D})$  is the optimal supporting hyperplane to  $\mathbf{D}$  regarding  $\mathbf{X}$  given in Definition 3.2, then there exists an index  $\bar{k} \geq 0$ , such that  $\mathbf{x}^k = \hat{\mathbf{x}}$ , for all  $k \geq \bar{k}$ .*

*Proof.* Since there is no common point to the target sets  $X_1, \dots, X_m$ , it follows that  $\mathbf{D} \cap \mathbf{X} = \emptyset$ . We note that, by [31, Proposition 4],  $\text{bap}_{\mathbf{X}}(\mathbf{D}) = \text{Fix } P_{\mathbf{D}}P_{\mathbf{X}} = \{(z, \dots, z) \in \mathbb{R}^{mn} \mid z \in \text{Fix } C\}$ , where  $\text{Fix } C$  is the set of fixed points of the Cimmino operator  $C$ . Now, since one of the sets  $X_i$  is bounded, [31, Proposition 7] guarantees that  $\text{bap}_{\mathbf{X}}(\mathbf{D}) \neq \emptyset$ . In particular, the distance between  $\mathbf{D}$  and  $\mathbf{X}$  is attained. That said, items (i) and (ii) are completely covered by [31, Theorems 1 and 4] and then item (iii) follows as a direct application of Lemma 3.5, with  $\mathbf{D}$  playing the role of  $X$  and  $\mathbf{X}$  playing the role of  $Y$ .  $\square$

One has to humble their expectations as for the range of instances covered by Corollary 4.11, since BAP-EB in the original space may not be transported to the product space. We present, in Example 5.4, one problem where BAP-EB is transported to the product space and one where it is not.

**5. Examples and applications discussing our results.** This section illustrates our results upon insightful examples, one of which raises a conjecture.

We start connecting Linear Programming with Lemma 3.1 and Theorem 4.1. Second, we present an example showing that local linear regularity (subtransversality) combined with infeasibility is not sufficient for finite convergence of MAP. Example 5.4 contains an important remark on BAP-EB in Pierra's product space. We then present an example in which we raise a conjecture regarding the substitution of the Lipschitzian error bound of BAP-EB by a Hölder one.



Finally, we look at an application of MAP for non-smooth convex minimization given by min-max problems.

*Example 5.1* (MAP for LP). A curious geometrical interpretation of Linear Programming (LP) follows from [Lemma 3.1](#) and [Theorem 4.1](#). Consider the following linear program

$$(5.1) \quad \begin{aligned} & \text{minimize} && \langle c, x \rangle \\ & \text{subject to} && \Omega := \{x \in \mathbb{R}^n \mid Ax \leq b, x \geq 0\} \end{aligned}$$

where  $A \in \mathbb{R}^{m \times n}$ ,  $c \in \mathbb{R}^n$  with  $c \neq 0$  and  $b \in \mathbb{R}^m$  are given. Let us assume that (5.1) is solvable and denote its solution set by  $\Omega^*$ . These hypotheses guarantee the existence of a dual feasible point  $\hat{y} \in \mathbb{R}^m$ , i.e.,  $A^\top \hat{y} \geq c$  and  $\hat{y} \geq 0$ . Take any  $\varepsilon > 0$ , some  $\hat{x} \in \mathbb{R}^m$  such that  $\langle c, \hat{x} \rangle = \langle b, \hat{y} \rangle - \varepsilon$ , and define  $H := \{x \in \mathbb{R}^n \mid \langle c, \hat{x} - x \rangle = 0\}$ . Note that finding such  $\hat{x}$  is a trivial task. Now, strong duality yields  $\Omega \cap H = \emptyset$ , so [Theorem 4.1](#) applies. This means that for any starting point  $x^0 \in \mathbb{R}^n$  the MAP sequence  $x^{k+1} := P_\Omega P_H(x^k)$  finds a best point  $\bar{x} \in \Omega$  after a finite number of iterations. It is easy to see that the displacement vector  $d = \bar{x} - P_H(\bar{x})$  between  $H$  and  $\Omega$  is a positive multiple of the cost  $c$ . Therefore,  $\bar{x}$  solves the LP (5.1). We could indeed find  $\bar{x}$  in a single MAP step by pushing the hyperplane  $H$  sufficiently away from  $\Omega$  in the direction of  $-c$ . A similar conclusion is derived in [47] for LP using a projection-based scheme, nevertheless uniqueness of primal and dual solutions of (5.1) is required. In turn, our results hold regardless of solutions being unique.

We now present an example showing that, in general, inconsistency combined with local linear regularity (subtransversality) from [Definition 3.6](#) alone does not imply finite convergence of MAP. The example features two disjoint reverse lines forming forty-five degrees. In this example, BAP-EB does not hold and thus, neither does intrinsic transversality.

*Example 5.2* (Linear regularity does not imply finite convergence of MAP). Let

$$X := \{(t, 0, 0) \in \mathbb{R}^3 \mid t \in \mathbb{R}\} \quad \text{and} \quad Y_\gamma := \{(t, t, \gamma) \in \mathbb{R}^3 \mid t \in \mathbb{R}\},$$

with  $\gamma \in \mathbb{R}$  fixed. In this example, the unique best approximation pair consists of  $(\bar{x}, \bar{y})$ , where  $\bar{x} = (0, 0, 0)$  and  $\bar{y} = (0, 0, \gamma)$ , but unilateral BAP-EB from [Definition 3.3](#) is not satisfied, for  $\gamma \neq 0$ . Indeed, in this case, the optimal supporting hyperplane to  $X$  regarding  $Y_\gamma$  is given by  $\mathbb{H}_Y(X) = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_3 = 0\}$  and obviously,  $X \cap Y_\gamma = \emptyset$ . Here,  $\text{dist}((0, 1, 0), \text{bap}_{Y_\gamma}(X)) = 1$ , however  $\text{dist}((0, 1, 0), \mathbb{H}_Y(X)) = 0$ , which prohibits the error bound inequality (3.3) to hold if  $\gamma \neq 0$ . Let us now investigate the behavior of MAP in this case. Consider an arbitrary starting point  $x^0 = (a, b, c) \in \mathbb{R}^3$  and the MAP sequence  $x^{k+1} = P_X P_{Y_\gamma}(x^k)$ . Straightforward manipulations give us

$$x^k = \left( \frac{a}{2^k}, 0, 0 \right).$$

This sequence converges linearly with rate  $1/2$  to the unique best point to  $X$ , namely the origin  $(0, 0, 0)$ . Note that the sequence does not depend on the parameter  $\gamma$ , showing that, in this example, infeasibility does not play a role at all. Thus, infeasibility together with polyhedrality is not enough to guarantee finite convergence of MAP. This example also shows that the hyperplane in both [Lemma 3.1](#) and [Theorem 4.1](#) cannot be replaced, in general, by an arbitrary polyhedron. This is not even the case when the set  $H$  in the theorem and corollary is affine only. Furthermore, this example satisfies local linear regularity, and thus this condition is not sufficient for finite convergence of MAP.

The next example shows that BAP-EB is fulfilled for more problems than intrinsic transversality.



*Example 5.3* (BAP-EB is more general than intrinsic transversality). Let us look at the sets  $X := \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 - x_2 + 1 \leq 0 \text{ and } -x_1 - x_2 + 1 \leq 0\}$  and  $Y := \{(y_1, y_2) \in \mathbb{R}^2 \mid y_1 + y_2 \leq 0 \text{ and } y_1^2 + y_2 \leq 0\}$ . They have empty intersection, only a single best approximation pair is associated to them, namely  $(x^*, y^*)$  with  $x^* = (1, 0)$  and  $y^* = (0, 0)$ , and clearly BAP-EB from [Definition 3.8](#) is satisfied. Thus, MAP applied to  $X$  and  $Y$  converges finitely. However, intrinsic transversality fails bilaterally. In order to see that, take the sequences determined by  $x^k := (\frac{1}{k}, \frac{1}{k^2} + 1) \in X \setminus \{x^*\}$  and  $y^k := (\frac{-1}{k}, \frac{-1}{k^2}) \in Y \setminus \{y^*\}$  and note that the unit vector  $(\frac{-2}{\sqrt{k^2+4}}, \frac{k}{\sqrt{k^2+4}})$  lies in  $-\mathcal{N}_X(x^k) \cap \mathcal{N}_Y(y^k)$ . Hence, intrinsic transversality does not hold, even if  $X$  and  $Y$  switch roles in [Definition 3.9](#).

The aim of the following example is to show that BAP-EB in the original space is not necessarily transported to the product space.

*Example 5.4* (BAP-EB in the original space and in the product space). Consider the polyhedron  $X := \{(x_1, x_2) \in \mathbb{R}^2 \mid |x_1| - x_2 + 2 \leq 0\}$ , the polyhedron  $Y := \{(y_1, y_2) \in \mathbb{R}^2 \mid |y_1| + y_2 \leq 0\}$  and the hyperplane  $H := \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 = 0\}$ . Note that  $\text{bap}(X, Y) = \text{bap}(X, H) = \{(x^*, y^*)\}$ , where  $x^* = (0, 2)$  and  $y^* = (0, 0)$ . MAP applied to either  $H$  and  $X$ , or  $Y$  and  $X$  has finite convergence wherever one starts in  $\mathbb{R}^2$ . Cimmino, however, only converges finitely to  $(x^* + y^*)/2 = (0, 1)$  when applied to the pair  $Y$  and  $X$ . The relation between MAP and Cimmino discussed in [Subsection 4.2](#) thus guarantees that MAP converges finitely in Pierra's product space regarding  $Y$  and  $X$ . Actually, the BAP-EB between  $Y$  and  $X$  is transported to the product space. Nevertheless, BAP-EB holds for  $H$  and  $X$  but Cimmino does not converge finitely to  $(0, 1)$  for this pair of sets, which means, again due to the relation between MAP and Cimmino, that MAP does not converge in a finite number of steps in the product space, preventing BAP-EB to hold there.

It is well-known that for the consistent case, Lipschitzian/Hölder error bound gives us in general linear/sublinear convergence of MAP [\[12, 21, 22, 40, 42\]](#). With that said, [Theorem 4.3](#) consists of a striking jump in the rate of convergence of MAP, namely from linear to finite. In this regard, the final example of this section incites the formulation of an intriguing question: Do we get linear convergence of MAP under infeasibility together with a Hölder error bound?

*Example 5.5* (A ball and a hyperplane in  $\mathbb{R}^2$ ). Consider the closed ball  $X$  in  $\mathbb{R}^2$  with radius 1 centered in  $(0, 1)$ , that is,  $X := \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + (x_2 - 1)^2 \leq 1\}$  and the family of horizontal hyperplanes of the form  $Y_\varepsilon := \{(t, -\varepsilon) \in \mathbb{R}^2 \mid t \in \mathbb{R}\}$ , where  $\varepsilon$  is a nonnegative parameter. Note that  $Y_0$  coincides with the optimal supporting hyperplane  $\mathbb{H}_{Y_\varepsilon}(X)$ , for all positive  $\varepsilon$ . We have that  $X \cap Y_0$  consists of the origin. If  $\varepsilon > 0$ ,  $X \cap Y_\varepsilon = \emptyset$  and there exists a unique best approximation pair, namely  $\{(0, 0), (0, -\varepsilon)\}$ .

We would like to draw our attention to a MAP iteration between  $X$  and  $Y_0$ , and one concerning  $X$  and  $Y_\varepsilon$ . These MAP iterations are given below and illustrated in [Figure 3](#). For a number  $a > 0$ , consider  $(a, 0) \in Y_0$  and  $(a, -\varepsilon) \in Y_\varepsilon$ . Note that

$$P_{Y_0}P_X(a, 0) = \left(\frac{a}{\sqrt{a^2 + 1}}, 0\right) \in Y_0$$

and

$$P_{Y_\varepsilon}P_X(a, -\varepsilon) = \left(\frac{a}{\sqrt{a^2 + (1 + \varepsilon)^2}}, -\varepsilon\right) \in Y_\varepsilon.$$

Note also that

$$(5.2) \quad \lim_{a \rightarrow 0^+} \frac{\|P_{Y_0}P_X(a, 0) - (0, 0)\|}{\|(a, 0) - (0, 0)\|} = \lim_{a \rightarrow 0^+} \frac{1}{\sqrt{a^2 + 1}} = 1,$$

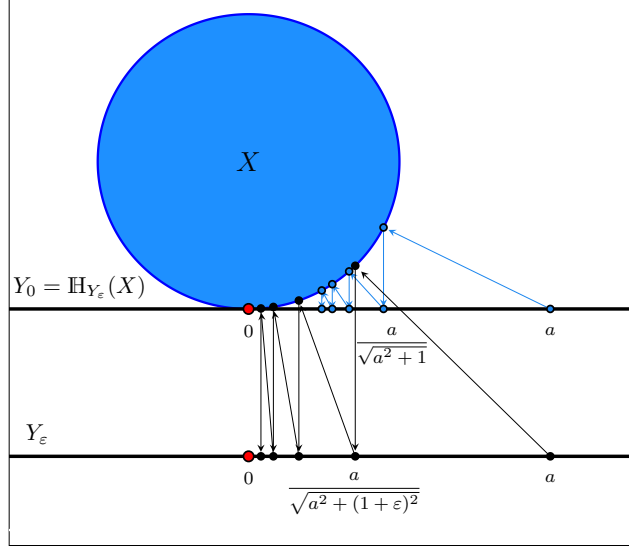


Fig. 3: Hölder error bound between  $X$  and  $\mathbb{H}_{Y_\epsilon}(X)$ .

and

$$(5.3) \quad \lim_{a \rightarrow 0^+} \frac{\|P_{Y_\epsilon} P_X(a, -\epsilon) - (0, -\epsilon)\|}{\|(a, -\epsilon) - (0, -\epsilon)\|} = \lim_{a \rightarrow 0^+} \frac{1}{\sqrt{a^2 + (1 + \epsilon)^2}} = \frac{1}{1 + \epsilon}.$$

The limit being 1 in (5.2) means that a MAP sequence between  $X$  and  $Y_0$ , monitored on  $Y_0$ , converges to the origin sublinearly. On the other hand, in view of (5.3), the gap of size  $\epsilon$  leads to a linear convergence of a MAP sequence to the best point  $(0, -\epsilon)$  with asymptotic rate  $\frac{1}{1+\epsilon}$ , where this MAP sequence regards  $X$  and  $Y_\epsilon$  and belongs to  $Y_\epsilon$ . For simplicity, we are looking at MAP shadows on  $Y_0/Y_\epsilon$ , but of course, due to the nonexpansiveness of projections, the rates would be sublinear/linear for the correspondent shadows on  $X$ .

Apparently, this leap from sublinear to linear convergence occurs in view of the presence of a Hölder error bound condition between  $X$  and  $\mathbb{H}_{Y_\epsilon}(X)$ . In fact, we have the existence of  $\omega > 0$  such that

$$(5.4) \quad \omega \operatorname{dist}(x, X \cap \mathbb{H}_{Y_\epsilon}(X))^{1+q} \leq \operatorname{dist}(x, \mathbb{H}_{Y_\epsilon}(X)),$$

$x \in X$  in a neighborhood of  $X \cap \mathbb{H}_{Y_\epsilon}(X) = \{(0, 0)\}$ , with  $q = 1$ . We remark that the Hölder parameter  $q = 1$  can be effortlessly derived in this example by taking into account that  $X$  is entirely contained in the epigraph of the quadratic function  $g(t) = \frac{t^2}{2}$ .

The previous example presented a particular instance where adding inconsistency led to a gain in the convergence rate of MAP from sublinear to linear. We understand this is not something one can expect without assuming some Hölder regularity. Probably, the improvement of MAP would be more limited if we considered MAP for a Cauchy bowl getting apart from a hyperplane. By Cauchy bowl [3] we mean something like the epigraph of the function  $f : D \rightarrow \mathbb{R}$  defined such that  $f(0) = 0$  and  $f(x) = e^{\|x\|^{-2}}$  elsewhere, where  $D = \{x \in \mathbb{R}^n \mid \|x\| \leq 1/\sqrt{3}\}$ . This function  $f$  is known to be infinitely differentiable in the interior of its domain, but not analytic. Furthermore, its epigraph does not satisfy a Hölder error bound with respect to the hyperplane  $H = \{(x, 0) \in \mathbb{R}^{n+1} \mid x \in \mathbb{R}^n\}$ .

We have just risen the conjecture of whether Hölder regularity, with  $q = 1$ , provides linear convergence for MAP in the inconsistent case. This looks very plausible and if indeed correct, it can establish MAP as a practical and (perhaps) competitive method for non-smooth optimization, at least for convex min-max problems featuring convex quadratic functions. Perhaps KL-theory [5, 46] may shed some light on the question.

*Example 5.6* (MAP for convex min-max optimization). Consider the following problem

$$(5.5) \quad \min_{x \in \mathbb{R}^n} g(x) := \max \{f_1(x), \dots, f_m(x)\},$$

where each  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$  is a convex function and assume that  $g$  has a minimizer. Let  $g^*$  be the optimal value of (5.5), take  $\beta < g^*$  and define the hyperplane  $H_\beta := \{(x, \beta) \in \mathbb{R}^{n+1} \mid x \in \mathbb{R}^n\}$ . Thus, the epigraph of  $g$ ,  $\text{epi } g := \{(x, t) \in \mathbb{R}^{n+1} \mid x \in \mathbb{R}^n, g(x) \leq t\}$ , does not intersect  $H_\beta$  and  $\text{bap}_{H_\beta}(\text{epi } g)$  consists of the minimizers of  $g$ . Moreover, due to Fact 2.3(i), for any starting point  $(x^0, t_0) \in \mathbb{R}^{n+1}$  the MAP sequence defined by

$$(5.6) \quad (x^{k+1}, t_{k+1}) := P_{\text{epi } g} P_{H_\beta}((x^k, t_k))$$

converges to a point  $(x^*, g^*)$  such that  $x^*$  is a solution of (5.5).

Note that if all  $f_i$  are affine, solving (5.5) relates to linear programming, in the sense discussed in Example 5.1, and finite convergence of the MAP sequence (5.6) is assured by Theorem 4.1.

Now, Theorem 4.2 opens the possibility for the MAP sequence (5.6) to achieve finite convergence in a more general setting. If, for instance, the solution  $x^*$  of (5.5) is unique and if the tangent cone of  $\text{epi } g$  at  $(x^*, g^*)$  is pointed, the Lipschitzian BAP-EB (3.4) holds and the MAP sequence (5.6) converges to  $(x^*, g^*)$  after a finite number of steps.

Of course that addressing a non-affine setting in practice relies on the calculation of projections onto epigraphs and a lower bound  $\beta$  for  $g$  in hand. Getting  $\beta$  is potentially an easier task. For instance, if there is an index  $i$  for which  $f_i$  has a minimizer  $\hat{x}$ , we have that any  $\beta < f_i(\hat{x})$  provides a strict lower bound for  $g$  as  $f_i(\hat{x}) \leq g^*$ .

With that said, one can think of using MAP for quadratic min-max, that is, when all  $f_i$ 's are convex quadratic functions. In this case, finding  $\beta$  is easy and projecting onto the epigraph of  $g$  is somehow manageable [32]. As for the BAP-EB (3.4), it may or may not be satisfied. However, the Hölder condition (5.4) holds with  $q = 1$ , in view of the quadratic growth of  $g$ . Hence, if the conjecture raised in Example 5.5 is true for  $q = 1$ , we would get either finite or linear convergence of the MAP sequence (5.6) for convex quadratic min-max problems.

Furthermore, the conjecture being true would echo as well in min-max problems featuring strongly convex functions. The maximum of a finite number of strongly convex functions is strongly convex, has a unique minimizer and one can bound  $g$  below by a strictly convex quadratic function with minimum value  $g^*$ . So, Hölder regularity holds with  $q = 1$  in the strongly convex setting of (5.5) and again, one would either get finite or linear convergence of the MAP sequence (5.6).

The bonds between min-max problems and Hölder regularity together with the validity of the conjecture formulated in Example 5.5 would be an asset in the field of non-smooth convex optimization.

**6. Concluding remarks.** We have derived finite convergence of alternating projections for two non-intersecting closed convex sets satisfying a Lipschitzian error bound condition, which has been to proven be more general than the well known concept of intrinsic transversality. This result strengthens the theory on MAP, a widely acclaimed method in Mathematics. In addition to being interesting from a theoretical point of view, our main theorem may also have an impact on practical issues regarding projection-type algorithms in general, as inconsistency has been seen favorable for MAP. A question left open is to what extent MAP can improve when embedding inconsistency to

a problem satisfying a Hölder error bound condition. As discussed within an example, a positive answer in this direction may turn MAP into an attractive method for minimizing the maximum of convex functions, a central problem in non-smooth optimization. Although our main results are for finite convergence of MAP, we have as well stated theorems on other projection/reflection-based methods for both two-set and multi-set inconsistent feasibility problems. In particular, the Cimmino method, the Douglas-Rachford method and Cyclic projections have been addressed.

**Acknowledgments.** The authors would like to thank the associate editor and the anonymous referees for their valuable suggestions, which significantly improved this manuscript.

## REFERENCES

- [1] R. AHARONI, Y. CENSOR, AND Z. JIANG, *Finding a best approximation pair of points for two polyhedra*, Comput Optim Appl, 71 (2018), pp. 509–523, <https://doi.org/10.1007/s10589-018-0021-3>.
- [2] S. ALWADANI, H. H. BAUSCHKE, J. P. REVALSKI, AND X. WANG, *The difference vectors for convex sets and a resolution of the geometry conjecture*, arXiv:2012.04784 [math], (2020), <https://arxiv.org/abs/2012.04784>.
- [3] R. AREFIDAMGHANI, R. BEHLING, J.-Y. BELLO-CRUZ, A. N. IUSEM, AND L.-R. SANTOS, *The circumcentered-reflection method achieves better rates than alternating projections*, Comput Optim Appl, 79 (2021), pp. 507–530, <https://doi.org/10.1007/s10589-021-00275-6>, <https://arxiv.org/abs/2007.14466>.
- [4] N. ARONSZAJN, *Theory of Reproducing Kernels*, Transactions of the American Mathematical Society, 68 (1950), pp. 337–404, <https://doi.org/10.1090/S0002-9947-1950-0051437-7>.
- [5] H. ATTOUCH, J. BOLTE, P. REDONT, AND A. SOUBEYRAN, *Proximal Alternating Minimization and Projection Methods for Nonconvex Problems: An Approach Based on the Kurdyka-Łojasiewicz Inequality*, Mathematics of OR, 35 (2010), pp. 438–457, <https://doi.org/10.1287/moor.1100.0449>.
- [6] C. BADEA, S. GRIVAUX, AND V. MÜLLER, *The rate of convergence in the method of alternating projections*, St. Petersburg Mathematical Journal, 23 (2012), pp. 413–434, <https://doi.org/10.1090/S1061-0022-2012-01202-1>.
- [7] J.-B. BAILLON, P. COMBETTES, AND R. COMINETTI, *There is no variational characterization of the cycles in the method of periodic projections*, Journal of Functional Analysis, 262 (2012), pp. 400–408, <https://doi.org/10.1016/j.jfa.2011.09.002>.
- [8] H. BAUSCHKE, *A norm convergence result on random products of relaxed projections in Hilbert space*, Trans. Amer. Math. Soc., 347 (1995), pp. 1365–1373, <https://doi.org/10.1090/S0002-9947-1995-1257097-1>.
- [9] H. H. BAUSCHKE, J.-Y. BELLO-CRUZ, T. T. A. NGHIA, H. M. PHAN, AND X. WANG, *Optimal Rates of Linear Convergence of Relaxed Alternating Projections and Generalized Douglas-Rachford Methods for Two Subspaces*, Numer. Algorithms, 73 (2016), pp. 33–76, <https://doi.org/10.1007/s11075-015-0085-4>.
- [10] H. H. BAUSCHKE AND J. M. BORWEIN, *On the convergence of von Neumann’s alternating projection algorithm for two sets*, Set-Valued Analysis, 1 (1993), pp. 185–212, <https://doi.org/10.1007/BF01027691>.
- [11] H. H. BAUSCHKE AND J. M. BORWEIN, *Dykstra’s Alternating Projection Algorithm for Two Sets*, J. Approx. Theory, 79 (1994), pp. 418–443, <https://doi.org/10.1006/jath.1994.1136>.
- [12] H. H. BAUSCHKE AND J. M. BORWEIN, *On Projection Algorithms for Solving Convex Feasibility Problems*, SIAM Review, 38 (1996), pp. 367–426, <https://doi.org/10.1137/S0036144593251710>.
- [13] H. H. BAUSCHKE, J. M. BORWEIN, AND A. S. LEWIS, *The method of cyclic projections for closed convex sets in Hilbert space*, Contemporary Math., 204 (1997), pp. 1–38.
- [14] H. H. BAUSCHKE AND P. L. COMBETTES, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics, Springer International Publishing, second ed., 2017, <https://doi.org/10.1007/978-3-319-48311-5>.
- [15] H. H. BAUSCHKE, P. L. COMBETTES, AND D. R. LUKE, *Finding best approximation pairs relative to two closed convex sets in Hilbert spaces*, Journal of Approximation Theory, 127 (2004), pp. 178–192, <https://doi.org/10.1016/j.jat.2004.02.006>.
- [16] H. H. BAUSCHKE, F. DEUTSCH, AND H. HUNDAL, *Characterizing arbitrarily slow convergence in the method of alternating projections*, International Transactions in Operational Research, 16 (2009), pp. 413–425, <https://doi.org/10.1111/j.1475-3995.2008.00682.x>.
- [17] R. BEHLING, J.-Y. BELLO-CRUZ, AND L.-R. SANTOS, *The block-wise circumcentered-reflection method*, Comput Optim Appl, 76 (2020), pp. 675–699, <https://doi.org/10.1007/s10589-019-00155-0>.
- [18] R. BEHLING, D. S. GONÇALVES, AND S. A. SANTOS, *Local Convergence Analysis of the Levenberg–Marquardt Framework for Nonzero-Residue Nonlinear Least-Squares Problems Under an Error Bound Condition*, J Optim Theory Appl, 183 (2019), pp. 1099–1122, <https://doi.org/10.1007/s10957-019-01586-9>.
- [19] R. BEHLING AND A. IUSEM, *The effect of calmness on the solution set of systems of nonlinear equations*, Math. Program., 137 (2013), pp. 155–165, <https://doi.org/10.1007/s10107-011-0486-7>.

- [20] D. P. BERTSEKAS, *Nonlinear Programming*, Athena Scientific, Belmont, USA, second ed., 1999.
- [21] J. M. BORWEIN, G. LI, AND M. K. TAM, *Convergence Rate Analysis for Averaged Fixed Point Iterations in Common Fixed Point Problems*, SIAM J. Optim., 27 (2017), pp. 1–33, <https://doi.org/10.1137/15M1045223>.
- [22] J. M. BORWEIN, G. LI, AND L. YAO, *Analysis of the Convergence Rate for the Cyclic Projection Algorithm Applied to Basic Semialgebraic Convex Sets*, SIAM J. Optim., 24 (2014), pp. 498–527, <https://doi.org/10.1137/130919052>.
- [23] H. T. BUI, R. LOXTON, AND A. MOEINI, *A note on the finite convergence of alternating projections*, Operations Research Letters, 49 (2021), pp. 431–438, <https://doi.org/10.1016/j.orl.2021.04.009>.
- [24] Y. CENSOR, *On variable block algebraic reconstruction techniques*, in Mathematical Methods in Tomography, G. T. Herman, A. K. Louis, and F. Natterer, eds., vol. 1497 of Lecture Notes in Mathematics, Springer, Berlin, Heidelberg, 1991, pp. 133–140, <https://doi.org/10.1007/BFb0084514>.
- [25] Y. CENSOR, M. D. ALTSCHULER, AND W. D. POWLIS, *On the use of Cimmino’s simultaneous projections method for computing a solution of the inverse problem in radiation therapy treatment planning*, Inverse Problems, 4 (1988), pp. 607–623, <https://doi.org/10.1088/0266-5611/4/3/006>.
- [26] Y. CENSOR AND A. CEGIELSKI, *Projection methods: An annotated bibliography of books and reviews*, Optimization, 64 (2014), pp. 2343–2358, <https://doi.org/10.1080/02331934.2014.957701>.
- [27] Y. CENSOR AND M. ZAKNOON, *Algorithms and Convergence Results of Projection Methods for Inconsistent Feasibility Problems: A Review*, Pure and Applied Functional Analysis, 3 (2018), pp. 565–586.
- [28] W. CHENEY AND A. A. GOLDSTEIN, *Proximity Maps for Convex Sets*, Proceedings of the American Mathematical Society, 10 (1959), pp. 448–450, <https://doi.org/10.2307/2032864>.
- [29] G. CIMMINO, *Calcolo approssimato per le soluzioni dei sistemi di equazioni lineari*, La Ricerca Scientifica, 9 (1938), pp. 326–333.
- [30] E. S. COAKLEY, V. ROKHLIN, AND M. TYGERT, *A Fast Randomized Algorithm for Orthogonal Projection*, SIAM Journal on Scientific Computing, 33 (2011), pp. 849–868, <https://doi.org/10.1137/090779656>.
- [31] P. L. COMBETTES, *Inconsistent signal feasibility problems: Least-squares solutions in a product space*, IEEE Trans. Signal Process., 42 (1994), pp. 2955–2966, <https://doi.org/10.1109/78.330356>.
- [32] Y.-H. DAI, *Fast Algorithms for Projection on an Ellipsoid*, SIAM J. Optim., 16 (2006), pp. 986–1006, <https://doi.org/10.1137/040613305>.
- [33] A. R. DE PIERRO AND A. N. IUSEM, *A finitely convergent “row-action” method for the convex feasibility problem*, Appl Math Optim, 17 (1988), pp. 225–235, <https://doi.org/10.1007/BF01448368>.
- [34] F. DEUTSCH, *The Method of Alternating Orthogonal Projections*, in Approximation Theory, Spline Functions and Applications, S. P. Singh, ed., no. 356 in NATO ASI Series, Springer, Dordrecht, 1992, pp. 105–121, [https://doi.org/10.1007/978-94-011-2634-2\\_5](https://doi.org/10.1007/978-94-011-2634-2_5).
- [35] F. R. DEUTSCH, *The Angle Between Subspaces of a Hilbert Space*, in Approximation Theory, Wavelets and Applications, S. P. Singh, ed., vol. 454 of NATO Science Series, Springer, Dordrecht, 1995, pp. 107–130, [https://doi.org/10.1007/978-94-015-8577-4\\_7](https://doi.org/10.1007/978-94-015-8577-4_7).
- [36] F. R. DEUTSCH AND H. HUNDAL, *The Rate of Convergence for the Method of Alternating Projections, II*, J. Math. Anal. Appl., 205 (1997), pp. 381–405, <https://doi.org/10.1006/jmaa.1997.5202>.
- [37] P. DIACONIS, K. KHARE, AND L. SALOFF-COSTE, *Stochastic alternating projections*, Illinois J. Math., 54 (2010), pp. 963–979, <https://doi.org/10.1215/ijm/1336568522>.
- [38] J. DOUGLAS AND H. H. RACHFORD JR., *On the numerical solution of heat conduction problems in two and three space variables*, Transactions of the American Mathematical Society, 82 (1956), pp. 421–421, <https://doi.org/10.1090/S0002-9947-1956-0084194-4>.
- [39] D. DRUSVYATSKIY, A. D. IOFFE, AND A. S. LEWIS, *Transversality and Alternating Projections for Nonconvex Sets*, Found Comput Math, 15 (2015), pp. 1637–1651, <https://doi.org/10.1007/s10208-015-9279-3>.
- [40] D. DRUSVYATSKIY, G. LI, AND H. WOLKOWICZ, *A note on alternating projections for ill-posed semidefinite feasibility problems*, Mathematical Programming, 162 (2016), pp. 537–548, <https://doi.org/10.1007/s10107-016-1048-9>.
- [41] C. FRANCHETTI AND W. LIGHT, *On the von Neumann alternating algorithm in Hilbert space*, J. Math. Anal. Appl., 114 (1986), pp. 305–314, [https://doi.org/10.1016/0022-247X\(86\)90085-5](https://doi.org/10.1016/0022-247X(86)90085-5).
- [42] L. G. GUBIN, B. T. POLYAK, AND E. V. RAIK, *The method of projections for finding the common point of convex sets*, USSR Computational Mathematics and Mathematical Physics, 7 (1967), pp. 1–24, [https://doi.org/10.1016/0041-5553\(67\)90113-9](https://doi.org/10.1016/0041-5553(67)90113-9).
- [43] S. KACZMARZ, *Angenäherte Auflösung von Systemen linearer Gleichungen*, Bull. Int. Acad. Pol. Sci. Lett. Class. Sci. Math. Nat. A, 35 (1937), pp. 355–357.
- [44] S. KAYALAR AND H. L. WEINERT, *Error bounds for the method of alternating projections*, Math. Control Signal Systems, 1 (1988), pp. 43–59, <https://doi.org/10.1007/BF02551235>.
- [45] A. Y. KRUGER, *About Intrinsic Transversality of Pairs of Sets*, Set-Valued Var. Anal, 26 (2018), pp. 111–142, <https://doi.org/10.1007/s11228-017-0446-3>.
- [46] G. LI AND T. K. PONG, *Calculus of the Exponent of Kurdyka–Łojasiewicz Inequality and Its Applications to Linear Convergence of First-Order Methods*, Found Comput Math, 18 (2018), pp. 1199–1232, <https://doi.org/10.1007/s10208-017-9366-8>.

- [47] E. A. NURMINSKI, *Single-projection procedure for linear optimization*, J Glob Optim, 66 (2016), pp. 95–110, <https://doi.org/10.1007/s10898-015-0337-9>.
- [48] G. PIERRA, *Decomposition through formalization in a product space*, Mathematical Programming, 28 (1984), pp. 96–115, <https://doi.org/10.1007/BF02612715>.
- [49] R. T. ROCKAFELLAR AND R. J.-B. WETS, *Variational Analysis*, no. 317 in Grundlehren Der Mathematischen Wissenschaften, Springer, Berlin, second ed., 2004.
- [50] K. T. SMITH, D. C. SOLMON, AND S. L. WAGNER, *Practical and mathematical aspects of the problem of reconstructing objects from radiographs*, Bull. Amer. Math. Soc., 83 (1977), pp. 1227–1271, <https://doi.org/10.1090/S0002-9904-1977-14406-6>.
- [51] J. VON NEUMANN, *Functional Operators, Volume 2: The Geometry of Orthogonal Spaces.*, no. 22 in Annals of Mathematics Studies, Princeton University Press, Princeton, 1950, <https://doi.org/10.2307/j.ctt1bc543b>.