

THE BEST APPROXIMATION TUPLE: AN EXTENSION OF THE CHENEY-GOLDSTEIN ALGORITHM AND RESULTS TO THE MULTIPLE SETS CASE

YAIR CENSOR AND DANIEL REEM

ABSTRACT. We extend the algorithm and several results published in the celebrated 1959 paper of Cheney and Goldstein about the best approximation pair (BAP) problem in two separate directions. One is, for the first time, the ability to consider more than two sets, a task which has been in the mind of researchers for many years without much success in fulfilling it because of a certain theoretical obstacle related to cycles and minimizers of general functionals. The other direction is the ability to handle each set as an intersection of a finite family of sets. We call the resulting problem the “Best Approximation Tuple (BAT) problem”. The fundamental observation that leads to these extensions is to recognize and handle one set (the “pivot set”) as different from the remaining sets (the “satellite sets”) instead of seeking cycles as the minimizers of a target functional. We prove the convergence of the algorithm to the unique solution of the problem in the Euclidean case with strictly convex and compact satellite sets. Because of the lack of Fejér monotonicity, our convergence analysis is not standard, and is based on almost unknown properties of orthogonal projections regarding equality and inequality in the definition of nonexpansiveness.

1. INTRODUCTION

1.1. Background and contributions. In the classical best approximation pair (BAP) problem we are given two nonempty, closed, convex and disjoint subsets Ω_1 and Ω_2 located in a real Hilbert space X , and we want to find a pair of points, each from each subset, which realizes the distance between the subsets. In other words, we want to find $a \in \Omega_1$ and $b \in \Omega_2$ such that

$$\|a - b\| = \text{dist}(\Omega_1, \Omega_2) := \inf\{\|x - y\| \mid x \in \Omega_1, y \in \Omega_2\}, \quad (1.1)$$

namely, (a, b) is a minimizer over $\Omega_1 \times \Omega_2$ of the function $f : X^2 \rightarrow [0, \infty)$ defined by $f(x, y) := \|x - y\|$ for all $(x, y) \in X^2$. The BAP problem goes back to the classical 1959 work of Cheney and Goldstein [28] and, briefly, also to the paper [67] of Nicolescu from 1938. Numerous papers discussing this problem in various settings have been published, and the problem has found applications in signal processing, solid modeling, computer graphics, robotics, collision detection, computer aided design, virtual reality and more. For a very partial list of works see, for instance, [1, 5, 6, 10, 11, 16, 17, 22, 27, 40, 42, 44,

Date: October 8, 2026.

2020 Mathematics Subject Classification. 90C25, 90C30, 47H10, 46N10, 47N10, 41A50.

Key words and phrases. Best Approximation Pair (BAP) problem, Best Approximation Tuple (BAT) problem, Cheney-Goldstein theorem, Fixed point, orthogonal projection, proximity function, strictly convex set.

[46](#), [47](#), [55](#), [57](#), [58](#), [60–62](#), [66](#), [68–70](#), [73–78](#), [80](#)]. For a review about the BAP problem (with over 120 citations) in various settings, as well as for many relevant results and applications, see our recent work [\[71\]](#).

It is not immediately clear how to generalize the BAP problem to more than two sets, and this task has been in the mind of researchers for many years without much success in fulfilling it, and one can find statements about the “interesting challenges” that poses the task of extending results related to the BAP problem (and, implicitly, the BAP problem itself) to more than two nonempty, closed and convex sets [\[4, p. 401\]](#). In fact, it has turned out there is a theoretical obstacle for a natural generalization. Indeed, we first recall the known fact, which is implicit in [\[28\]](#) (and follows from [\[28, Theorem 2\]](#) and additional simple arguments; see also Lemma [4.6](#) below for a generalization), that $(a, b) \in \Omega_1 \times \Omega_2$ is a BAP relative to (Ω_1, Ω_2) if and only if (a, b) is a 2-cycle, namely $a = P_1(b)$ and $b = P_2(a)$. Thus, given $m \in \mathbb{N}$, a natural candidate for being a generalization of the concept of a BAP to any $m + 1 \geq 2$ sets $\Omega_1, \dots, \Omega_{m+1}$ is an $(m + 1)$ -cycle relative to $(\Omega_1, \Omega_2, \dots, \Omega_{m+1})$, namely a tuple $(a^1, a^2, \dots, a^{m+1}) \in \Omega_1 \times \Omega_2 \times \dots \times \Omega_{m+1}$ which satisfies $a^i = P_i(a^{i+1})$ for all $i \in \{1, 2, \dots, m\}$ and $a^{m+1} = P_{m+1}(a^1)$.

While there are various relevant results related to cycles, such as [\[8, Section 5\]](#), [\[50, Theorem 2\]](#), [\[76, Theorem 1.6, Corollary 1.7\]](#), [\[12, Corollary 4.10, Corollary 4.11\]](#) and [\[2, Theorem 9\]](#), none of them shows that when $m + 1 \geq 3$, then there is a certain functional such that the minimizers of this functional over $\Omega_1 \times \Omega_2 \times \dots \times \Omega_{m+1}$ (for all nonempty, closed and convex sets $\Omega_i \subseteq X$, $i \in \{1, 2, \dots, m + 1\}$) must be $(m + 1)$ -cycles relative to $(\Omega_1, \Omega_2, \dots, \Omega_{m+1})$ and vice versa, as in the case of two sets mentioned above (simple candidates to such a functional, such as $(x^i)_{i=1}^{m+1} \mapsto \|x^1 - x^2\| + \|x^2 - x^3\| + \dots + \|x^m - x^{m+1}\| + \|x^{m+1} - x^1\|$, have been considered and rejected: see [\[8, Remarks 5.1.5\]](#) and [\[59, p. 36\]](#) where $m + 1 = 3$).

Apparently this is not a coincidence, since a remarkable result of Baillon, Combettes and Cominetti [\[4, Theorem 2.3\]](#) shows that when X has dimension 2 or above (possibly infinite dimension) and $m + 1 \geq 3$, then there exists no functional $\Phi : X^{m+1} \rightarrow \mathbb{R}$ having the property that for all tuple $(\Omega_1, \Omega_2, \dots, \Omega_{m+1})$ of nonempty, closed and convex sets $\Omega_i \subseteq X$, $i \in \{1, 2, \dots, m + 1\}$, the set of $(m + 1)$ -cycles relative to $(\Omega_1, \Omega_2, \dots, \Omega_{m+1})$ is the set of minimizers of Φ over $\Omega_1 \times \Omega_2 \times \dots \times \Omega_{m+1}$. This theoretical obstacle has led to statements such as “the Cheney and Goldstein...result of minimizing the distance between two disjoint sets cannot be extended to more than two sets” [\[24, p. 579, after Theorem 2.18\]](#) and “The problem of best approximation pair relative to two sets cannot be extended to more than two sets in view of Baillon, Combettes and Cominetti’s result” [\[24, p. 581, Subsection 2.16\]](#).

We present here a way to bypass this theoretical obstacle by singling out and focusing our attention on one specific component of the tuple of sets which we call the “pivot set”. Then we pose the problem of minimizing a natural functional, which is defined on X and not on X^{m+1} , over the pivot set.

More precisely, assume for the sake of simplicity that we single out and focus on the $(m + 1)$ -th index for the pivot set. If $(\beta_i)_{i=1}^m$ are positive numbers satisfying $\sum_{i=1}^m \beta_i = 1$, and if P_i denotes the orthogonal projection onto any remaining sets Ω_i , $i \in \{1, 2, \dots, m\}$ (which we call the “satellite sets”), then we consider the following minimization problem,

wherein $d(x, \Omega_i) := \inf\{\|x - y\| \mid y \in \Omega_i\}$ is the distance between the point x and the set Ω_i :

$$\min_{x \in \Omega_{m+1}} \frac{1}{2} \sum_{i=1}^m \beta_i d(x, \Omega_i)^2 = \min_{x \in \Omega_{m+1}} \frac{1}{2} \sum_{i=1}^m \beta_i \|x - P_i(x)\|^2. \quad (1.2)$$

Once we have a solution $x^* \in \Omega_{m+1}$ to **(1.2)** at our hands, we consider the tuple $(P_1(x^*), P_2(x^*), \dots, P_m(x^*), x^*)$ as the Best Approximation Tuple (BAT). This tuple is not necessarily an $(m+1)$ -cycle when $m+1 > 2$, but the case $m+1 = 2$ reduces to the BAP problem considered by Cheney and Goldstein [28].

Indeed, in this case β_1 must be 1, and $x^* \in \Omega_2$ solves **(1.2)** if and only if x^* solves the minimization problem $\min_{x \in \Omega_2} \frac{1}{2} d(x, \Omega_1)^2$, i.e., if and only if x^* is a minimizer of the minimization problem $\min_{x \in \Omega_2} d(x, \Omega_1)$. Since [28, Theorem 2] ensures that x^* solves the minimization problem $\min_{x \in \Omega_2} d(x, \Omega_1)$ if and only if $x^* = P_2(P_1(x^*))$, by recalling what we explained a bit after **(1.1)**, namely that $(a, b) \in \Omega_1 \times \Omega_2$ is a BAP relative to (Ω_1, Ω_2) if and only if (a, b) is a 2-cycle, and by letting $a := P_1(x^*)$ and $b := x^*$, we conclude that $x^* \in \Omega_2$ solves **(1.2)** if and only if $(P_1(x^*), x^*)$ is a BAP and a BAT relative to (Ω_1, Ω_2) .

Proposition 3.7 below shows that a tuple $(y^1, y^2, \dots, y^{m+1}) \in \Omega_1 \times \Omega_2 \times \dots \times \Omega_{m+1}$ is a BAT if and only if it solves the multivariate minimization problem

$$\min_{(x^1, x^2, \dots, x^{m+1}) \in \Omega_1 \times \Omega_2 \times \dots \times \Omega_{m+1}} \frac{1}{2} \sum_{i=1}^m \beta_i \|x^{m+1} - x^i\|^2. \quad (1.3)$$

Thus, in contrast with the case forbidden in the Baillon, Combettes and Cominetti result [4, Theorem 2.3], there is a (non-symmetric) multivariate functional over X^{m+1} , that is $(x^1, x^2, \dots, x^{m+1}) \mapsto 0.5 \sum_{i=1}^m \beta_i \|x^{m+1} - x^i\|^2$, $(x^1, x^2, \dots, x^{m+1}) \in X^{m+1}$, whose set of minimizers over $\Omega_1 \times \Omega_2 \times \dots \times \Omega_{m+1}$ (for every tuple $(\Omega_1, \Omega_2, \dots, \Omega_{m+1})$ of nonempty, closed and convex subsets $\Omega_i \subseteq X$, $i \in \{1, 2, \dots, m+1\}$) coincides with the set of “generalized best approximation pairs” (namely BATs). Again, a BAT is not necessarily a cycle, and so there is no contradiction with [4, Theorem 2.3].

Our generalization **(1.2)** of the BAP problem was inspired by the, so-called, “inconsistent convex feasibility problem with soft and hard constraints” and by the generalized Heron problem. Before presenting the former problem, recall first the classical convex feasibility problem (CFP): we are given a family $(\Omega_i)_{i \in I}$ of nonempty, closed and convex subsets in a real Hilbert space X , (where $I \neq \emptyset$ and it is finite or infinite) and our goal is to find a feasible point, that is, a point which belongs to the intersection $\cap_{i \in I} \Omega_i$. See, for example, [7, 14, 15, 18, 20, 23, 26, 31, 50] for related results, algorithms and applications.

There are, however, various real-world and theoretical instances in which either it is not clear in advance that $\cap_{i \in I} \Omega_i \neq \emptyset$, or it is known that $\cap_{i \in I} \Omega_i = \emptyset$ and yet one wants to find a point which somehow generalizes the notion of a feasible point. A typical example for this is the inconsistent convex feasibility problem with soft and hard constraints [35]. In this problem the sets Ω_i , $i \in I$ are divided into two groups: one group is the hard constraints group Ω_i , $i \in I_{\text{hard}} \neq \emptyset$, and the second group is the soft constraints group Ω_i , $i \in I_{\text{soft}} \neq \emptyset$, where $I_{\text{hard}} \cup I_{\text{soft}} = I$, $I_{\text{hard}} \cap I_{\text{soft}} = \emptyset$, and I has at least two indices. The goal is to find a point $x^* \in \cap_{i \in I_{\text{hard}}} \Omega_i$ (assumed or known to be nonempty) which violates

as little as possible the constraints Ω_i , $i \in I_{\text{soft}}$. More precisely, we want to find a point $x^* \in \cap_{i \in I_{\text{hard}}} \Omega_i$ which is a minimizer, over $\cap_{i \in I_{\text{hard}}} \Omega_i$, of the functional $\Phi_{\text{soft}} : X \rightarrow \mathbb{R}$ defined by

$$\Phi_{\text{soft}}(x) := \frac{1}{2} \sum_{i \in I_{\text{soft}}} \beta_i d(x, \Omega_i)^2 = \frac{1}{2} \sum_{i \in I_{\text{soft}}} \beta_i \|x - P_i(x)\|^2, \quad \forall x \in X, \quad (1.4)$$

where, for each $i \in I_{\text{soft}}$, $P_i : X \rightarrow X$ is the orthogonal projection onto Ω_i , $\beta_i > 0$ and $\sum_{i \in I_{\text{soft}}} \beta_i = 1$.

The functional Φ_{soft} is a proximity function (see, e.g., [18, p. 28]). It has the property that $\Phi_{\text{soft}}(x) = 0$ if and only if $x \in \cap_{i \in I_{\text{soft}}} \Omega_i$, and $\Phi_{\text{soft}}(x) > 0$ if and only if $x \notin \cap_{i \in I_{\text{soft}}} \Omega_i$ for at least one index $i \in I_{\text{soft}}$. One can think about Φ_{soft} as a cost function which assigns to each $x \in X$ the price tag $\Phi_{\text{soft}}(x)$ which represents by how much x violates the constraints: the higher the violation, the higher the price, and we want to find a point $x^* \in \cap_{i \in I_{\text{hard}}} \Omega_i$ having the least price tag. Alternatively, Φ_{soft} can be thought of as being an energy function which assigns to each $x \in X$ the energy level $\Phi_{\text{soft}}(x)$, and our goal is to find a point $x^* \in \cap_{i \in I_{\text{hard}}} \Omega_i$ having the lowest possible energy level.

The inconsistent convex feasibility problem with soft and hard constraints has applications in signal processing [35, 49, 79], network bandwidth allocation problem [53], power control [52], network analysis [54], sound velocity reconstruction [51] and more; see also [3, 18, 19, 21, 25, 30, 32–34, 36–39] for various algorithms and other results related to this problem, and see [24] for a general review related to the inconsistent convex feasibility problem.

Back to (1.2), one can see that it is a particular case of inconsistent convex feasibility problem with soft and hard constraints where $I_{\text{hard}} = \{m+1\}$ and $I_{\text{soft}} = \{1, 2, \dots, m\}$. Of course, we could let $I_{\text{hard}} := \{h\}$ and $I_{\text{soft}} := I \setminus \{h\}$ for any $h \in \{1, 2, \dots, m+1\}$, and by doing this, we would have obtained a different functional to be minimized and a different generalization of the BAP problem.

The second problem which has inspired us is the generalized Heron problem [29, 63, 64] (see also [79]). Its original version (the classical Heron problem) goes back to antiquity, and it has applications in optics and engineering. Here instead of minimizing the quadratic-type functional (1.2) one minimizes, over the pivot set, a non-weighted sum of the distances between a point and the satellite sets:

$$\min_{x \in \Omega_{m+1}} \sum_{i=1}^m d(x, \Omega_i). \quad (1.5)$$

In addition to extending the BAP problem to any finite number of sets (greater than or equal to 2), we also extend the problem in another aspect, by considering sets Ω_i having the property that each of them is an intersection of finitely many sets, namely $\Omega_i = \cap_{j=1}^{m_i} \Omega_{i,j}$ for some sets $\Omega_{i,j}$, where $i \in \{1, 2, \dots, m\}$, $m_i \in \mathbb{N}$ and $j \in \{1, 2, \dots, m_i\}$. Such a scenario appears, for instance, in bandwidth allocation problems [53, Subsections 2.1 and 2.2] and in reconstruction of sound velocity distribution [51, Sections III, IV]. See also [35, Fig. 1] and [31, Subsection 6.1].

Our extension, mentioned above, is what we call the Best Approximation Tuple (BAT) problem. Inspired by the “Alternating Simultaneous Halpern–Lions–Wittmann–Bauschke (A-S-HLWB)” algorithm that we presented in [22], and by other material which appears there, we formulate precisely the BAT problem (Problem 3.1 below), formulate conditions which ensure the existence and uniqueness of a solution to this problem, present an algorithmic scheme for solving the problem (Algorithm 3.3 below), and prove its convergence under some reasonable assumptions (Theorem 3.5 below; see also Remark 5.5 for an extension), among them that the space is Euclidean and the satellite sets are strictly convex and compact.

Algorithm 3.3 below has the advantage that instead of projecting onto the individual sets Ω_i , $i \in \{1, 2, \dots, m\}$, a task which may be demanding from the computational point of view, one projects onto the building bricks $\Omega_{i,j}$, $i \in \{1, 2, \dots, m\}$, $j \in \{1, 2, \dots, m_i\}$ which induce the intersection set Ω_i , $i \in \{1, 2, \dots, m\}$, a task which is less expensive, particularly when these building bricks have a simple form.

Because of the lack of Fejér monotonicity, our convergence analysis is not standard, and is based on almost unknown properties of orthogonal projections regarding equality and inequality in the definition of nonexpansiveness: see Lemma 4.1, Remark 4.2, Lemma 4.3 and Lemma 4.4 below. These properties allow us to derive strict estimates at crucial stages (such as the inequality (5.4) and the inequality $r_\varepsilon + \|P_i(x^k) - P_i(x^*)\| < \varepsilon$ for all $i \in \{1, 2, \dots, m\}$ mentioned below (5.5)).

1.2. Paper layout: In Section 2 we present necessary terminology and notation needed for the rest of the paper. In Section 3 we present precisely the BAT problem, the algorithm for solving it, the main existence, uniqueness and convergence theorem (Theorem 3.5), a characterization property of BATs as minimizers of a multivariate target functional (Proposition 3.7), and we also compare briefly our algorithm and results to the ones presented by Cheney and Goldstein [28]. In Section 4 we formulate and prove various auxiliary assertions, and in Section 5 we present the proof of Theorem 3.5.

2. PRELIMINARIES

Our setting is a real Hilbert space $(X, \|\cdot\|)$ with an inner product $\langle \cdot, \cdot \rangle$ and $X \neq \{0\}$, although some of the definitions below hold in a normed space or a metric space setting, and from time to time we will also consider these and other (perhaps more general) settings.

Given a nonempty subset $S \subseteq X$ and an operator $F : S \rightarrow X$, we say that F is nonexpansive on S if $\|F(x) - F(y)\| \leq \|x - y\|$ for all $(x, y) \in S^2$. We say that F is contractive on S in the sense of Edelstein, or Edelstein-contractive, or simply contractive (following Edelstein [43, p. 74]), if $\|F(x) - F(y)\| < \|x - y\|$ whenever $x \neq y$ and $(x, y) \in S^2$. Note that in [9, p. 69, Definition 4.1(iii)] a different term is used for describing Edelstein-contractive operators: “strictly nonexpansive” (we chose to not use this term because of a reason which is explained in Remark 4.2 below). An Edelstein-contractive mapping on S is nonexpansive on S since given $(x, y) \in S^2$, obviously $\|F(x) - F(y)\| \leq \|x - y\|$ whenever $x = y$ or $x \neq y$.

Given $\nu > 0$, we say that F is ν -firmly nonexpansive on S if $\|F(x) - F(y)\|^2 \leq \|x - y\|^2 - \nu\|(x - y) - (F(x) - F(y))\|^2$ for all $(x, y) \in S^2$. If F is 1-firmly nonexpansive on S , then we also say that F is firmly nonexpansive. Various examples of ν -firmly nonexpansive operators for $\nu > 0$ can be found in [18, Chapter 2, Section 2.2].

If S is nonempty, closed and convex, then it is well-known [18, Theorem 1.2.3 on p. 18] that for all $x \in X$ there exists a unique point in S , denoted by $P_S(x)$, such that $\|x - P_S(x)\| = d(x, S) := \inf\{\|x - y\| \mid y \in S\}$. We call the operator $P_S : X \rightarrow S$, which assigns to each $x \in X$ the unique point $P_S(x)$, the orthogonal projection onto S (also called best approximation projection, metric projection, projector, and proximity map). It is well known that P_S is firmly nonexpansive [18, Theorem 2.2.21(iii) on p. 76 and Theorem 2.2.10(v) on p. 70] and satisfies the so-called Kolmogorov property $\langle x - P_S(x), y - P_S(x) \rangle \leq 0$ for all $x \in X$ and $y \in S$ (see, e.g., [18, Theorem 2.2.21(ii) on p. 76 and Theorem 2.2.5 on p. 67] or [28, The Lemma on p. 448]).

We say that S is strictly convex if for all distinct points $a \in S$ and $b \in S$, the open interval (a, b) connecting a and b is contained in the interior $\text{int}(S)$ of S . We denote by Id the identity operator on X which satisfies $\text{Id}(x) = x$ for each $x \in X$. Given $x \in X$ and $r > 0$, we denote by $B[x, r] := \{y \in X \mid \|y - x\| \leq r\}$ and $B(x, r) := \{y \in X \mid \|y - x\| < r\}$ the closed and open balls, respectively, of radius r , about x (and $B[x, 0] := \{x\}$).

Given $m \in \mathbb{N}$ and a vector $\beta := (\beta_i)_{i=1}^m \in \mathbb{R}^m$, we say that β is a “positive weight vector” if $\beta_i > 0$ for all $i \in \{1, 2, \dots, m\}$ and $\sum_{i=1}^m \beta_i = 1$. A sequence $(\tau_k)_{k=0}^\infty$ of real numbers is said to be a “steering parameter sequence” whenever the following conditions hold: $\tau_k \in (0, 1)$ for all $k \in \mathbb{N} \cup \{0\}$, $\lim_{k \rightarrow \infty} \tau_k = 0$, $\sum_{k=0}^\infty \tau_k = \infty$, and $\sum_{k=0}^\infty |\tau_{k+1} - \tau_k| < +\infty$. For example, the sequence $(\tau_k)_{k=0}^\infty$ which satisfies $\tau_k := 1/(k+1)$ for all $k \in \mathbb{N} \cup \{0\}$ is a steering parameter sequence.

3. THE BEST APPROXIMATION TUPLE PROBLEM

In this section we formulate the BAT problem (Subsection 3.1), our algorithm for solving it (Subsection 3.2), and the main existence, uniqueness and convergence theorem as well as a characterization of a BAT (Subsection 3.3). In Subsection 3.4 we make a short comparison between our algorithm and results to the ones obtained by Cheney and Goldstein [28].

3.1. Problem formulation. (See Figure 1 for an illustration.)

Problem 3.1. *The Best Approximation Tuple (BAT) problem:* *In a real Hilbert space X we are given $m \in \mathbb{N}$ sets $(\Omega_i)_{i=1}^m$ (which we call “the satellite sets”), each of which is itself a nonempty intersection of a given family $(\Omega_{i,j})_{j=1}^{m_i}$ of $m_i \in \mathbb{N}$ closed and convex subsets of X , i.e., $\Omega_i := \cap_{j=1}^{m_i} \Omega_{i,j} \neq \emptyset$ for all $i \in \{1, 2, \dots, m\}$. We are also given another set (which we call “the pivot set”) Ω_{m+1} which is nonempty, closed and convex, and a positive weight vector $(\beta_i)_{i=1}^m$. The goal of the “Best Approximation Tuple” (BAT) problem is to solve the following constrained minimization problem:*

$$\min_{x \in \Omega_{m+1}} \frac{1}{2} \sum_{i=1}^m \beta_i (d(x, \Omega_i))^2. \quad (3.1)$$

In other words, if $\Phi : X \rightarrow \mathbb{R}$ is defined by

$$\Phi(x) := \frac{1}{2} \sum_{i=1}^m \beta_i d(x, \Omega_i)^2 = \frac{1}{2} \sum_{i=1}^m \beta_i \|x - P_i(x)\|^2, \quad \forall x \in X, \quad (3.2)$$

where $P_i : X \rightarrow X$ is the orthogonal projection onto Ω_i , $i \in \{1, 2, \dots, m\}$, then the problem is to provide conditions which ensure the existence and perhaps uniqueness of a minimizer of Φ over Ω_{m+1} , and to present algorithms for finding such a minimizer.

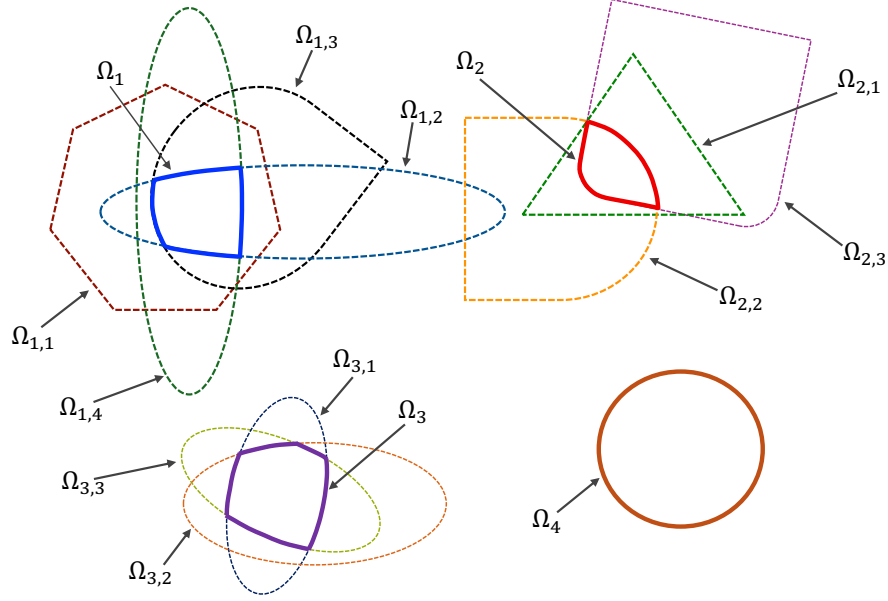


FIGURE 1. Illustration of Problem 3.1 in the Euclidean plane. Here $m = 3$, $m_1 = 4$, $m_2 = 3$, $m_3 = 3$, the boundaries of the sets $\Omega_1, \Omega_2, \Omega_3, \Omega_4$ are represented by solid lines, and the boundaries of the sets $\Omega_{i,j}$ are represented by dashed lines, $i \in \{1, 2, \dots, m\}$, $j \in \{1, 2, \dots, m_i\}$. The satellite sets are $\Omega_1, \Omega_2, \Omega_3$ and the pivot set is Ω_4 .

Definition 3.2. A tuple $(y^1, y^2, \dots, y^{m+1}) \in \Omega_1 \times \Omega_2 \times \dots \times \Omega_{m+1}$ is called a BAT (relative to $(\Omega_i)_{i=1}^{m+1}$) if y^{m+1} solves (1.2) and $y^i = P_i(y^{m+1})$ for each $i \in \{1, 2, \dots, m\}$.

3.2. The Multiple-Sets Best Approximation Tuple algorithm. Here we propose a projection-type algorithm for the solution of the BAT problem.

Algorithm 3.3. (Algorithm for the BAT problem)

- (1) **Input:** A real Hilbert space $(X, \|\cdot\|)$, a natural number m , a set of m natural numbers m_i and sequences $(\tau_{i,t})_{t=0}^\infty$ of steering parameters for each $i \in \{1, 2, \dots, m\}$, positive weight vectors $w_i = (w_{i,j})_{j=1}^{m_i}$ where $i \in \{1, 2, \dots, m\}$, closed and convex subsets Ω_i (the satellite sets) and $\Omega_{i,j}$ such that $\Omega_i = \bigcap_{j=1}^{m_i} \Omega_{i,j} \neq \emptyset$.

$\emptyset$ for all $i \in \{1, 2, \dots, m\}$ and $j \in \{1, 2, \dots, m_j\}$, a nonempty closed and convex subset Ω_{m+1} (the pivot set), a positive weight vector $\beta = (\beta_i)_{i=1}^m$, and two sequences $(\lambda_k)_{k=0}^\infty$ and $(\gamma_k)_{k=0}^\infty$ of real numbers in $(0, 1]$.

(2) **Initialization:** Choose an arbitrary starting point $x^0 \in X$.

(3) **Iterative Step:** For each $k \in \mathbb{N} \cup \{0\}$ and $i \in \{1, 2, \dots, m\}$, define

$$\widehat{P}_{i,k}(x) := \prod_{t=0}^k \left(\tau_{i,t}x + (1 - \tau_{i,t}) \sum_{j=1}^{m_i} w_{i,j} P_{i,j}(x) \right), \quad \forall x \in X, \quad (3.3)$$

$$G_k(x) := (1 - \lambda_k)x + \lambda_k P_{m+1} \left((1 - \gamma_k)x + \gamma_k \sum_{i=1}^m \beta_i \widehat{P}_{i,k}(x) \right), \quad \forall x \in X, \quad (3.4)$$

where $P_{i,j}$ is the orthogonal projection onto the closed and convex subset $\Omega_{i,j}$, $i \in \{1, 2, \dots, m\}$ and $j \in \{1, 2, \dots, m_j\}$ and P_{m+1} is the orthogonal projection onto the pivot set Ω_{m+1} . Now, given a current iterate x^k , $k \in \mathbb{N} \cup \{0\}$, define the next iterate x^{k+1} by the following iterative process:

$$x^{k+1} := G_k(x^k) = (1 - \lambda_k)x^k + \lambda_k P_{m+1} \left((1 - \gamma_k)x^k + \gamma_k \sum_{i=1}^m \beta_i \widehat{P}_{i,k}(x^k) \right). \quad (3.5)$$

Remark 3.4. (i) The vector $(\beta_i)_{i=1}^m$ in the functional Φ from (3.2) is assumed to be a positive weight vector. This is not a real restriction as long as the coefficients β_i , $i \in \{1, 2, \dots, m\}$ are assumed to be positive. Indeed, suppose that instead of a positive weight vector we just assume that $(\beta_i)_{i=1}^m$ is a vector of positive numbers. Then by letting $\bar{\beta}_i := \beta_i / (\sum_{j=1}^m \beta_j)$ for each $i \in \{1, 2, \dots, m\}$ and $\bar{\Phi}(x) := \frac{1}{2} \sum_{i=1}^m \bar{\beta}_i d(x, \Omega_i)^2$, $x \in X$, we see that $(\bar{\beta}_i)_{i=1}^m$ is a weight vector, and an immediate verification shows that $\bar{\Phi}$ and $\Phi(x) := \frac{1}{2} \sum_{i=1}^m \beta_i d(x, \Omega_i)^2$, $x \in X$, have the same minimizers over any subset S of X .

(ii) The iterative scheme (3.5) is inspired by a different iterative scheme presented in [35, (20) on p. 2463], whose goal is to solve the inconsistent convex feasibility problem with soft and hard constraints, where each subset is formed by itself (namely $m_i = 1$ for all $i \in \{1, 2, \dots, m\}$).

3.3. The main existence, uniqueness and convergence theorem. The proof of the next theorem, which is the main convergence result, is long and is based on several lemmata which we formulate and prove in the next sections.

Theorem 3.5. *In the setting of Algorithm 3.3, suppose that either Ω_{m+1} is bounded or that Ω_{m+1} is unbounded and Ω_q is bounded for some $q \in \{1, 2, \dots, m\}$. Then (3.1) has at least one solution. If, in addition, Ω_s is strictly convex for some $s \in \{1, 2, \dots, m\}$ and also $\Omega_s \cap \Omega_{m+1} = \emptyset$, then (3.1) has a unique solution x^* , and x^* has the property that it is also the unique fixed point of each of the functions $G_{\lambda,\gamma} : X \rightarrow X$, $\lambda \in (0, 1]$, $\gamma \in (0, 2)$, defined by*

$$G_{\lambda,\gamma}(x) := (1 - \lambda)x + \lambda P_{m+1}((1 - \gamma)x + \gamma \sum_{i=1}^m \beta_i P_i(x)), \quad \forall x \in X. \quad (3.6)$$

Assume further that X is finite dimensional, that there is some $\rho > 0$ such that $\Omega_{m+1} \cup (\cup_{i=1}^m \cup_{j=1}^{m_i} \Omega_{i,j}) \subseteq B[0, \rho]$, that Ω_i is strictly convex and satisfies $\Omega_i \cap \Omega_{m+1} = \emptyset$ for each $i \in \{1, 2, \dots, m\}$, and that $\lambda_\infty := \lim_{k \rightarrow \infty} \lambda_k$ and $\gamma_\infty := \lim_{k \rightarrow \infty} \gamma_k$ exist and satisfy $\lambda_\infty > 0$ and $\gamma_\infty > 0$. If $x^0 \in \Omega_{m+1}$, then any sequence $(x^k)_{k=0}^\infty$ generated by Algorithm 3.3 is contained in Ω_{m+1} and converges to x^* .

- Remark 3.6.** (i) Non-uniqueness of the optimal solution x^* of (3.1) can occur without imposing the strict convexity assumption on the sets in Theorem 3.5. Indeed, suppose that X is the Euclidean plane, $\Omega_1 := [-10, -\eta_1] \times [0, 1]$, $\Omega_2 := [\eta_2, 10] \times [0, 1]$, $\eta_1, \eta_2 \in (0, 10)$, $\Omega_3 := [-0.5\eta_1, 0.5\eta_2] \times [0, 1]$, $m = 2$, $m_1 = m_2 = 1$, and (β_1, β_2) is an arbitrary positive weight vector. In this case there exists a solution $x^* = (x_1^*, x_2^*)$ to (3.1) because of considerations mentioned in Lemma 4.5 below. Since Φ is constant on intervals of the form $\{t\} \times [0, 1]$ for each $t \in [-0.5\eta_1, 0.5\eta_2]$ as a simple verification shows (because $d((x_1, x_2), \Omega_1) = x_1 + \eta_1$ and $d((x_1, x_2), \Omega_2) = \eta_2 - x_1$ for all $(x_1, x_2) \in \Omega_3$), it follows that for every $x_2 \in [0, 1]$ the pair (x_1^*, x_2) solves (3.1).
- (ii) Non-uniqueness of the optimal solution of (3.1) can occur by replacing the strictly convex functional Φ defined in (3.2) by the non-strictly convex functional $\Psi(x) := \frac{1}{2} \sum_{i=1}^m \beta_i d(x, \Omega_i)$, $x \in X$. Indeed, just take $X := \mathbb{R}$, $m := 2$, $m_1 := m_2 := 1$, $\Omega_1 := \{0\}$, $\Omega_2 := \{2\}$, $\Omega_3 := [0.5, 1.5]$, $\beta_1 := \beta_2 := 0.5$. In this case $\Psi(x) = 0.5$ for all $x \in \Omega_3$. Similarly, if $X := \mathbb{R}^2$, $m := 2$, $m_1 := m_2 := 1$, $\Omega_1 := B[(-1, 0), 1]$, $\Omega_2 := B[(3, 0), 1]$, $\Omega_3 := B[(1, 0), 0.5]$, $\beta_1 := \beta_2 := 0.5$, then $\Psi(x) = 0.5 \leq \Psi(z)$ for all $x \in [0.5, 1.5] \times \{0\} \subset \Omega_3$ and all $z \in \Omega_3$.
- (iii) While in the last part of Theorem 3.5 we require that Ω_i is strictly convex and satisfies $\Omega_i \cap \Omega_{m+1} = \emptyset$ for each $i \in \{1, 2, \dots, m\}$, we do not require that $\Omega_{i,j}$ is strictly convex and satisfies $\Omega_{i,j} \cap \Omega_{m+1} = \emptyset$ and $\Omega_{i,j} \cap \Omega_{i',j'} = \emptyset$ for each $i, i' \in \{1, 2, \dots, m\}$ and each $j \in \{1, 2, \dots, m_i\}$ and $j' \in \{1, 2, \dots, m_{i'}\}$. Hence, in particular, it may happen that Ω_{i_1,j_1} intersects both Ω_{i_2,j_2} and Ω_{m+1} for some $i_1, i_2 \in \{1, 2, \dots, m\}$ and some $j_1 \in \{1, 2, \dots, m_{i_1}\}$ and $j_2 \in \{1, 2, \dots, m_{i_2}\}$: see Figure 1. We note that a sufficient condition for Ω_i to be strictly convex is that $\Omega_{i,j}$ is strictly convex for all $j \in \{1, 2, \dots, m_i\}$, but, as shown in Figure 1, this condition is not necessary.
- (iv) In the special case where $m = 1$ in Theorem 3.5 we do not need to impose the boundedness of either Ω_1 or Ω_2 for the existence of a solution to (3.1) (which, as explained earlier, is equivalent to the existence of a BAP relative to (Ω_1, Ω_2)): see [71, Theorem 5.1] for less restrictive assumptions.

The next proposition characterizes a BAT as the solution of a multivariate minimization problem. This characterization can serve as an alternative definition of a BAT, a definition which actually holds in any normed space and with any tuple of nonempty sets $(\Omega_i)_{i=1}^{m+1}$ (not necessarily closed and convex).

Proposition 3.7. A tuple $(y^1, y^2, \dots, y^{m+1}) \in \Omega_1 \times \Omega_2 \times \dots \times \Omega_{m+1}$ is a BAT if and only it solves the minimization problem (1.3).

Proof. Indeed, suppose by way of contradiction that $(y^1, y^2, \dots, y^{m+1})$ is a BAT which does not solve (1.3). Then $\sum_{i=1}^m \beta_i \|x^{m+1} - x^i\|^2 < \sum_{i=1}^m \beta_i \|y^{m+1} - y^i\|^2$ for some

$(x^1, x^2, \dots, x^{m+1}) \in \Omega_1 \times \Omega_2 \times \dots \times \Omega_{m+1}$. Since $x^i \in \Omega_i$ for each $i \in \{1, 2, \dots, m\}$, we have $\|x^{m+1} - P_i(x^{m+1})\| \leq \|x^{m+1} - x^i\|$ by the definition of the orthogonal projection; thus $0.5 \sum_{i=1}^m \beta_i \|x^{m+1} - P_i(x^{m+1})\|^2 \leq 0.5 \sum_{i=1}^m \beta_i \|x^{m+1} - x^i\|^2$. Since $(y^1, y^2, \dots, y^{m+1})$ is a BAT, we have $y^i = P_i(y^{m+1})$ for all $i \in \{1, 2, \dots, m\}$; as a result, we conclude from the previous lines that $0.5 \sum_{i=1}^m \beta_i \|x^{m+1} - P_i(x^{m+1})\|^2 < 0.5 \sum_{i=1}^m \beta_i \|y^{m+1} - P_i(y^{m+1})\|^2$, in contrast to the assumption y^{m+1} solves (1.2) (since $(y^1, y^2, \dots, y^{m+1})$ is a BAT). This contradiction implies that any BAT must solve (1.3).

Conversely, suppose that $(y^1, y^2, \dots, y^{m+1}) \in \Omega_1 \times \Omega_2 \times \dots \times \Omega_{m+1}$ solves (1.3) and assume by way of contradiction that it is not BAT. Then there are two possibilities (see Definition 3.2): either $y^j \neq P_j(y^{m+1})$ for some $j \in \{1, 2, \dots, m\}$ or $y^i = P_i(y^{m+1})$ for all $i \in \{1, 2, \dots, m\}$ but y^{m+1} does not solve (1.2).

Consider the first possibility. Basic properties of the orthogonal projection and the fact that $y^i \in \Omega_i$ for all $i \in \{1, 2, \dots, m\}$ imply the inequality $\|y^{m+1} - P_i(y^{m+1})\| \leq \|y^{m+1} - y^i\|$ for each $i \in \{1, 2, \dots, m\}$, where for $i = j$ we have a strict inequality. Hence $0.5 \sum_{i=1}^m \beta_i \|y^{m+1} - P_i(y^{m+1})\|^2 < 0.5 \sum_{i=1}^m \beta_i \|y^{m+1} - y^i\|^2$, namely the target functional which appears in (1.3) has a strictly lower value at the tuple $(x^i)_{i=1}^{m+1} := (P_1(y^{m+1}), P_2(y^{m+1}), \dots, P_m(y^{m+1}), y^{m+1})$, which is in $\Omega_1 \times \Omega_2 \times \dots \times \Omega_{m+1}$, than it has at its assumed minimizer $(y^1, y^2, \dots, y^{m+1})$, a contradiction. Thus any solution of (1.3) must be a BAT in the first possibility.

Consider the second possibility in which $y^i = P_i(y^{m+1})$ for all $i \in \{1, 2, \dots, m\}$ but y^{m+1} does not solve (1.2). This possibility implies that there is some $x^{m+1} \in \Omega_{m+1}$ such that $0.5 \sum_{i=1}^m \beta_i \|x^{m+1} - P_i(x^{m+1})\|^2 < 0.5 \sum_{i=1}^m \beta_i \|y^{m+1} - P_i(y^{m+1})\|^2$. Because $(y^1, y^2, \dots, y^{m+1}) \in \Omega_1 \times \Omega_2 \times \dots \times \Omega_{m+1}$ solves (1.3), we obtain the inequality $0.5 \sum_{i=1}^m \beta_i \|y^{m+1} - y^i\|^2 \leq 0.5 \sum_{i=1}^m \beta_i \|x^{m+1} - x^i\|^2$ for every $i \in \{1, 2, \dots, m\}$ and every $x^i \in \Omega_i$, and in particular for $x^i := P_i(x^{m+1})$. As a result we obtain the inequality $0.5 \sum_{i=1}^m \beta_i \|y^{m+1} - P_i(y^{m+1})\|^2 \leq 0.5 \sum_{i=1}^m \beta_i \|x^{m+1} - P_i(x^{m+1})\|^2$ and this is a contradiction to a previous line from which we already know that $0.5 \sum_{i=1}^m \beta_i \|x^{m+1} - P_i(x^{m+1})\|^2 < 0.5 \sum_{i=1}^m \beta_i \|y^{m+1} - P_i(y^{m+1})\|^2$. We conclude that any solution to (1.3) must be a BAT in the second possibility too. $\square$

3.4. A short comparison with Cheney and Goldstein [28]. It is worth comparing briefly our results and algorithm with the ones obtained by Cheney and Goldstein in their celebrated 1959 paper [28].

First, if in the BAT Problem 3.1 we have $m = 1 = m_1$, then by what we explained earlier (below (1.2)) the BAT problem, namely Problem 3.1, reduces to BAT problem considered in [28]. Hence the BAT problem generalizes the BAP problem from two subsets to any finite number (greater than or equal 2) of subsets and from stand-alone satellite subsets to satellite subsets that each one of them is obtained from intersections of finitely many subsets.

Second, if in Algorithm 3.3 we have $m = 1 = m_1$, assume that $\lambda_k = \gamma_k = 1$ for all $k \in \mathbb{N} \cup \{0\}$, and replace $\hat{P}_{1,k}$ by P_1 (see also Remark 5.5 below for the justification for this latter replacement), then (3.5) leads to $x^{k+1} = P_2 P_1(x^k)$ for each $k \in \mathbb{N} \cup \{0\}$. In other words, we restore the alternating projection method presented in [28].

Third, if we take $\lambda := 1 =: \gamma$ in Lemma 4.6 below, then we see that Lemma 4.6 generalizes [28, Theorem 2] which says that x^* solves the minimization problem $\min_{x \in \Omega_2} d(x, \Omega_1)$ if and only if $x^* = P_2(P_1(x^*))$.

Fourth, the convergence part of Theorem 3.5 above extends, but does not generalize, [28, Theorem 4] since the setting in [28, Theorem 4] is a possibly infinite-dimensional real Hilbert space, with possibly unbounded and non-strictly convex subsets, while in the convergence part of Theorem 3.5 above we require that the space is a finite-dimensional, that all the subsets (pivot and satellite) are compact, and we also require the satellite subsets to be strictly convex.

4. PREPARATORY RESULTS

In this section we formulate and prove (or refer to known proofs in the literature) several preparatory results which are needed for the proof of Theorem 3.5. Some of the results hold in a setting which is more general than the setting of this theorem.

Lemma 4.1. *Suppose that $(Y, \|\cdot\|)$ is a normed space, $\emptyset \neq S \subseteq Y$, and $F : S \rightarrow Y$ is ν -firmly nonexpansive for some $\nu > 0$. Given $(x, y) \in S^2$, one has $\|F(x) - F(y)\| = \|x - y\|$ if and only if $F(x) - F(y) = x - y$, namely, if and only if the points $x, y, F(x)$ and $F(y)$ are the four vertices of the possibly degenerate parallelogram $\{x + t(y - x) + s(F(x) - x) \mid (s, t) \in [0, 1]^2\}$ (see Figures 2 and 3).*

Proof. If $F(x) - F(y) = x - y$, then obviously $\|F(x) - F(y)\| = \|x - y\|$. Conversely, if $\|F(x) - F(y)\| = \|x - y\|$, then since F is ν -firmly nonexpansive, the inequality $\|F(x) - F(y)\|^2 \leq \|x - y\|^2 - \nu\|(x - y) - (F(x) - F(y))\|^2$ holds. This inequality and the nonnegativity of the norm imply that $0 \leq \|(x - y) - (F(x) - F(y))\|^2 \leq (1/\nu)(\|F(x) - F(y)\|^2 - \|x - y\|^2) = 0$. Hence $\|(x - y) - (F(x) - F(y))\|^2 = 0$ and $(x - y) - (F(x) - F(y)) = 0$, as required. $\square$



FIGURE 2. The setting of Lemma 4.1: a nondegenerate parallelogram.



FIGURE 3. The settings of Lemma 4.1: a degenerate parallelogram.

Remark 4.2. (i) A nonexpansive operator $F : S \rightarrow Y$ for which the equality $\|F(x) - F(y)\| = \|x - y\|$ implies $F(x) - F(y) = x - y$ is called in [56, p. 556] and [45, Definition 2] (in the case where $S = Y = \mathbb{R}^n$ for some $n \in \mathbb{N}$) strictly nonexpansive (unfortunately, the term “strictly nonexpansive” has other meanings in the literature: see Section 2 above). We conclude from Lemma 4.1 that any ν -firmly nonexpansive operator ($\nu > 0$) is strictly nonexpansive. For an interesting class of

strictly nonexpansive operators which is related to Lemma 4.3 below see [41, Definition 1, Proposition 1, Proposition 2, Proposition 8].

- (ii) Since the orthogonal projection P_S onto a nonempty, closed and convex subset S of a real Hilbert space X is 1-firmly nonexpansive, we conclude from Lemma 4.1 that for all $(x, y) \in X^2$ one has $\|P_S(x) - P_S(y)\| = \|x - y\|$ if and only if $P_S(x) - P_S(y) = x - y$. In particular, if $\|P_S(x) - P_S(y)\| = \|x - y\|$ for some $(x, y) \in X^2$, then $P_S(x) - x = P_S(y) - y$ and so $\|P_S(x) - x\| = \|P_S(y) - y\|$. This is the necessary condition mentioned in [28, Theorem 3] for equality in the nonexpansiveness inequality $\|P_S(x) - P_S(y)\| \leq \|x - y\|$. Lemma 4.1 presents a more general necessary and sufficient condition.

Lemma 4.3 below is closely related to [41, Proposition 8]. Our arguments, however, are different from the ones presented in the proof of [41, Proposition 8] and have a more geometric flavour.

Lemma 4.3. *Let S be a nonempty, closed, and convex subset of a real Hilbert space $(Y, \langle \cdot, \cdot \rangle)$ with norm $\|\cdot\|$, $Y \neq \{0\}$. Let $x, y \in Y$ be arbitrary such that $x \neq y$ and either $x \notin S$ or $y \notin S$. If $\dim(Y) = 1$, then $\|P_S(x) - P_S(y)\| < \|x - y\|$. If $\dim(Y) > 1$ (possibly $\dim(Y) = \infty$) and $\|P_S(x) - P_S(y)\| = \|x - y\|$, then all the points $x, y, P_S(x)$ and $P_S(y)$ are distinct, they are located on the same affine plane, the convex hull spanned by them is a rectangle where each of these points is a corner of the rectangle, and $x - y$ is orthogonal to $P_S(y) - y$ (see Figure 5).*

Proof. Since either $x \notin S$ or $y \notin S$, we can assume that $y \notin S$ (the case where $x \notin S$ can be treated in a similar manner). First we show that if $\|P_S(x) - P_S(y)\| = \|x - y\|$, then the points x, y and $P_S(y)$ are distinct and are not located on the same line. Indeed, under our assumptions, since $\|P_S(x) - P_S(y)\| = \|x - y\|$ and P_S is 1-firmly nonexpansive [9, Proposition 4.16, p. 74], it follows from Lemma 4.1 that

$$P_S(x) - P_S(y) = x - y. \quad (4.1)$$

Since $y \notin S$, we have $y \neq P_S(y) \in S$. If $x = P_S(y)$, then $x \in S$ and hence $x = P_S(x)$. Thus (4.1) implies that $y - P_S(y) = 0$, a contradiction. Hence $x \neq P_S(y)$, and since we already assume that $x \neq y$ and showed above that $y \neq P_S(y)$, we conclude that all the points x, y and $P_S(y)$ are distinct.

Next we show that the points x, y and $P_S(y)$ cannot be located on the same line. Indeed, assume by way of contradiction that they are located on the same line L . Since $y \neq P_S(y)$ and both y and $P_S(y)$ belong to L , we can represent L as $L = \{y + t(P_S(y) - y) \mid t \in \mathbb{R}\}$. Since $x \in L$, we have $x = y + t_x(P_S(y) - y)$ for some $t_x \in \mathbb{R}$. Since $x \neq y$ and $x \neq P_S(y)$, we have $t_x \neq 0$ and $t_x \neq 1$. Hence there are two cases to consider:

Case 1. Either $t_x < 0$ or $t_x \in (0, 1)$: In this case the point x is on the ray $\{P_S(y) + s(y - P_S(y)) \mid s \in [0, \infty)\}$ emanating from $P_S(y)$ in the direction of y , and hence $P_S(x) = P_S(y)$ because the orthogonal projection is a sunny retraction, namely $P_S(z) = P_S(y)$ for all z on the ray emanating from $P_S(y)$ in the direction of y : see [48, p. 17] (this fact can be proved directly, using basic properties of orthogonal projections such

as the so-called Kolmogorov property mentioned below in (4.2); see [72, Lemma 2.7] for a much more general result). Thus (4.1) implies that $x = y$, which is a contradiction.

Case 2. $t_x > 1$: In this case either $x \in S$ or $x \notin S$. If $x \in S$, then $x = P_S(x)$, and then (4.1) implies that $P_S(y) = y$, a contradiction. Hence $x \notin S$ and, therefore, $x \neq P_S(x) \in S$. From (4.1) and the equality $x = y + t_x(P_S(y) - y)$ we have $P_S(x) = x + P_S(y) - y = y + t_x(P_S(y) - y) + P_S(y) - y = y + (t_x + 1)(P_S(y) - y)$, namely $P_S(x) \in L$ with $t := t_x + 1$.

Since $1 < t_x < t_x + 1$, it follows that $x = y + t_x(P_S(y) - y)$ is between $P_S(y) = y + 1 \cdot (P_S(y) - y)$ and $P_S(x) = y + (t_x + 1)(P_S(y) - y)$, and so the convexity of S implies that $x \in S$. But then $x = P_S(x)$, and hence (4.1) implies that $y = P_S(y) \in S$, a contradiction to what we assumed above.

The previous analysis shows that if $\dim(Y) = 1$, $x \neq y$ and either $x \notin S$ or $y \notin S$, then the equality $\|P_S(x) - P_S(y)\| = \|x - y\|$ cannot hold because obviously when $\dim(Y) = 1$, then x , y and $P_S(y)$ are located on the same line. As a result, since P_S is nonexpansive, we conclude that $\|P_S(x) - P_S(y)\| < \|x - y\|$ in this case.

From now on we assume that $\dim(Y) \geq 2$ (possibly $\dim(Y) = \infty$), together with the assumptions from the beginning of the proof (that is, $x \neq y$, $\|P_S(x) - P_S(y)\| = \|x - y\|$ and $y \notin S$). In this case, by what we have shown above, we know that x , y and $P_S(y)$ are distinct and are not located on the same line. Hence the affine space generated by x , y and $P_S(y)$, namely $E := P_S(y) + \text{span}\{x - P_S(y), y - P_S(y)\}$, is two-dimensional.

Since $x \neq y$, we have $\|P_S(x) - P_S(y)\| = \|x - y\| > 0$, and, therefore, $P_S(x) \neq P_S(y)$. In addition, it must be true that $x \neq P_S(x)$: indeed, if $x = P_S(x)$, then (4.1) implies that $y = P_S(y) \in S$, a contradiction. Since $y \notin S$, we also have $y \neq P_S(x) \in S$, and we conclude that all the four points x , y , $P_S(x)$ and $P_S(y)$ are distinct.

It follows from (4.1) that $P_S(x) = P_S(y) + 1 \cdot (x - P_S(y)) + (-1) \cdot (y - P_S(y)) \in E$, and so all the four points x , y , $P_S(x)$ and $P_S(y)$ are on the same affine plane E . Moreover, from (4.1) and the fact that all the points are different from each other it follows that the nondegenerate interval $[P_S(x), P_S(y)]$ is parallel to the nondegenerate interval $[x, y]$.

The equality (4.1) implies that $P_S(x) - x = P_S(y) - y$, and hence the non-degenerate interval $[P_S(x), x]$ is parallel to the nondegenerate interval $[P_S(y), y]$. Thus x , y , $P_S(y)$, and $P_S(x)$ form a parallelogram: see Figure 4.

By the, so-called, Kolmogorov property of orthogonal projections, we have

$$\langle x - P_S(x), z - P_S(x) \rangle \leq 0, \quad \forall z \in S. \quad (4.2)$$

In particular, (4.2) holds with $z := P_S(y)$, namely the angle $\alpha_1 := \angle xP_S(x)P_S(y)$ between $x - P_S(x)$ and $P_S(y) - P_S(x)$ satisfies $\alpha_1 \geq \pi/2$. Similarly (namely by replacing x by y in (4.2) and letting there $z := P_S(x)$), $\alpha_2 := \angle yP_S(y)P_S(x)$ satisfies $\alpha_2 \geq \pi/2$. Since $[P_S(x), x]$ is parallel to $[P_S(y), y]$, basic Euclidean plane geometry tells us that $\alpha_1 + \alpha_2 = \pi$. Thus, it must be that $\alpha_1 = \pi/2$, since, otherwise, $\alpha_1 > 0.5\pi$ and hence $\pi = 0.5\pi + 0.5\pi < \alpha_1 + \alpha_2 = \pi$, a contradiction. Therefore, $\alpha_2 = \pi - \alpha_1 = \pi/2$, and we conclude that the parallelogram $xyP_S(y)P_S(x)$ is actually a rectangle and $x - y$ is orthogonal to $P_S(y) - y$: see Figure 5. $\square$

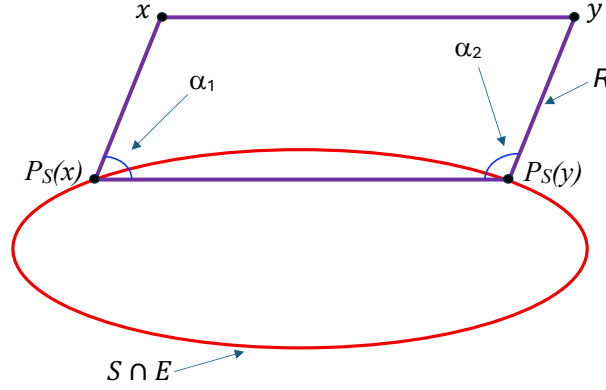


FIGURE 4. The setting of Lemma 4.3, where the points $x, y, P_S(y), P_S(x)$ are the corners of a parallelogram located on the affine plane E .

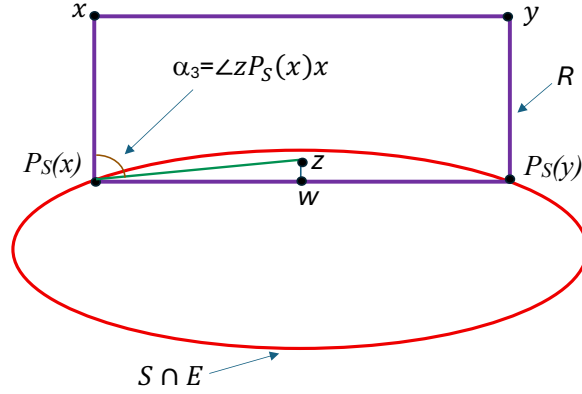


FIGURE 5. The settings of Lemma 4.3 and Lemma 4.4, where the points $x, y, P_S(y), P_S(x)$ are the corners of a rectangle R located on the affine plane E , the point z is an interior point of $R \cap S$, and $\alpha_3 := \angle zP_S(x)x$ satisfies $0 < \alpha_3 < 0.5\pi$.

The next lemma is known [50, Lemma 2], but is not very well-known, and we present a new proof of it. We also remark that there are issues in the English translation of [50, Lemma 2] (and of [50] in general), and these issues may have contributed to the fact that [50, Lemma 2] is not very known. The issues are that it should be “strictly convex” instead of “strongly convex” in the line located between (5) and (5’) in the formulation of [50, Lemma 1] and also in the sixth line of the proof of this lemma (this lemma is used in the proof of [50, Lemma 2]), and it should be “strictly convex” instead of “strongly convex” also in the line located between (6) and (6’) in the formulation [50, Lemma 2].

Lemma 4.4. *Let S be a nonempty, closed, and strictly convex subset of a real Hilbert space $(Y, \langle \cdot, \cdot \rangle)$ with norm $\|\cdot\|$, $Y \neq \{0\}$. Let $x, y \in Y$ be arbitrary such that $x \neq y$ and either $x \notin S$ or $y \notin S$. Then $\|P_S(x) - P_S(y)\| < \|x - y\|$. Stated differently, for all x*

and y in Y which satisfy $\|P_S(x) - P_S(y)\| = \|x - y\|$ it must hold that either $x = y$ or both $x \in S$ and $y \in S$.

Proof. The case $\dim(Y) = 1$ follows from Lemma 4.3. Hence from now on $\dim(Y) \geq 2$ (possibly $\dim(Y) = \infty$). Assume for a contradiction that $\|P_S(x) - P_S(y)\| = \|x - y\|$. Then Lemma 4.3 implies that all the four points x , y , $P_S(x)$ and $P_S(y)$ are distinct and each of them is a corner of a rectangle R in which $x - y$ is orthogonal to $P_S(y) - y$ (and hence $P_S(y) - P_S(x)$ is orthogonal to $x - P_S(x)$). See Figure 5.

Since S is strictly convex and $P_S(x) \in S$ and $P_S(y) \in S$, any point in the open line segment $(P_S(x), P_S(y))$ is in the interior of S . In particular, $w := 0.5(P_S(x) + P_S(y))$ is in the interior of S and hence there is some $r \in (0, \|x - P_S(x)\|)$ such that the open ball of center w and radius r is contained in S . In particular, the point $z := w + (0.5r/\|x - P_S(x)\|)(x - P_S(x))$ is in S , and it is also strictly inside R since $r < \|x - P_S(x)\|$ and $z = P_S(x) + 0.5(P_S(y) - P_S(x)) + (0.5r/\|x - P_S(x)\|)(x - P_S(x))$ and $R = \{P_S(x) + \eta_1(P_S(y) - P_S(x)) + \eta_2(x - P_S(x)) \mid (\eta_1, \eta_2) \in [0, 1]^2\}$ (and the interior of R is $\{P_S(x) + \eta_1(P_S(y) - P_S(x)) + \eta_2(x - P_S(x)) \mid (\eta_1, \eta_2) \in (0, 1)^2\}$). This fact and the fact that $P_S(y) - P_S(x)$ is orthogonal to $x - P_S(x)$ imply that

$$\begin{aligned} & \langle x - P_S(x), z - P_S(x) \rangle \\ &= \frac{1}{2} \langle x - P_S(x), P_S(y) - P_S(x) \rangle + \frac{r}{2\|x - P_S(x)\|} \langle x - P_S(x), x - P_S(x) \rangle = \frac{1}{2} r \|x - P_S(x)\|. \end{aligned}$$

Hence, if we let $\alpha_3 := \angle z P_S(x) x$, i.e., the angle between $z - P_S(x)$ and $x - P_S(x)$, then $\cos(\alpha_3) = \langle x - P_S(x), z - P_S(x) \rangle / (\|x - P_S(x)\| \|z - P_S(x)\|) = 0.5r / \|z - P_S(x)\| > 0$. Since

$$\|z - P_S(x)\| = \sqrt{\|z - w\|^2 + \|w - P_S(x)\|^2} = \sqrt{(0.5r)^2 + \|0.5(P_S(y) - P_S(x))\|^2} > \frac{r}{2},$$

we have $\cos(\alpha_3) = 0.5r / \|z - P_S(x)\| < 1$. Thus, $0 < \alpha_3 < \pi/2$ (see Figure 5), a contradiction to (4.2) which holds since $z \in S$. The contradiction shows that the equality $\|P_S(x) - P_S(y)\| = \|x - y\|$ is impossible, and since P_S is nonexpansive we conclude that $\|P_S(x) - P_S(y)\| < \|x - y\|$. $\square$

Lemma 4.5. *In the setting of Algorithm 3.3, if, either Ω_{m+1} is bounded or it is unbounded and Ω_q is bounded for some $q \in \{1, 2, \dots, m\}$, then there exists a solution to (3.1) in Ω_{m+1} .*

Proof. The real-valued functional Φ defined in (3.2) is continuous and convex over X , as a sum of squares of nonnegative, convex and continuous functions. Hence (as follows from [13, Corollary 3.9, p. 61]), it is lower semicontinuous in the weak topology defined on X .

Now, if the nonempty, closed and convex subset Ω_{m+1} is bounded, then Ω_{m+1} is weakly compact since X is reflexive [13, Corollary 3.22, p. 71] and hence, by the Weirestrass Extreme Value Theorem, Φ attains a minimum over Ω_{m+1} , namely there exists a solution to (3.1). If Ω_{m+1} is unbounded, then since we assume that Ω_q is bounded for some $q \in \{1, \dots, m\}$, we have that $\|P_q\|$ is bounded over X (because if $x \in X$ is arbitrary, and $v \in \Omega_q$ is fixed independently of x , then the nonexpansivity of P_q and the triangle inequality imply that $\|P_q(x)\| \leq \|P_q(x) - P_q(v)\| + \|P_q(v)\| \leq \text{diam}(\Omega_q) + \|P_q(v)\| < \infty$).

Hence, the triangle inequality implies that $\infty = \lim_{\|x\| \rightarrow \infty, x \in \Omega_{m+1}} (\|x\| - \|P_q(x)\|) \leq \lim_{\|x\| \rightarrow \infty, x \in \Omega_{m+1}} \|x - P_q(x)\|$. Therefore, $\infty = \lim_{\|x\| \rightarrow \infty, x \in \Omega_{m+1}} 0.5\beta_q \|x - P_q(x)\|^2 \leq \lim_{\|x\| \rightarrow \infty, x \in \Omega_{m+1}} \Phi(x)$ (also because we assume that $\beta_i > 0$ for all $i \in \{1, 2, \dots, m\}$: see (3.2)). Thus, Φ is coercive, convex and continuous (in particular, lower semicontinuous), and so we can use [13, Corollary 3.23, p. 71] to conclude that Φ has a minimizer in Ω_{m+1} . Hence, (3.1) has a solution. $\square$

Lemma 4.6. *In the setting of Algorithm 3.3, denote by $\text{Argmin}_{\Omega_{m+1}}(\Phi)$ the set of solutions to (3.1), and denote by $\text{Fix}_S(G_{\lambda,\gamma})$ the set of fixed points of the operator $G_{\lambda,\gamma}$ over a subset S of X for each $\lambda \neq 0$ and $\gamma > 0$, where $G_{\lambda,\gamma}$ is defined in (3.6). Then $\text{Argmin}_{\Omega_{m+1}}(\Phi) = \text{Fix}_X(G_{\lambda,\gamma}) = \text{Fix}_{\Omega_{m+1}}(G_{\lambda,\gamma})$ for each $\lambda \neq 0$ and each $\gamma > 0$, and therefore, x solves (3.1) if and only if x is a fixed point of $G_{\lambda,\gamma}$ for each $\lambda \neq 0$ and each $\gamma > 0$, if and only if x is a fixed point of $G_{\lambda,\gamma}$ for some $\lambda \neq 0$ and some $\gamma > 0$ (and x also belongs to Ω_{m+1}).*

Proof. Part of this assertion essentially appears in [35] (in a somewhat different form), and for the sake of completeness we provide its proof here. We first observe that if $x = G_{\lambda,\gamma}(x)$ for some $x \in X$, $\lambda \neq 0$ and $\gamma \in \mathbb{R}$, then the following equality holds: $x = (1 - \lambda)x + \lambda P_{m+1}((1 - \gamma)x + \gamma \sum_{i=1}^m \beta_i P_i(x))$. Hence, simple arithmetic implies that $x = P_{m+1}((1 - \gamma)x + \gamma \sum_{i=1}^m \beta_i P_i(x)) = G_{1,\gamma}(x)$ and $x \in \Omega_{m+1}$ because P_{m+1} maps X onto Ω_{m+1} . Conversely, if $x = G_{1,\gamma}(x) = P_{m+1}((1 - \gamma)x + \gamma \sum_{i=1}^m \beta_i P_i(x))$ for some $x \in X$ and $\gamma \in \mathbb{R}$, then obviously $x \in \Omega_{m+1}$, and simple arithmetic implies that $x = (1 - \lambda)x + \lambda x = (1 - \lambda)x + \lambda P_{m+1}((1 - \gamma)x + \gamma \sum_{i=1}^m \beta_i P_i(x)) = G_{\lambda,\gamma}(x)$ for all $\lambda \in \mathbb{R}$ (even for $\lambda = 0$). Therefore, $\text{Fix}_X(G_{\lambda,\gamma}) = \text{Fix}_{\Omega_{m+1}}(G_{\lambda,\gamma}) = \text{Fix}_X(G_{1,\gamma})$ for all $\lambda \neq 0$ and all $\gamma \in \mathbb{R}$.

Now denote $f_i(x) := \frac{1}{2}d(x, \Omega_i)^2$ for all $i \in \{1, \dots, m\}$ and all $x \in X$. This function is continuously differentiable and its gradient satisfies $\nabla f_i = \text{Id} - P_i$ (see [18, Lemma 2.2.27, p. 79]). Thus, the function $f : X \rightarrow \mathbb{R}$ defined by $f(x) := \frac{1}{2} \sum_{i=1}^m \beta_i d(x, \Omega_i)^2$ for all $x \in X$, namely, $f = \sum_{i=1}^m \beta_i f_i$, is continuously differentiable. Since $\sum_{i=1}^m \beta_i = 1$ we have $\nabla f = \sum_{i=1}^m \beta_i \nabla f_i = \text{Id} - \sum_{i=1}^m \beta_i P_i$.

From [35, Proposition 3] we have $\text{Argmin}_{\Omega_{m+1}}(\Phi) = \text{Fix}_X(P_{m+1}(\text{Id} - \gamma \nabla f))$ for each $\gamma > 0$. However, the previous lines imply that $P_{m+1}(\text{Id} - \gamma \nabla f) = P_{m+1}((1 - \gamma)\text{Id} + \gamma \sum_{i=1}^m \beta_i P_i) = G_{1,\gamma}$, and we conclude from this fact, as well as from the previous lines and the previous paragraph, the equality $\text{Argmin}_{\Omega_{m+1}}(\Phi) = \text{Fix}_X(G_{1,\gamma}) = \text{Fix}_X(G_{\lambda,\gamma}) = \text{Fix}_{\Omega_{m+1}}(G_{\lambda,\gamma})$ for each $\lambda \neq 0$ and each $\gamma > 0$. $\square$

Lemma 4.7. *In the setting of Algorithm 3.3, suppose that $\lambda \in [0, 1]$ and $\gamma \in [0, 2]$, and let $G_{\lambda,\gamma} : X \rightarrow X$ be defined by (3.6). Then $G_{\lambda,\gamma}$ is nonexpansive over X and maps Ω_{m+1} to itself. Suppose further that Ω_s is strictly convex for some $s \in \{1, 2, \dots, m\}$ and $\Omega_s \cap \Omega_{m+1} = \emptyset$, and assume also that $\lambda \in (0, 1]$ and $\gamma \in (0, 2)$. Then $G_{\lambda,\gamma}$ is Edelstein-contractive over Ω_{m+1} , has at most one fixed point in X , and this fixed point - if it exists - is in Ω_{m+1} . If, in addition, either Ω_{m+1} is bounded or Ω_{m+1} is unbounded and Ω_q is bounded for some $q \in \{1, 2, \dots, m\}$, then $G_{\lambda,\gamma}$ has a unique fixed point in X and this fixed point is in Ω_{m+1} .*

Proof. Let $T := \sum_{i=1}^m \beta_i P_i$ be the weighted sum operator of the operators P_i , $i \in \{1, 2, \dots, m\}$, and let $T_\gamma := (1 - \gamma)\text{Id} + \gamma T$ be its γ -relaxation whenever $\gamma \in [0, 2]$.

The triangle inequality, the definition of $G_{\lambda,\gamma}$ (see (3.6)), the nonexpansivity of P_{m+1} , the fact that $(\beta_i)_{i=1}^m$ is a positive weight vector and the fact that $\lambda \in [0, 1]$, imply, for each $x \in X$ and $y \in X$, that

$$\begin{aligned}
\|G_{\lambda,\gamma}(x) - G_{\lambda,\gamma}(y)\| &= \left\| \left((1-\lambda)x + \lambda P_{m+1}((1-\gamma)x + \gamma \sum_{i=1}^m \beta_i P_i(x)) \right) \right. \\
&\quad \left. - \left((1-\lambda)y + \lambda P_{m+1}((1-\gamma)y + \gamma \sum_{i=1}^m \beta_i P_i(y)) \right) \right\| \\
&\leq (1-\lambda)\|x - y\| + \lambda \|P_{m+1}((1-\gamma)x + \gamma \sum_{i=1}^m \beta_i P_i(x)) - P_{m+1}((1-\gamma)y + \gamma \sum_{i=1}^m \beta_i P_i(y))\| \\
&\leq (1-\lambda)\|x - y\| + \lambda \|((1-\gamma)x + \gamma \sum_{i=1}^m \beta_i P_i(x)) - ((1-\gamma)y + \gamma \sum_{i=1}^m \beta_i P_i(y))\| \\
&= (1-\lambda)\|x - y\| + \lambda \|T_\gamma(x) - T_\gamma(y)\|. \quad (4.3)
\end{aligned}$$

Since the orthogonal projection onto a nonempty, closed and convex subset is firmly nonexpansive [18, Theorem 2.2.21(iii) on p. 76 and Theorem 2.2.10(v) on p. 70], and since a positively weighted sum (namely, with a positive weight vector) of firmly nonexpansive operators is also firmly nonexpansive [18, Corollary 2.2.20, p. 75], we conclude that the operator $T = \sum_{i=1}^m \beta_i P_i$ is firmly nonexpansive. Hence [18, Theorem 2.2.20(ii)] implies that its relaxation T_γ is nonexpansive for all $\gamma \in [0, 2]$. Thus from (4.3) and the fact that $\lambda \in [0, 1]$ we conclude that $G_{\lambda,\gamma}$ is nonexpansive.

Now let $x \in \Omega_{m+1}$ be arbitrary. The definition of $G_{\lambda,\gamma}$ (see (3.6)), the assumption that $\lambda \in [0, 1]$, the fact that $P_{m+1}(z) \in \Omega_{m+1}$ for every $z \in X$, and the convexity of Ω_{m+1} , all imply that $G_{\lambda,\gamma}(x)$ is a convex combination of x and $P_{m+1}(z) \in \Omega_{m+1}$ for $z := (1-\gamma)x + \gamma \sum_{i=1}^m \beta_i P_i(x)$. Hence $G_{\lambda,\gamma}(x) \in \Omega_{m+1}$ for each $x \in \Omega_{m+1}$ and therefore $G_{\lambda,\gamma}$ maps Ω_{m+1} to itself.

Now we want to show that under certain assumptions $G_{\lambda,\gamma}$ is Edelstein-contractive over Ω_{m+1} . Before recalling these assumptions, we observe that from the calculation presented in [18, Top of p. 71] (where we take γ instead of λ there) we have for all $\gamma \in [0, 2]$ and all $(x, y) \in X^2$,

$$\|T_\gamma(x) - T_\gamma(y)\|^2 \leq \gamma(2-\gamma)\|T(x) - T(y)\|\|x - y\| + (1-\gamma)^2\|x - y\|^2. \quad (4.4)$$

(In fact, the calculation done in [18, Top of p. 71] and also (4.4) hold for any firmly nonexpansive operator T and not just for $T = \sum_{i=1}^m \beta_i P_i$.) Therefore (4.4), the definition of T , the fact that $(\beta_i)_{i=1}^m$ is a weight vector, and the triangle inequality, all imply that

$$\begin{aligned}
\|T_\gamma(x) - T_\gamma(y)\|^2 &\leq \gamma(2-\gamma) \left\| \sum_{i=1}^m \beta_i (P_i(x) - P_i(y)) \right\| \cdot \|x - y\| + (1-\gamma)^2\|x - y\|^2 \\
&\leq \gamma(2-\gamma)\|x - y\| \sum_{i=1}^m \beta_i \|P_i(x) - P_i(y)\| + (1-\gamma)^2\|x - y\|^2. \quad (4.5)
\end{aligned}$$

Now assume that Ω_s is strictly convex for some $s \in \{1, 2, \dots, m\}$, that $\Omega_s \cap \Omega_{m+1} = \emptyset$, that $\lambda \in (0, 1]$ and that $\gamma \in (0, 2)$. Take any pair of points $(x, y) \in \Omega_{m+1}^2$ such that $x \neq y$. Then both $x \notin \Omega_s$ and $y \notin \Omega_s$, and hence Lemma 4.4 (applied to $S := \Omega_s$), the nonexpansivity of the orthogonal projections P_i for every $i \in \{1, 2, \dots, m\}$, the fact that $(\beta_i)_{i=1}^m$ is a positive weight vector and hence $\beta_s > 0$, and the assumptions that Ω_s is strictly convex, $x \neq y$, $\gamma \in (0, 2)$ (which implies that $\gamma(2 - \gamma) > 0$) and (4.5), all imply that

$$\begin{aligned}
\|T_\gamma(x) - T_\gamma(y)\|^2 &\leq \gamma(2 - \gamma)\|x - y\| \sum_{i=1}^m \beta_i \|P_i(x) - P_i(y)\| + (1 - \gamma)^2 \|x - y\|^2 \\
&= \gamma(2 - \gamma)\|x - y\| \sum_{i \neq s} \beta_i \|P_i(x) - P_i(y)\| \\
&\quad + \gamma(2 - \gamma)\beta_s \|x - y\| \|P_s(x) - P_s(y)\| + (1 - \gamma)^2 \|x - y\|^2 \\
&< \gamma(2 - \gamma)\|x - y\| \sum_{i \neq s} \beta_i \|x - y\| + \gamma(2 - \gamma)\beta_s \|x - y\|^2 \quad (4.6) \\
&\quad + (1 - \gamma)^2 \|x - y\|^2 \\
&= \gamma(2 - \gamma)\|x - y\|^2 \sum_{i=1}^m \beta_i + (1 - \gamma)^2 \|x - y\|^2 \\
&= \|x - y\|^2.
\end{aligned}$$

Thus $\|T_\gamma(x) - T_\gamma(y)\| < \|x - y\|$, and when we combine this inequality with (4.3) and the assumption that $\lambda > 0$, we see that $G_{\lambda, \gamma}$ is Edelstein contractive under the assumptions imposed before (4.6).

As for the fixed points of $G_{\lambda, \gamma}$ over X , Lemma 4.6 tells us that any fixed point of $G_{\lambda, \gamma}$, if it exists, belongs to Ω_{m+1} . Now, if x and y are two distinct fixed points of $G_{\lambda, \gamma}$ in X , then they belong to Ω_{m+1} and $\|x - y\| = \|G_{\lambda, \gamma}(x) - G_{\lambda, \gamma}(y)\| < \|x - y\|$ since $G_{\lambda, \gamma}$ is Edelstein-contractive over Ω_{m+1} as we showed above. This contradiction shows that $G_{\lambda, \gamma}$ has at most one fixed point in Ω_{m+1} and hence (Lemma 4.6) at most one fixed point in X .

It remains to consider the issue of existence of a fixed point of $G_{\lambda, \gamma}$ under the further assumption that either Ω_{m+1} is bounded or Ω_{m+1} is unbounded but Ω_q is bounded for some $q \in \{1, 2, \dots, m\}$. This is an immediate consequence of Lemma 4.5 and Lemma 4.6. $\square$

The next lemma is recent and is essential for the convergence analysis.

Lemma 4.8. *(a particular case of [22, Lemma 27], with a different notation) Under the conditions of Algorithm 3.3, suppose that X is finite-dimensional. Then for each $i \in \{1, 2, \dots, m\}$ and each $\rho > 0$ the sequence $(\hat{P}_{i,k})_{k=0}^\infty$ converges uniformly, over $B[0, \rho]$, to the orthogonal projection P_i onto Ω_i .*

We finish this section with the following well-known lemma whose proof is just an exercise: see [65, Ex. 8, p. 134].

Lemma 4.9. *Suppose that $(X_1, \mathcal{T}_1)$ is a topological space and (X_2, d_2) is a metric space. Assume also that $(g_k)_{k=0}^\infty$ is a sequence of continuous functions from X_1 to X_2 which converges uniformly to a function $g : X_1 \rightarrow X_2$. Suppose also that $(z^k)_{k=0}^\infty$ is a sequence in X_1 which converges to $z \in X_1$. Then $\lim_{k \rightarrow \infty} g_k(z^k) = g(z)$.*

5. EXISTENCE, UNIQUENESS AND CONVERGENCE ANALYSIS FOR THE BAT ALGORITHM

In this section we present the analysis needed for proving our main existence, uniqueness and convergence result, namely, Theorem 3.5 formulated above. The theorem itself is proved at the end of the section.

Lemma 5.1. *In the setting of Algorithm 3.3, for each $k \in \mathbb{N} \cup \{0\}$ the function G_k maps Ω_{m+1} into itself. In particular, if $x^0 \in \Omega_{m+1} \subseteq B[0, \rho]$ for some $\rho > 0$, then any sequence $(x^k)_{k \in \mathbb{N} \cup \{0\}}$ generated by Algorithm 3.3 is contained in Ω_{m+1} and hence bounded.*

Proof. Fix $k \in \mathbb{N} \cup \{0\}$, let $x \in \Omega_{m+1}$ and let $y := (1 - \gamma_k)x + \sum_{i=1}^m \beta_i \hat{P}_{i,k}(x) \in X$. Then $P_{m+1}(y) \in \Omega_{m+1}$ and hence (3.4) implies that $G_k(x) = (1 - \lambda_k)x + \lambda_k P_{m+1}(y) \in \Omega_{m+1}$ as a convex combination of two elements in the convex set Ω_{m+1} . Thus, G_k maps Ω_{m+1} into itself for each $k \in \mathbb{N} \cup \{0\}$. Now assume that $x^0 \in \Omega_{m+1}$. Then $x^1 = G_0(x^0) \in \Omega_{m+1}$, by (3.5) and by what we have just proved. Thus, induction, as well as (3.5), show that $x^k \in \Omega_{m+1}$ for each $k \in \mathbb{N} \cup \{0\}$. Since Ω_{m+1} is contained in $B[0, \rho]$, we conclude that $(x^k)_{k \in \mathbb{N} \cup \{0\}}$ is bounded. $\square$

Lemma 5.2. *In the setting of Algorithm 3.3, suppose that X is finite-dimensional. If there is an infinite subset K of $\mathbb{N} \cup \{0\}$ such that the sequences of scalars $(\lambda_k)_{k \in K}$ and $(\gamma_k)_{k \in K}$ converge to the real numbers $\lambda \in [0, 1]$ and $\gamma \in [0, 1]$, respectively, and if there is $\rho > 0$ such that $\Omega_i \subseteq B[0, \rho]$ for all $i \in \{1, 2, \dots, m\}$, then the sequence of operators $(G_k)_{k \in K}$ defined by (3.4) for all $k \in K$ converges uniformly to $G_{\lambda, \gamma} := (1 - \lambda)Id + \lambda P_{m+1}((1 - \gamma)Id + \gamma \sum_{i=1}^m \beta_i P_i)$ over $B[0, \rho]$.*

Proof. Let $\varepsilon > 0$ be arbitrary. Since P_i is nonexpansive for all $i \in \{1, 2, \dots, m+1\}$, these operators are bounded on the bounded set $B[0, \rho]$. Hence there is a positive number M such that $\|P_i(x)\| < M$ for all $x \in B[0, \rho]$ and all $i \in \{1, 2, \dots, m+1\}$.

Since $\lambda = \lim_{k \in K} \lambda_k$ and $\gamma = \lim_{k \in K} \gamma_k$, there is $k_0 \in K$ such that $|\lambda_k - \lambda| < (1/6)\varepsilon/(\rho + M)$ and $|\gamma_k - \gamma| < (1/6)\varepsilon/(\rho + M)$ for every $k_0 \leq k \in K$. Because $(\tau_{i,k})_{k=0}^\infty$ is a sequence of steering parameters for each $i \in \{1, \dots, m+1\}$ and because X is finite-dimensional, it follows from Lemma 4.8 that for each $i \in \{1, \dots, m\}$ the sequence $(\hat{P}_{i,k})_{k=0}^\infty$ converges uniformly, over $B[0, \rho]$, to the orthogonal projection P_i onto Ω_i . Thus, its subsequence $(\hat{P}_{i,k})_{k \in K}$ also converges uniformly, over $B[0, \rho]$, to P_i for each $i \in \{1, \dots, m\}$, and hence there is $k_0 \leq k_1 \in K$ such that $\|\hat{P}_{i,k}(z) - P_i(z)\| < (1/6)\varepsilon$ for all $k_1 \leq k \in K$ and all $i \in \{1, 2, \dots, m\}$ and all $z \in B[0, \rho]$.

These facts, as well as the nonexpansiveness of P_{m+1} , the fact that $\lambda_k \in [0, 1]$ and $\gamma_k \in [0, 1]$ for all $k \in K$, the fact that $(1 - \gamma)x + \gamma \sum_{i=1}^m \beta_i P_i(x) \in B[0, \rho]$ for each $x \in B[0, \rho]$ (due to the convexity of $B[0, \rho]$, the inclusions $\Omega_i \subseteq B[0, \rho]$, $i \in \{1, 2, \dots, m\}$, because $\gamma \in [0, 1]$ and because $(\beta_i)_{i=1}^m$ is a weight vector), and simple algebra, all imply

that for every $x \in B[0, \rho]$ and every $k_1 \leq k \in K$,

$$\begin{aligned}
\|G_k(x) - G_{\lambda, \gamma}(x)\| &= \|((1 - \lambda_k)x + \lambda_k P_{m+1}((1 - \gamma_k)x + \gamma_k \sum_{i=1}^m \beta_i \hat{P}_{i,k}(x))) \\
&\quad - ((1 - \lambda)x + \lambda P_{m+1}((1 - \gamma)x + \gamma \sum_{i=1}^m \beta_i P_i(x)))\| \\
&\leq |\lambda_k - \lambda| \|x\| + \lambda_k \|P_{m+1}((1 - \gamma_k)x + \gamma_k \sum_{i=1}^m \beta_i \hat{P}_{i,k}(x)) - P_{m+1}((1 - \gamma)x + \gamma \sum_{i=1}^m \beta_i P_i(x))\| \\
&\quad + \|\lambda_k P_{m+1}((1 - \gamma)x + \gamma \sum_{i=1}^m \beta_i P_i(x)) - \lambda P_{m+1}((1 - \gamma)x + \gamma \sum_{i=1}^m \beta_i P_i(x))\| \\
&\leq |\lambda_k - \lambda| \|x\| + \lambda_k \|((1 - \gamma_k)x + \gamma_k \sum_{i=1}^m \beta_i \hat{P}_{i,k}(x)) - ((1 - \gamma)x + \gamma \sum_{i=1}^m \beta_i P_i(x))\| + \\
&\quad \|\lambda_k P_{m+1}((1 - \gamma)x + \gamma \sum_{i=1}^m \beta_i P_i(x)) - \lambda P_{m+1}((1 - \gamma)x + \gamma \sum_{i=1}^m \beta_i P_i(x))\| \\
&\leq \frac{(1/6)\varepsilon\rho}{\rho + M} + \|\gamma_k - \gamma\| \|x\| + \|\gamma_k \sum_{i=1}^m \beta_i \hat{P}_{i,k}(x) - \gamma \sum_{i=1}^m \beta_i P_i(x)\| + \frac{(1/6)\varepsilon M}{\rho + M} \\
&\leq \frac{1}{6}\varepsilon + \frac{(1/6)\varepsilon\rho}{\rho + M} + \gamma_k \|\sum_{i=1}^m \beta_i \hat{P}_{i,k}(x) - \sum_{i=1}^m \beta_i P_i(x)\| + |\gamma_k - \gamma| \sum_{i=1}^m \beta_i \|P_i(x)\| + \frac{1}{6}\varepsilon \\
&\leq \frac{3}{6}\varepsilon + \sum_{i=1}^m \beta_i \|\hat{P}_{i,k}(x) - P_i(x)\| + \frac{(1/6)\varepsilon M}{\rho + M} < \frac{3}{6}\varepsilon + \frac{1}{6}\varepsilon + \frac{1}{6}\varepsilon < \varepsilon.
\end{aligned}$$

Since ε was an arbitrary positive number we conclude that $G_{\lambda, \gamma} = \lim_{k \rightarrow \infty, k \in K} G_k$ uniformly on $B[0, \rho]$, as claimed. $\square$

Lemma 5.3. *In the setting of Algorithm 3.3, if either Ω_{m+1} is bounded or Ω_{m+1} is unbounded and Ω_q is bounded for some $q \in \{1, 2, \dots, m\}$, then (3.1) has at least one solution. If, in addition, Ω_s is strictly convex for some $s \in \{1, 2, \dots, m\}$ and also $\Omega_s \cap \Omega_{m+1} = \emptyset$, then (3.1) has a unique solution x^* , and x^* has the property that it is the unique fixed point of each of the functions $G_{\lambda, \gamma} : X \rightarrow X$, $\lambda \in (0, 1]$, $\gamma \in (0, 2)$ defined in (3.6).*

Proof. Existence of a solution to (3.1) in Ω_{m+1} when either Ω_{m+1} is bounded or when it is unbounded and Ω_q is bounded for some $q \in \{1, 2, \dots, m\}$ follows from Lemma 4.5. Uniqueness of the solution to (3.1) when there is some $s \in \{1, 2, \dots, m\}$ such that Ω_s is strictly convex and $\Omega_s \cap \Omega_{m+1} = \emptyset$, as well as the fact that this unique solution is a fixed point of $G_{\lambda, \gamma}$ for each $\lambda \in (0, 1]$ and $\gamma \in (0, 2)$, follow from the combination of Lemma 4.6 and Lemma 4.7. $\square$

Lemma 5.4. *Suppose that all the conditions mentioned in Theorem 3.5 hold. Then there exists an infinite subset K of $\mathbb{N} \cup \{0\}$ and a subsequence $(x^k)_{k \in K}$ of $(x^k)_{k=0}^\infty$ which converges to x^* . Moreover, $\lim_{k \rightarrow \infty, k \in K} \|x^{k+1} - x^k\| = 0$.*

Proof. Lemma 5.3 ensures the existence and uniqueness of the solution x^* to (3.1), as well as the fact that it is the unique fixed point of $G_{\lambda, \gamma}$ for each of the functions $G_{\lambda, \gamma} : X \rightarrow X$, $\lambda \in (0, 1]$, $\gamma \in (0, 2)$ defined in (3.6).

Now assume for a contradiction that no subsequence of $(x^k)_{k=0}^\infty$ converges to x^* . Then there is an $\varepsilon > 0$ such that only finitely many elements of the sequence $(x^k)_{k=0}^\infty$ are in the open ball $B(x^*, \varepsilon)$, and so there is $k_0 \in \mathbb{N} \cup \{0\}$ such that $x^k \notin B(x^*, \varepsilon)$ for all $k_0 \leq k \in \mathbb{N} \cup \{0\}$. We claim that this implies that $\liminf_{k \rightarrow \infty} \|x^{k+1} - x^k\| > 0$.

Indeed, if this is false, then this assumption, as well as the fact that the sequence $(\|x^{k+1} - x^k\|_{k \in \mathbb{N} \cup \{0\}})$ is a sequence of nonnegative numbers, imply that there is an infinite set N_1 of $\{k_0, k_0 + 1, k_0 + 2, \dots\}$ such that $\lim_{k \rightarrow \infty, k \in N_1} \|x^{k+1} - x^k\| = 0$. Since $(x^k)_{k \in N_1}$ is contained in the bounded set Ω_{m+1} , it has a convergent subsequence $(x^k)_{k \in N_2}$ (where N_2 is an infinite subset of N_1) whose limit, which we denote by $x^{\infty, 2}$, is in the closed set Ω_{m+1} .

Knowing that $x^{k+1} = G_k(x^k)$ for every $k \in \mathbb{N} \cup \{0\}$ (see (3.5)), we conclude from Lemma 4.9 and Lemma 5.2 (with $K := \mathbb{N} \cup \{0\}$ there) that

$$\lim_{k \rightarrow \infty, k \in N_2} \|x^{k+1} - x^k\| = \lim_{k \rightarrow \infty, k \in N_2} \|G_k(x^k) - x^k\| = \|G_{\lambda_\infty, \gamma_\infty}(x^{\infty, 2}) - x^{\infty, 2}\|,$$

where $0 < \lambda_\infty = \lim_{k \rightarrow \infty} \lambda_k$ and $0 < \gamma_\infty = \lim_{k \rightarrow \infty} \gamma_k$ as we assume in the formulation of Theorem 3.5 (since $\lambda_k \in (0, 1]$ and $\gamma_k \in (0, 1]$ for each $k \in \mathbb{N} \cup \{0\}$, we also have $\lambda_\infty \leq 1$ and $\gamma_\infty \leq 1$). Since $N_2 \subseteq N_1$, we have $\lim_{k \rightarrow \infty, k \in N_2} \|x^{k+1} - x^k\| = \lim_{k \rightarrow \infty, k \in N_1} \|x^{k+1} - x^k\| = 0$. Hence, $\|G_{\lambda_\infty, \gamma_\infty}(x^{\infty, 2}) - x^{\infty, 2}\| = 0$, namely, $x^{\infty, 2}$ is a fixed point of $G_{\lambda_\infty, \gamma_\infty}$.

The operator $G_{\lambda_\infty, \gamma_\infty}$ has a unique fixed point in Ω_{m+1} according to Lemma 4.7, and this fixed point is x^* according to Lemma 4.6. This yields $x^{\infty, 2} = x^*$ which is a contradiction because the fact that $x^k \notin B(x^*, \varepsilon)$ for all $k_0 \leq k \in \mathbb{N} \cup \{0\}$ implies that $\|x^k - x^*\| \geq \varepsilon$ for all $k_0 \leq k \in \mathbb{N} \cup \{0\}$, and hence $\varepsilon \leq \lim_{k \rightarrow \infty, k \in N_2} \|x^k - x^*\| = \|x^{\infty, 2} - x^*\|$. Consequently, $\liminf_{k \rightarrow \infty} \|x^{k+1} - x^k\| > 0$, as claimed, and hence there is a nonnegative integer $k_1 \geq k_0$ such that $\|x^{k+1} - x^k\| \geq \liminf_{\ell \rightarrow \infty} \|x^{\ell+1} - x^\ell\| > 0$ for all $k_1 \leq k \in \mathbb{N} \cup \{0\}$.

Denote $r := \inf\{\|x^{k+1} - x^k\| \mid k_1 \leq k \in \mathbb{N} \cup \{0\}\}$. According to what we showed in the previous paragraph $r > 0$. Additionally, the definition of r implies that there is an infinite subset N_3 of $\{k_1, k_1 + 1, k_1 + 2, k_1 + 3, \dots\}$ such that $r = \lim_{k \rightarrow \infty, k \in N_3} \|x^{k+1} - x^k\|$. The sequence $(x^k)_{k \in N_3}$ is contained in the compact set Ω_{m+1} and hence has an accumulation point $x^{\infty, 3} \in \Omega_{m+1}$, namely $x^{\infty, 3} = \lim_{k \rightarrow \infty, k \in N_4} x^k$ for some infinite subset N_4 of N_3 . Lemma 4.9 and Lemma 5.2 imply that

$$r = \lim_{k \rightarrow \infty, k \in N_4} \|x^{k+1} - x^k\| = \lim_{k \rightarrow \infty, k \in N_4} \|G_k(x^k) - x^k\| = \|G_{\lambda_\infty, \gamma_\infty}(x^{\infty, 3}) - x^{\infty, 3}\|. \quad (5.1)$$

Since $r > 0$, we conclude that $x^{\infty, 3}$ is not a fixed point of $G_{\lambda_\infty, \gamma_\infty}$. In addition, since $G_{\lambda_\infty, \gamma_\infty}(x^{\infty, 3}) = \lim_{k \rightarrow \infty, k \in N_4} G_k(x^k)$ as follows from Lemma 4.9 and Lemma 5.2 (with $K := \mathbb{N} \cup \{0\}$), another application of these lemmata (again with $K := \mathbb{N} \cup \{0\}$) implies

that

$$\|x^{k+2} - x^{k+1}\| = \|G_{k+1}(G_k(x^k)) - G_k(x^k)\| \xrightarrow{k \rightarrow \infty, k \in N_4} \|G_{\lambda_\infty, \gamma_\infty}^2(x^{\infty,3}) - G_{\lambda_\infty, \gamma_\infty}(x^{\infty,3})\|.$$

On the one hand, the definition of r implies that $r \leq \|x^{k+2} - x^{k+1}\|$ for all $k_1 \leq k \in \mathbb{N} \cup \{0\}$, and, in particular, $r \leq \|G_{\lambda_\infty, \gamma_\infty}^2(x^{\infty,3}) - G_{\lambda_\infty, \gamma_\infty}(x^{\infty,3})\|$. On the other hand the fact that $G_{\lambda_\infty, \gamma_\infty}$ is Edelstein-contractive on Ω_{m+1} according to Lemma 4.7, together with (5.1) and the facts that $G_{\lambda_\infty, \gamma_\infty}(x^{\infty,3}) \neq x^{\infty,3}$, $x^{\infty,3} \in \Omega_{m+1}$ and $G_{\lambda_\infty, \gamma_\infty}(x^{\infty,3}) \in \Omega_{m+1}$ (since Lemma 4.7 implies that $G_{\lambda_\infty, \gamma_\infty}$ maps Ω_{m+1} into itself), all imply that $\|G_{\lambda_\infty, \gamma_\infty}^2(x^{\infty,3}) - G_{\lambda_\infty, \gamma_\infty}(x^{\infty,3})\| < \|G_{\lambda_\infty, \gamma_\infty}(x^{\infty,3}) - x^{\infty,3}\| = r$. We conclude that $r \leq \|G_{\lambda_\infty, \gamma_\infty}^2(x^{\infty,3}) - G_{\lambda_\infty, \gamma_\infty}(x^{\infty,3})\| < r$, a contradiction.

The previous paragraph shows that the initial assumption that there does not exist a subsequence of $(x^k)_{k=0}^\infty$ which converges to x^* cannot hold, namely there exists a subsequence of $(x^k)_{k=0}^\infty$ which converges to x^* . Thus, there exists an infinite subset K of $\mathbb{N} \cup \{0\}$ such that $\lim_{k \rightarrow \infty, k \in K} x^k = x^*$, and therefore Lemma 4.9 and Lemma 5.2, as well as Lemma 4.6, imply that

$$\lim_{k \rightarrow \infty, k \in K} \|x^{k+1} - x^k\| = \lim_{k \rightarrow \infty, k \in K} \|G_k(x^k) - x^k\| = \|G_{\lambda_\infty, \gamma_\infty}(x^*) - x^*\| = 0.$$

□

Now we are ready to present the proof of Theorem 3.5.

Proof of Theorem 3.5. Lemma 5.3 addresses the issues of existence of a solution to (3.1) under the conditions mentioned in the formulation of Theorem 3.5, its uniqueness under further assumptions, as well as the fact that it is the unique fixed point of each of the functions $G_{\lambda, \gamma} : X \rightarrow X$, $\lambda \in (0, 1]$, $\gamma \in (0, 2)$ defined in (3.6).

It remains to show that $(x^k)_{k=0}^\infty$ converges to the unique solution $x^* \in \Omega_{m+1}$ to (3.1) when the following assumptions are imposed: $x^0 \in \Omega_{m+1}$, the space X is finite-dimensional, there is some $\rho > 0$ such that $\Omega_{m+1} \cup (\cup_{i=1}^m \cup_{j=1}^{m_i} \Omega_{i,j}) \subseteq B[0, \rho]$, Ω_i is strictly convex and satisfies $\Omega_i \cap \Omega_{m+1} = \emptyset$ for each $i \in \{1, 2, \dots, m\}$, and $\lambda_\infty := \lim_{k \rightarrow \infty} \lambda_k$ and $\gamma_\infty := \lim_{k \rightarrow \infty} \gamma_k$ exist and are positive (and hence they satisfy $\lambda_\infty \leq 1$ and $\gamma_\infty \leq 1$ since $\lambda_k \in (0, 1]$ and $\gamma_k \in (0, 1]$ for every $k \in \mathbb{N} \cup \{0\}$).

Since Ω_{m+1} is closed and bounded and X is finite-dimensional, both Ω_{m+1} and Ω_{m+1}^2 are compact metric spaces, where the metrics are induced by the norm. Now let $\varepsilon > 0$ be arbitrary. Denote

$$A_{0.5\varepsilon} := \{(x, y) \in \Omega_{m+1}^2 \mid \|x - y\| \geq 0.5\varepsilon\}. \quad (5.2)$$

This is a closed set of the compact metric space Ω_{m+1}^2 and hence a compact metric space. Therefore, for all $i \in \{1, 2, \dots, m\}$ the continuous function $f_i : A_{0.5\varepsilon} \rightarrow \mathbb{R}$, defined by $f_i(x, y) := \|x - y\| - \|P_i(x) - P_i(y)\|$ for all $(x, y) \in A_{0.5\varepsilon}$, attains a minimum over $A_{0.5\varepsilon}$ at some point $(\tilde{x}(i), \tilde{y}(i))$.

Since $\Omega_i \cap \Omega_{m+1} = \emptyset$ and Ω_i is strictly convex for every $i \in \{1, 2, \dots, m\}$, Lemma 4.4 implies that P_i is Edelstein-contractive on Ω_{m+1} for all $i \in \{1, 2, \dots, m\}$. Since $\|\tilde{x}(i) - \tilde{y}(i)\| \geq 0.5\varepsilon$ by the definition of $A_{0.5\varepsilon}$, it follows that $\tilde{x}(i) \neq \tilde{y}(i)$ and, hence, $\|P_i(\tilde{x}(i)) - P_i(\tilde{y}(i))\| < \|\tilde{x}(i) - \tilde{y}(i)\|$. Therefore,

$$0 < \|\tilde{x}(i) - \tilde{y}(i)\| - \|P_i(\tilde{x}(i)) - P_i(\tilde{y}(i))\| = f_i(\tilde{x}(i), \tilde{y}(i)) \leq \|x - y\| - \|P_i(x) - P_i(y)\| \quad (5.3)$$

for all $(x, y) \in A_{0.5\varepsilon}$, and hence $r_\varepsilon := \min_{i \in \{1, 2, \dots, m\}} \{f_i(\tilde{x}(i), \tilde{y}(i)), 0.5\varepsilon\}$ satisfies both $r_\varepsilon \leq 0.5\varepsilon$ and

$$0 < r_\varepsilon \leq \|x - y\| - \|P_i(x) - P_i(y)\|, \quad \forall (x, y) \in A_{0.5\varepsilon} \text{ and } \forall i \in \{1, \dots, m\}. \quad (5.4)$$

We know from Lemma 4.8 that for each $i \in \{1, 2, \dots, m\}$ the sequence $(\hat{P}_{i,k})_{k=0}^\infty$ converges uniformly, over $B[0, \rho]$, to the orthogonal projection P_i onto Ω_i . Hence, in particular, $(\hat{P}_{i,k})_{k=0}^\infty$ converges uniformly to P_i over Ω_{m+1} and, therefore, there is $k_0 \in \mathbb{N} \cup \{0\}$ such that $\|\hat{P}_{i,k}(x) - P_i(x)\| < r_\varepsilon$ for all $x \in \Omega_{m+1}$, all $i \in \{1, 2, \dots, m\}$ and all $k_0 \leq k \in \mathbb{N} \cup \{0\}$.

Lemma 5.4 implies that there is an infinite subset K of $\mathbb{N} \cup \{0\}$ such that $x^* = \lim_{k \rightarrow \infty, k \in K} x^k$. Thus there is $k_0 \leq k_1 \in K$ such that $x^{k_1} \in B(x^*, \varepsilon)$. We show by induction that $x^k \in B(x^*, \varepsilon)$ for all $k_1 \leq k \in \mathbb{N} \cup \{0\}$.

Indeed, for $k = k_1$ this is the induction hypothesis. Assume now that the assertion is true up to $k_1 \leq k \in \mathbb{N} \cup \{0\}$. We need to show that it holds also for $k + 1$. Due to $x^* = G_{\lambda_\infty, \gamma_\infty}(x^*) = (1 - \lambda_\infty)x^* + \lambda_\infty P_{m+1}((1 - \gamma_\infty)x^* + \gamma_\infty \sum_{i=1}^m P_i(x^*))$ according to Lemma 4.6, we conclude from (3.5), simple arithmetic, the triangle inequality, the nonexpansivity of the orthogonal projections P_i (for each $i \in \{1, 2, \dots, m+1\}$), from the fact that $(\lambda_\infty, \gamma_\infty) \in (0, 1]^2$, from the fact that $\beta = (\beta_i)_{i=1}^m$ is a positive weight vector, from the fact that $x^* \in \Omega_{m+1}$, and from the fact that $x^k \in \Omega_{m+1}$ (according to Lemma 5.1) and hence $\|\hat{P}_{i,k}(x^k) - P_i(x^k)\| < r_\varepsilon$ (since $k_0 \leq k_1 \leq k$), that

$$\begin{aligned} \|x^{k+1} - x^*\| &= \|((1 - \lambda_\infty)x^k + \lambda_\infty P_{m+1}((1 - \gamma_\infty)x^k + \gamma_\infty \sum_{i=1}^m \hat{P}_{i,k}(x^k))) \\ &\quad - ((1 - \lambda_\infty)x^* + \lambda_\infty P_{m+1}((1 - \gamma_\infty)x^* + \gamma_\infty \sum_{i=1}^m P_i(x^*)))\| \\ &\leq (1 - \lambda_\infty)\|x^k - x^*\| + \lambda_\infty\|P_{m+1}((1 - \gamma_\infty)x^k + \gamma_\infty \sum_{i=1}^m \beta_i \hat{P}_{i,k}(x^k)) \\ &\quad - P_{m+1}((1 - \gamma_\infty)x^* + \gamma_\infty \sum_{i=1}^m \beta_i P_i(x^*))\| \\ &\leq (1 - \lambda_\infty)\varepsilon + \lambda_\infty((1 - \gamma_\infty)\|x^k - x^*\| + \gamma_\infty \sum_{i=1}^m \beta_i \|\hat{P}_{i,k}(x^k) - P_i(x^*)\|) \\ &\leq (1 - \lambda_\infty)\varepsilon + \lambda_\infty(1 - \gamma_\infty)\|x^k - x^*\| \\ &\quad + \lambda_\infty \gamma_\infty \sum_{i=1}^m \beta_i (\|\hat{P}_{i,k}(x^k) - P_i(x^k)\| + \|P_i(x^k) - P_i(x^*)\|) \\ &\leq (1 - \lambda_\infty)\varepsilon + \lambda_\infty(1 - \gamma_\infty)\varepsilon + \lambda_\infty \gamma_\infty \sum_{i=1}^m \beta_i (r_\varepsilon + \|P_i(x^k) - P_i(x^*)\|). \end{aligned} \quad (5.5)$$

We claim that $r_\varepsilon + \|P_i(x^k) - P_i(x^*)\| < \varepsilon$ for all $i \in \{1, 2, \dots, m\}$. Indeed, if we are in the case $\|x^k - x^*\| < 0.5\varepsilon$, then the nonexpansivity of P_i (for all $i \in \{1, 2, \dots, m\}$) and the fact that $r_\varepsilon \leq 0.5\varepsilon$ imply that $r_\varepsilon + \|P_i(x^k) - P_i(x^*)\| \leq 0.5\varepsilon + \|x^k - x^*\| < 0.5\varepsilon + 0.5\varepsilon = \varepsilon$. If we are in the case $\|x^k - x^*\| \geq 0.5\varepsilon$, then $(x^k, x^*) \in A_{0.5\varepsilon}$ (see (5.2)). This fact, as well as the fact that $\|x^k - x^*\| < \varepsilon$ and the fact that $r_\varepsilon \leq \|x^k - x^*\| - \|P_i(x^k) - P_i(x^*)\|$ for all $i \in \{1, 2, \dots, m\}$ (see (5.4)), imply that $r_\varepsilon < \varepsilon - \|P_i(x^k) - P_i(x^*)\|$ for all $i \in \{1, 2, \dots, m\}$. Hence $r_\varepsilon + \|P_i(x^k) - P_i(x^*)\| < \varepsilon$ for all $i \in \{1, 2, \dots, m\}$ in this case too.

Therefore, (5.5), as well as the facts that $(\beta_i)_{i=1}^m$ is a positive weight vector and both λ_∞ and γ_∞ are positive, imply that

$$\|x^{k+1} - x^*\| < (1 - \lambda_\infty)\varepsilon + \lambda_\infty(1 - \gamma_\infty)\varepsilon + \lambda_\infty\gamma_\infty \sum_{i=1}^m \beta_i \varepsilon = \varepsilon,$$

namely $x^{k+1} \in B(x^*, \varepsilon)$. We conclude by induction that $x^k \in B(x^*, \varepsilon)$ for all $k_1 \leq k \in \mathbb{N} \cup \{0\}$, and since ε was an arbitrary positive number it follows that $(x^k)_{k=0}^\infty$ converges to the unique solution x^* of (3.1), as required. $\square$

Remark 5.5. A careful examination of the proof of Theorem 3.5 and the lemmata on which it is based, shows that for each $i \in \{1, 2, \dots, m\}$ if the sequence of operators $(\hat{P}_{i,k})_{k=0}^\infty$ which is defined in (3.3) is replaced by any sequence of operators which converges uniformly to P_i over $B[0, \rho]$, then the conclusion of Theorem 3.5 still holds.

Acknowledgements: This work was supported by U.S. National Institutes of Health (NIH) Grant Number R01CA266467 and by the Cooperation Program in Cancer Research of the German Cancer Research Center (DKFZ) and Israel's Ministry of Innovation, Science and Technology (MOST).

Data availability. Not applicable.

Ethics/Ethical. Not applicable.

Clinical trial. Not applicable.

Competing Interests. The authors declare no competing financial or non-financial interests.

REFERENCES

- [1] AHARONI, R., CENSOR, Y., AND JIANG, Z. Finding a best approximation pair of points for two polyhedra. *Comput. Optim. Appl.* 71 (2018), 509–523.
- [2] ALWADANI, S., BAUSCHKE, H. H., REVALSKI, J. P., AND WANG, X. The difference vectors for convex sets and a resolution of the geometry conjecture. *Open J. Math. Optim.* 2 (2021), Art. No. 5, 18.
- [3] ATTOUCH, H., PEYPOUQUET, J., AND REDONT, P. A dynamical approach to an inertial forward-backward algorithm for convex minimization. *SIAM J. Optim.* 24, 1 (2014), 232–256.
- [4] BAILLON, J.-B., COMBETTES, P. L., AND COMINETTI, R. There is no variational characterization of the cycles in the method of periodic projections. *J. Funct. Anal.* 262, 1 (2012), 400–408.
- [5] BAUSCHKE, H. H., AND BORWEIN, J. M. On the convergence of von Neumann's alternating projection algorithm for two sets. *Set-Valued Anal.* 1 (1993), 185–212.

- [6] BAUSCHKE, H. H., AND BORWEIN, J. M. Dykstra's alternating projection algorithm for two sets. *J. Approx. Theory* 79 (1994), 418–443.
- [7] BAUSCHKE, H. H., AND BORWEIN, J. M. On projection algorithms for solving convex feasibility problems. *SIAM Rev.* 38 (1996), 367–426.
- [8] BAUSCHKE, H. H., BORWEIN, J. M., AND LEWIS, A. S. The method of cyclic projections for closed convex sets in Hilbert space. *Contemp. Math.* 204 (1997), 1–38.
- [9] BAUSCHKE, H. H., AND COMBETTES, P. L. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2 ed. CMS Books in Mathematics. Springer International Publishing, Cham, Switzerland, 2017.
- [10] BAUSCHKE, H. H., COMBETTES, P. L., AND LUKE, D. R. Finding best approximation pairs relative to two closed convex sets in Hilbert spaces. *J. Approx. Theory* 127 (2004), 178–192.
- [11] BAUSCHKE, H. H., SINGH, S., AND WANG, X. Finding best approximation pairs for two intersections of closed convex sets. *Comput. Optim. Appl.* 81 (2022), 289–308.
- [12] BEHLING, R., BELLO-CRUZ, Y., AND SANTOS, L.-R. Infeasibility and error bound imply finite convergence of alternating projections. *SIAM J. Optim.* 31, 4 (2021), 2863–2892.
- [13] BREZIS, H. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext. Springer, New York, 2011.
- [14] BUTNARIU, D., CENSOR, Y., GURFIL, P., AND HADAR, E. On the behavior of subgradient projections methods for convex feasibility problems in Euclidean spaces. *SIAM J. Optim.* 19 (2008), 786–807.
- [15] BYRNE, C., AND CENSOR, Y. Proximity function minimization using multiple Bregman projections, with applications to split feasibility and Kullback-Leibler distance minimization. *Ann. Oper. Res.* 105 (2001), 77–98.
- [16] CAMERON, S., AND CULLEY, R. Determining the minimum translational distance between two convex polyhedra. In *Proceedings of the 1986 IEEE International Conference on Robotics and Automation* (1986), vol. 3, pp. 591–596.
- [17] CASEIRO, R., FACAS VICENTE, M. A., AND VITÓRIA, J. Projection method and the distance between two linear varieties. *Linear Algebra Appl.* 563 (2019), 446–460.
- [18] CEGIELSKI, A. *Iterative Methods for Fixed Point Problems in Hilbert Spaces*, vol. 2057 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2012.
- [19] CENSOR, Y., BORTFELD, T., MARTIN, B., AND TROFIMOV, A. A unified approach for inversion problems in intensity-modulated radiation therapy. *Phys. Med. Biol.* 51, 10 (2006), 2353.
- [20] CENSOR, Y., AND CEGIELSKI, A. Projection methods: an annotated bibliography of books and reviews. *Optimization* 64 (2015), 2343–2358.
- [21] CENSOR, Y., ELFVING, T., KOPF, N., AND BORTFELD, T. The multiple-sets split feasibility problem and its applications for inverse problems. *Inverse Problems* 21, 6 (2005), 2071–2084.
- [22] CENSOR, Y., MANSOUR, R., AND REEM, D. The alternating simultaneous Halpern-Lions-Wittmann-Bauschke algorithm for finding the best approximation pair for two disjoint intersections of convex sets. *J. Approx. Theory* 301 (2024), Paper No. 106045, 22 pp.
- [23] CENSOR, Y., AND REEM, D. Zero-convex functions, perturbation resilience, and subgradient projections for feasibility-seeking methods. *Math. Prog. (Ser. A)* 152 (2015), 339–380.
- [24] CENSOR, Y., AND ZAKNOON, M. Algorithms and convergence results of projection methods for inconsistent feasibility problems: a review. *Pure Appl. Funct. Anal.* 3 (2018), 565–586.
- [25] CENSOR, Y., ZAKNOON, M., AND ZASLAVSKI, A. J. Data-compatibility of algorithms for constrained convex optimization. *J. Appl. Numer. Optim.* 3, 1 (2021), 21–41.
- [26] CENSOR, Y., AND ZENIOS, A. S. *Parallel Optimization: Theory, Algorithms, and Applications*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 1997. With a foreword by George B. Dantzig. Available on arXiv for free download at: <https://arxiv.org/abs/2609.24505>.

- [27] CHANG, J.-W., CHOI, Y.-K., KIM, M.-S., AND WANG, W. Computation of the minimum distance between two Bézier curves/surfaces. *Comput. Graphics* 35, 3 (2011), 677–684. Shape Modeling International (SMI) Conference 2011.
- [28] CHENEY, W., AND GOLDSTEIN, A. A. Proximity maps for convex sets. *Proc. Amer. Math. Soc.* 10 (1959), 448–450.
- [29] CHI, E. C., AND LANGE, K. A look at the generalized Heron problem through the lens of majorization-minimization. *Amer. Math. Monthly* 121, 2 (2014), 95–108.
- [30] COMBETTES, P. L. Inconsistent signal feasibility problems: least-squares solutions in a product space. *IEEE Trans. Signal Process* 42 (1994), 2955–2966.
- [31] COMBETTES, P. L. *The Convex Feasibility Problem in Image Recovery*, vol. 95 of *Adv. Imaging Electron Phys.* Elsevier, 1996, pp. 155–270.
- [32] COMBETTES, P. L. A block-iterative surrogate constraint splitting method for quadratic signal recovery. *IEEE Trans. Signal Process.* 51, 7 (2003), 1771–1782.
- [33] COMBETTES, P. L. Solving monotone inclusions via compositions of nonexpansive averaged operators. *Optimization* 53, 5-6 (2004), 475–504.
- [34] COMBETTES, P. L. Systems of structured monotone inclusions: duality, algorithms, and applications. *SIAM J. Optim.* 23, 4 (2013), 2420–2447.
- [35] COMBETTES, P. L., AND BONDON, P. Hard-constrained inconsistent signal feasibility problems. *IEEE Trans. Signal Process* 47 (1999), 2460–2468.
- [36] COMBETTES, P. L., AND GLAUDIN, L. E. Proximal activation of smooth functions in splitting algorithms for convex image recovery. *SIAM J. Imaging Sci.* 12, 4 (2019), 1905–1935.
- [37] COMBETTES, P. L., AND PESQUET, J.-C. A proximal decomposition method for solving convex variational inverse problems. *Inverse Problems* 24, 6 (2008), 065014, 27.
- [38] COMBETTES, P. L., AND WAJS, V. R. Signal recovery by proximal forward-backward splitting. *Multiscale Model. Simul.* 4 (2005), 1168–1200.
- [39] COMBETTES, P. L., AND WOODSTOCK, Z. C. A variational inequality model for the construction of signals from inconsistent nonlinear equations. *SIAM J. Imaging Sci.* 15, 1 (2022), 84–109.
- [40] DAX, A. The distance between two convex sets. *Linear Algebra Appl.* 416 (2006), 184–213.
- [41] DE PIERRO, A. R., AND IUSEM, A. N. On the asymptotic behavior of some alternate smoothing series expansion iterative methods. *Linear Algebra Appl.* 130 (1990), 3–24.
- [42] DOBKIN, D. P., AND KIRKPATRICK, D. G. A linear algorithm for determining the separation of convex polyhedra. *J. Algorithms* 6, 3 (1985), 381–392.
- [43] EDELSTEIN, M. On fixed and periodic points under contractive mappings. *J. London Math. Soc.* 37 (1962), 74–79.
- [44] EHMANN, S. A., AND LIN, M. C. Accelerated proximity queries between convex polyhedra by multi-level Voronoi marching. In *Proceedings of the 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2000)* (2000), vol. 3, pp. 2101–2106.
- [45] ELSNER, L., KOLTRACHT, I., AND NEUMANN, M. Convergence of sequential and asynchronous nonlinear paracontractions. *Numer. Math.* 62, 3 (1992), 305–319.
- [46] FAN, P., WANG, W., TONG, R., LI, H., AND TANG, M. gDist: Efficient distance computation between 3D meshes on GPU. In *SIGGRAPH Asia 2024 Conference Papers* (New York, NY, USA, 2024), SA '24, Association for Computing Machinery.
- [47] GILBERT, E. G., JOHNSON, D. W., AND KEERTHI, S. S. A fast procedure for computing the distance between complex objects in three-dimensional space. *IEEE J. Robot. Autom.* 4, 2 (1988), 193–203.
- [48] GOEBEL, K., AND REICH, S. *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings*, vol. 83 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker Inc., New York, 1984.
- [49] GOLDBURG, M., AND MARKS, II, R. J. Signal synthesis in the presence of an inconsistent set of constraints. *IEEE Trans. Circuits and Systems* 32 (1985), 647–663.

- [50] GUBIN, L., POLYAK, B., AND RAIK, E. The method of projections for finding the common point of convex sets. *USSR Comput. Math. Math. Phys.* 7 (1967), 1–24.
- [51] HUANG, S.-W., AND LI, P.-C. Computed tomography sound velocity reconstruction using incomplete data. *IEEE Trans. Ultrason., Ferroelect., Freq. Control* 51, 3 (2004), 329–342.
- [52] IDUKA, H. Fixed point optimization algorithm and its application to power control in CDMA data networks. *Math. Program. (Ser. A)* 133, 1-2 (2012), 227–242.
- [53] IDUKA, H. Iterative algorithm for triple-hierarchical constrained nonconvex optimization problem and its application to network bandwidth allocation. *SIAM J. Optim.* 22, 3 (2012), 862–878.
- [54] IDUKA, H. Decentralized hierarchical constrained convex optimization. *Optim. Eng.* 21, 1 (2020), 181–213.
- [55] JOHNSON, D., AND COHEN, E. A framework for efficient minimum distance computations. In *Proceedings. 1998 IEEE International Conference on Robotics and Automation* (1998), vol. 4, pp. 3678–3684.
- [56] KOLTRACHT, I., AND LANCASTER, P. Constraining strategies for linear iterative processes. *IMA J. Numer. Anal.* 10, 4 (1990), 555–567.
- [57] KOPECKÁ, E., AND REICH, S. A note on the von Neumann alternating projections algorithm. *J. Nonlinear Convex Anal.* 5 (2004), 379–386.
- [58] KOPECKÁ, E., AND REICH, S. A note on alternating projections in Hilbert space. *J. Fixed Point Theory Appl.* 12 (2012), 41–47.
- [59] KOSMOL, P. Über die sukzessive wahl des kürzesten weges. In *Ökonomie und Mathematik: Rudolf Henn zum 65. Geburtstag*, O. Opitz and B. Rauhut, Eds. Springer, Berlin, Heidelberg, 1987, pp. 35–42.
- [60] LIN, M. C., AND GOTTSCHALK, S. Collision detection between geometric models: A survey. In *Proc. of IMA Conference on Mathematics of Surfaces* (1998), vol. 1, pp. 602–608.
- [61] LIN, M. C., MANOCHA, D., AND KIM, Y. J. Collision and proximity queries. In *Handbook of Discrete and Computational Geometry*, J. E. Goodman, J. O’Rourke, and C. D. Tóth, Eds. CRC Press, Boca Raton, FL, 2018, ch. 39. (28 pages).
- [62] LUO, X. Characterizations and uniqueness of mutually nearest points for two sets in normed spaces. *Numer. Funct. Anal. Optim.* 35, 5 (2014), 611–622.
- [63] MORDUKHOVICH, B. S., NAM, N. M., AND SALINAS, J. Applications of variational analysis to a generalized Heron problem. *Appl. Anal.* 91, 10 (2012), 1915–1942.
- [64] MORDUKHOVICH, B. S., NAM, N. M., AND SALINAS, JR., J. Solving a generalized Heron problem by means of convex analysis. *Amer. Math. Monthly* 119, 2 (2012), 87–99.
- [65] MUNKRES, J. *Topology*, second ed. Pearson New International Edition, Essex, England, 2014.
- [66] NARANG, T. D. Some remarks on a paper of K. B. Xu: “A result on best proximity pair of two sets” [J. Approx. Theory 54 (1988), no. 3, 322–325; MR0960054 (90a:41034)]. *Math. Ed. (Siwan)* 25 (1991), 126–128.
- [67] NICOLESCU, M. Sur la meilleure approximation d’une fonction donnée par les fonctions d’une famille donnée. *Bull. Fac. Sti. Cernauti* 12 (1938), 120–128.
- [68] PAI, D. V. Proximal points of convex sets in normed linear spaces. *Yokohama Math. J.* 22 (1974), 53–78.
- [69] PATOGLU, V., AND GILLESPIE, R. Extremal distance maintenance for parametric curves and surfaces. In *Proceedings of the 2002 IEEE International Conference on Robotics and Automation* (2002), vol. 3, pp. 2817–2823.
- [70] QUINLAN, S. Efficient distance computation between non-convex objects. In *Proceedings of the 1994 IEEE International Conference on Robotics and Automation* (1994), vol. 4, pp. 3324–3329.
- [71] REEM, D., AND CENSOR, Y. The best approximation pair problem relative to two subsets in a normed space. *arXiv:2403.18767 [math.OC]*. ([v4] 7 Sep 2026).
- [72] REICH, S. Asymptotic behavior of contractions in Banach spaces. *J. Math. Anal. Appl.* 44 (1973), 57–70.

- [73] SATO, Y., HIRATA, M., MARUYAMA, T., AND ARITA, Y. Efficient collision detection using fast distance-calculation algorithms for convex and non-convex objects. In *Proceedings of IEEE International Conference on Robotics and Automation* (1996), vol. 1, pp. 771–778.
- [74] SCHWARTZ, J. T. Finding the minimum distance between two convex polygons. *Inform. Process. Lett.* 13, 4-5 (1981), 168–170.
- [75] SON, S.-H., YOON, S.-H., KIM, M.-S., AND ELBER, G. Efficient minimum distance computation for solids of revolution. *Comput. Graphics Forum* 39, 2 (2020), 535–544. (Special issue dedicated to Eurographics 2020).
- [76] STILES, W. J. Closest-point maps and their products. II. *Nieuw Arch. Wisk. (3)* 13 (1965), 212–225.
- [77] VOISEI, M. D., AND ZĂLINESCU, C. A counter-example to ‘minimal distance between two non-convex surfaces’. *Optimization* 60, 5 (2011), 593–602.
- [78] XU, X. A result on best proximity pair of two sets. *J. Approx. Theory* 54 (1988), 322–325.
- [79] YOULA, D. C., AND VELASCO, V. Extensions of a result on the synthesis of signals in the presence of inconsistent constraints. *IEEE Trans. Circuits and Systems* 33 (1986), 465–468.
- [80] ZEGHLOUL, S., AND RAMBEAUD, P. A fast algorithm for distance calculation between convex objects using the optimization approach. *Robotica* 14, 4 (1996), 355–363.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF HAIFA, MT. CARMEL, HAIFA 3498838, ISRAEL
Email address: yair@math.haifa.ac.il

THE CENTER FOR MATHEMATICS AND SCIENTIFIC COMPUTATION (CMSC), UNIVERSITY OF HAIFA, MT. CARMEL, HAIFA 3498838, ISRAEL
Email address: dream@math.haifa.ac.il