Finding a best approximation pair of points for two polyhedra

Ron Aharoni[*,†] Yair Censor[‡] Zilin Jiang[*,§]

July 30, 2017

Abstract

Given two disjoint convex polyhedra, we look for a best approximation pair relative to them, i.e., a pair of points, one in each polyhedron, attaining the minimum distance between the sets. Cheney and Goldstein showed that alternating projections onto the two sets, starting from an arbitrary point, generate a sequence whose two interlaced subsequences converge to a best approximation pair. We propose a process based on projections onto the half-spaces defining the two polyhedra, which are more negotiable than projections on the polyhedra themselves. A central component in the proposed process is the Halpern–Lions–Wittmann–Bauschke algorithm for approaching the projection of a given point onto a convex set.

Keywords: Best approximation pair, convex polyhedra, alternating projections, half-spaces, Cheney–Goldstein theorem, Halpern–Lions–Wittmann–Bauschke algorithm

Mathematics Subject Classification: 65K05, 90C20, 90C25

1 Introduction

A *best approximation pair* relative to two closed convex sets A and B is a pair $(a, b) \in A \times B$ attaining $\|a - b\| = \min \|A - B\|$, where $A - B := \{x - y \mid x \in A, y \in B\}$.

For a closed convex set C denote by P_C the metric projection operator onto C. Take an arbitrary starting point $a_0 \in \mathbb{R}^d$, the d-dimensional Euclidean space, and consider the sequence:

$$
\begin{tikzcd}
a_0 \arrow[dr, "P_B"'] & & a_2 \arrow[dr, "P_B"'] & & \cdots & & a_{2k} \arrow[dr, "P_B"'] & & a_{2k+2} & & \cdots \\
& b_1 \arrow[ur, "P_A"'] & & b_3 & & b_{2k-1} \arrow[ur, "P_A"'] & & b_{2k+1} \arrow[ur, "P_A"'] & & &
\end{tikzcd}
$$

A well-known theorem of Cheney and Goldstein [CG59] specifies conditions under which alternating metric projections onto the two sets are guaranteed to converge to the best approximation pair.

[*]Department of Mathematics, the Technion – Israel Institute of Technology, Technion City, Haifa 3200003, Israel.

[†]Supported in part by the United States–Israel Binational Science Foundation (BSF) grant no. 2012031, the Israel Science Foundation (ISF) grant no. 2023464 and the Discount Bank Chair at the Technion. Email: `ra@tx.technion.ac.il`.

[‡]Department of Mathematics, University of Haifa, Mt. Carmel, Haifa 3498838, Israel. Supported in part by BSF grants no. 2013003. Email: `yair@math.haifa.ac.il`.

[§]Supported in part by ISF grant nos 1162/15, 936/16. Email: `jiangzilin@technion.ac.il`.

In fact, their result applies when A and B are closed convex sets in a Hilbert space, and one of them is compact. For related results, using the averaged alternating reflections method and applying it to not necessarily convex sets, see [BCL04, Luk08].

In real-life problems, convex polyhedra are usually represented by a set of linear constraints, namely as the intersection of half-spaces. Projecting onto the polyhedron can then be done using projections onto the half-spaces. We propose a projection method for finding the best approximation pair that uses directly projections onto the half-spaces, instead of on the polyhedra.

An algorithm for approaching the projection of a point a onto a polyhedron, using projections onto the half-spaces defining the polyhedron, was proposed by Halpern, Lions, Wittmann and Bauschke (HLWB)[1] [Hal67, Lio77, Wit92, Bau96]. The HLWB algorithm is designed to find the projection. It works by projecting successively and cyclically onto the half-spaces, the main stratagem being that after each projection the algorithm "pulls" a bit back in the direction of a. The latter guarantees that the algorithm does not "forget" the point a whose projection onto the polyhedron is sought after.

In the algorithm proposed and studied in this paper, the HLWB algorithm is applied alternatingly to the two polyhedra. Its application is divided into sweeps — in the odd numbered sweeps we project successively onto half-spaces defining A, and in even numbered sweeps onto half-spaces defining B. A critical point is that the number of successive projections onto each set's half-spaces increases from sweep to sweep. The proof of convergence of the algorithm is rather standard in the case that the best approximation pair is unique. The non-uniqueness case, however, poses some difficulties and its proof is more involved.

The algorithm belongs to a family known as *projection methods*. These are iterative algorithms that use projections onto individual sets, to converge to a point in the intersection of these sets, or images of them under some transformation. They were originally used to solve systems of linear equations in Euclidean space (see, e.g., [Byr08, Ceg12, CZ97, ER11, Gal04]), and later were extended to solve general convex feasibility problems in a Hilbert space, see, e.g., [BC17]. On the low computational cost of projection methods, see [BK15, CCC$^+$12]. Consult also [CC15].

The paper is organized as follows. In Section 2 we present the alternating HLWB algorithm. In Section 3 we prove some preliminary results needed for the proof of convergence of the algorithm. The proof itself is given in Section 4. Finally, in Section 5 we discuss possible choices of the parameters for the algorithm.

2 An Alternating HLWB Algorithm

Throughout the paper, we assume that $A := \cap_{i=1}^{M} A_i$ and $B := \cap_{j=1}^{N} B_j$ are two nonempty convex polyhedra, where $\{A_i\}_{i=1}^{M}$ and $\{B_j\}_{j=1}^{N}$ are two families of closed half-spaces. By adding A_i or B_j that are equal to the entire space $\mathbb{R}^d$ (or alternatively repeat the same half-space) we may assume that $M = N$. For the purpose of performing unboundedly many projections, we extend the sequences

[1]This acronym was dubbed in [Cen06].

$\{A_i\}$ and $\{B_j\}$ to all $i, j \in \mathbb{N}$ by the rules $A_i = A_{i \bmod N}$ and $B_j = B_{j \bmod N}$, where the mod$N$ function takes values in $\{1, 2, \ldots, N\}$.

We incorporate into our algorithm the HLBW algorithm, which is designed to find the projection of a point a onto a polyhedron C, using the projections onto the half-spaces defining C. Let $P_1, P_2, \ldots, P_N$ be the respective projections onto these half-spaces. The HLBW algorithm starts by choosing an arbitrary starting point x_0 and numbers λ_n satisfying:

$$\lim_{n\to\infty} \lambda_n = 0, \quad \textstyle\sum_n \lambda_n = \infty, \quad \textstyle\sum_n |\lambda_n - \lambda_{n+N}| < \infty. \tag{1}$$

A sequence $\{x_n\}_{n=1}^{\infty}$ is then recursively generated by the rule:

$$x_n := \lambda_n a + (1 - \lambda_n) P_{n \bmod N}(x_{n-1}),$$

Bauschke [Bau96, Theorem 3.1] proved that the sequence $\{x_n\}_{n=0}^{\infty}$ generated by this HLBW algorithm convergences to $P_C(a)$. Some computational performance results with the HLWB and the Dykstra [Dyk83] algorithms were presented in [Cen06].

In our proposed algorithm, we apply the HLBW algorithm alternatingly to A and to B. We call this method "A-HLBW" ("A" for "alternating"). Like in HLWB, we choose numbers λ_n satisfying (1). For points $a, x \in \mathbb{R}^d$ and $n \in \mathbb{N}$, we recursively define

$$Q_{B,0}(a; x) := x \quad \text{and} \quad Q_{B,n}(a; x) := \lambda_n a + (1 - \lambda_n) P_{B_n}(Q_{B,n-1}(a; x)).$$

For $b, x \in \mathbb{R}^d$ and $n \in \mathbb{N}$, $Q_{A,n}(b; x)$ is similarly defined

$$Q_{A,0}(b; x) := x \quad \text{and} \quad Q_{A,n}(b; x) := \lambda_n b + (1 - \lambda_n) P_{A_n}(Q_{A,n-1}(b; x)).$$

Thus, $Q_{A,n}(b; x)$ and $Q_{B,n}(a; x)$ are operators, each being defined by a sequence of n iterations.

We also choose an arbitrary starting point $a_0 \in \mathbb{R}^d$ and a non-decreasing sequence (n_k) such that

$$n_k \to \infty \quad \text{and} \quad \sup_{k_0} \left\{ \sum_{k>k_0} \prod_{n>n_{k_0}}^{n_k} (1 - \lambda_n) \right\} < \infty. \tag{2}$$

Once the sequence (λ_n) is chosen so that (1) holds, one can always make (n_k) increase rapidly enough so that (2) holds. For example, $\lambda_n = \frac{1}{n+1}, n_k = \lfloor 1.1^k \rfloor$ satisfy both (1) and (2).

The kth sweep of the A-HLWB algorithm uses n_k iterations of the HLWB algorithm to generate:

$$b_{k+1} = Q_{B,n_k}(a_k; a'_k) \text{ if } k \text{ is even}; \quad a_{k+1} = Q_{A,n_k}(b_k; b'_k) \text{ if } k \text{ is odd}, \tag{3}$$

where the auxiliary parameter a'_k or b'_k is chosen before each sweep. The validity of the algorithm is guaranteed if the auxiliary sequence (a'_{2k}, b'_{2k+1}) is bounded. For example, one may simply take $a'_{2k} = b'_{2k+1} = a_0$. We now state our main convergence result.

Theorem 1. *If the above assumptions on the mappings and on the parameters hold and the auxiliary sequence (a'_{2k}, b'_{2k+1}) is bounded, then the pairs (a_{2k}, b_{2k+1}), generated by the A-HLWB algorithm* (3)*, converge to a best approximation pair relative to* (A, B)*.*

The second inequality in (2) is technical and could possibly be redundant. In fact, if the best approximation pair is unique, the convergence is assured without this inequality (see Remark 1). However, we are unable to remove it for Theorem 1 in the non-uniqueness case.

3 Preliminaries for the Proof of Convergence

We present several preliminary results that will be used to prove Theorem 1, the first of which says, in Lemma 3 below, that the set of points generated by the A-HLWB algorithm is bounded. This follows from a result of Aharoni, Duchet and Wajnryb [ADW84], see also Meshulam [Mes96].

Theorem 2 ([ADW84])**.** *Any sequence of points in $\mathbb{R}^d$ obtained by successive projections of a point onto elements of a finite set of hyperplanes is bounded.*

Lemma 3. *For every bounded set $D \subset \mathbb{R}^d$, there exists a compact set $C \subset \mathbb{R}^d$ containing D such that $Q_{A,m}(b;x), Q_{B,n}(a;x) \in C$ for all $a, b, x \in C$ and $m, n \in \mathbb{N}$.*

Proof. Take a simplex with vertices $x_0, x_1, \ldots, x_d \in \mathbb{R}^d$ that contains D. Denote the bounding hyperplanes of the half-spaces A_i and B_i by ∂A_i and ∂B_i, respectively. Let X be the set of points obtained by successive projections of $x_0, x_1, \ldots, x_d$ on $\{\partial A_i\}_{i=1}^M$ and on $\{\partial B_i\}_{i=1}^N$. By Theorem 2, we know that X is bounded, and so is its convex hull $Y := \mathrm{conv}(X)$.

Notice that $Q_{A,m}(b;x)$ is either $\lambda_m b + (1-\lambda_m)Q_{A,m-1}(b;x)$ or $\lambda_m b + (1-\lambda_m)P_{\partial A_m}Q_{A,m-1}(b;x)$ depending on whether $Q_{A,m-1}(b;x)$ is in A_m. One can then show by induction on m that $Q_{A,m}(b;x) \in Y$ for every $b, x \in Y$.

The same argument shows that $Q_{B,n}(a;x) \in Y$ for every $a, x \in Y$. Finally, let C be the closure of Y in $\mathbb{R}^d$. Since $Q_{A,m}$ and $Q_{B,n}$ are continuous, $Q_{A,m}(b;x)$ and $Q_{B,n}(a;x)$ are in C for every $a, b, x \in C$. $\square$

The following result is well-known, see, e.g., [CG59, Theorem 3].

Theorem 4. *If B is a closed convex set in Hilbert space, then the projection map P_B onto B satisfies the Lipschitz condition $\|P_B(x) - P_B(y)\| \leq \|x - y\|$, equality holding only if $\|x - P_B(x)\| = \|y - P_B(y)\|$.*

The classical 1959 result of Cheney and Goldstein, repeatedly referred to in this paper, is given next as a paraphrased version of their Theorems 2 and 4.

Theorem 5. *Let A and B be two closed convex sets in Hilbert space. A point of A is nearest to B if and only if it is a fixed point of $P_A P_B$. If one set is finite-dimensional and the distance between the sets is attained, then convergence of $((P_A P_B)^n(x))$ to a fixed point of $P_A P_B$ is assured.*

We need also the following result.

Lemma 6. *Let A and B be two closed convex sets in Hilbert space, one of which being finite-dimensional. Suppose that the distance between the sets is attained. If S is a nonempty compact set such that $P_AP_B(S) = S$, then S consists of points of A nearest to B.*

Proof. Define $S' := \{s \in S \mid P_AP_B(s) = s\}$. Since S is compact and $P_AP_B(S) \subset S$, by the second part of Theorem 5, for any $x \in S$, $(P_AP_B)^n(x)$ converges in S and its limit is a fixed point of P_AP_B.

Since S is nonempty, so is S' and it is easy to see that S' is compact as well.

Let $d := \max_{s\in S} \inf \|s - S'\|$ and let $y \in S$ be such that $\inf \|y - S'\| = d$. Since $P_AP_B(S) \supset S$, there exists $x \in S$ such that $P_AP_B(x) = y$. Since $\min \|x - S'\| \le d$, we can take $s' \in S'$ such that $\|x - s'\| \le d$.

By way of contradiction, assume that $x \notin S'$. By the first part of Theorem 5,

$$\left\|P_B(s') - s'\right\| = \inf \|A - B\| < \|P_B(x) - x\|.$$

By Theorem 4, we obtain

$$\left\|y - s'\right\| = \left\|P_AP_B(x) - P_AP_B(s')\right\| \le \left\|P_B(x) - P_B(s')\right\| < \left\|x - s'\right\|.$$

This contradicts with $\|x - s'\| \le d \le \|y - s'\|$. Therefore, $x \in S'$, hence $y = x$ and $d = 0$. By the first part of Theorem 5, $S = S'$ implies that S consists only of points of A nearest to B. □

The last ingredient that will be used in our proof of Theorem 1 is the following.

Theorem 7. *Let B be a polyhedron in Hilbert space, and assume that $B = \cap_{i=1}^N B_i \neq \emptyset$, where $\{B_i\}_{i=1}^N$ are closed convex sets. If the sequence (λ_n) satisfies (1), then $\lim_{n\to\infty} \|Q_{B,n}(a;x) - P_B(a)\| = 0$ for any points a and x_0.*

Proof. This follows from Bauschke's Theorem 3.1 in [Bau96]. In fact, Bauschke's theorem applies to a broader setting, in which the B_i's are sets of fixed points of nonexpansive mappings in Hilbert space. □

It is easy to check that

$$\left\|Q_{B,n}(a;x) - Q_{B,n}(a';x')\right\| \le \left\|a - a'\right\| + \left\|x - x'\right\|$$

for all n. Together with the fact that P_B is nonexpansive (i.e., 1-Lipschitz), it is routine to check the uniform convergence of $(Q_{B,n})$ on any compact set, leading to the next lemma.

Lemma 8. *Let B be as in Theorem 7, and let C be a compact set in the Hilbert space. If the sequence (λ_n) satisfies (1), then*

$$\lim_{n\to\infty} \left(\sup_{a,x\in C} \|Q_{B,n}(a;x) - P_B(a)\| \right) = 0.$$

4 Convergence of the A-HLWB Algorithm

In this section we present a proof of the convergence theorem of the A-HLWB algorithm.

Proof of Theorem 1. In order to prove this theorem we inspect the set of accumulation points of (a_{2k}). We show that it is compact, fixed under $P_A P_B$ and, finally, that it is a singleton. By Lemma 3, there exists a compact set $C \subset \mathbb{R}^d$ containing $\{a_0\} \cup \{a'_{2k}\}_{k=0}^{\infty} \cup \{b'_{2k+1}\}_{k=0}^{\infty}$ such that both $Q_{A,m}$ and $Q_{B,n}$ map $C \times C$ to C, hence the sequences (a_{2k}) and (b_{2k+1}) are contained in C.

Let S be the set of accumulation points of (a_{2k}). By the Bolzano–Weierstrass theorem $S \neq \emptyset$. Moreover, since S is closed and $S \subset C$, it is compact.

We claim that $P_A P_B(S) = S$. Pick any point $s \in S$ and any $\epsilon > 0$. Using Lemma 8 and the assumption that $n_k \to \infty$, one can choose k sufficiently large such that

$$\|a_{2k} - s\| < \epsilon, \quad \sup_{a,x \in C} \|Q_{B,n_{2k}}(a;x) - P_B(a)\| < \epsilon/2, \quad \sup_{a,x \in C} \left\|Q_{A,n_{2k+1}}(b;x) - P_A(b)\right\| < \epsilon/2.$$

In particular, since $b_{2k+1} = Q_{B,n_{2k}}(a_{2k}; a'_{2k})$ and $a_{2k+2} = Q_{A,n_{2k+1}}(b_{2k+1}; b'_{2k+1})$, we have

$$\|P_B(a_{2k}) - b_{2k+1}\| < \epsilon/2, \quad \|P_A(b_{2k+1}) - a_{2k+2}\| < \epsilon/2.$$

By the triangle inequality and the fact that P_A and P_B are nonexpansive, we obtain that

$$\begin{aligned} \|P_A P_B(s) - a_{2k+2}\| &\leq \|P_A P_B(s) - P_A P_B(a_{2k})\| + \|P_A P_B(a_{2k}) - P_A(b_{2k+1})\| \\ &+ \|P_A(b_{2k+1}) - a_{2k+2}\| \leq \|s - a_{2k}\| + \|P_B(a_{2k}) - b_{2k+1}\| + \|P_A(b_{2k+1}) - a_{2k+2}\| < 2\epsilon. \end{aligned} \tag{4}$$

This implies that $P_A P_B(s)$ is also an accumulation point of (a_{2k}). Thus, $P_A P_B(S) \subset S$. On the other hand, suppose that $s \in S$ is the limit of the subsequence (a_{2k_l}). Let $s' \in S$ be an accumulation point of the subsequence $(a_{2k_l - 2})$. The same argument for $P_A P_B(S) \subset S$ shows that $P_A P_B(s')$ is an accumulation point of (a_{2k_l}), and so $P_A P_B(s') = s$. This means that $P_A P_B(S) \supset S$. By Lemma 6, S consists of points of A nearest to B.

It remains to be shown that S is a singleton, namely that it contains only one point, which is then the limit of a_{2k}. This is clear if there is only one best approximation pair.

From here on we consider the case that A and B have parallel closest faces. This situation requiers a deeper and more delicate analysis which we give now.

Let v be the shortest vector of the form $a - b$, where $a \in A$ and $b \in B$. Put differently, v is the projection of the origin onto $A - B$. Set $T := A \cap (B + v)$. Clearly, T is precisely the set of all points in A nearest to B. Therefore $S \subset T$. Moreover, T is a convex polyhedron inside the supporting hyperplane of A that is perpendicular to v (see Figure 1.)

We decompose the polyhedron T into the relative interiors of its faces (see Figure 2.) Let e be the largest integer such that the relative interior of some e-dimensional face F_e intersects S, say at point s. We shall prove that (a_{2k}) converges to s. Since, by Lemma 8, $\lim_{k\to\infty} \|b_{2k+1} - P_B(a_{2k})\| = 0$, this

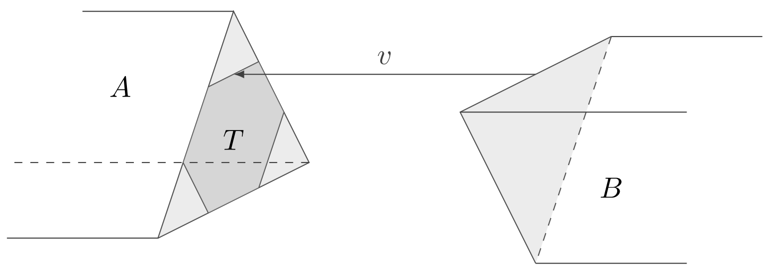

Figure 1

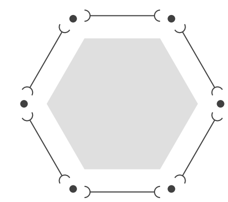

Figure 2

will imply that (b_{2k+1}) converges to $P_B(s)$, and (a_{2k}, b_{2k+1}) thus converges to the best approximation pair $(s, P_B(s))$.

The proof that $\lim_{k\to\infty} a_{2k} = s$ combines ideas from the two extreme cases for e, namely $e = 0$ or $e = d-1$. We first handle these two cases and then present the general case.

Case 1: $e = 0$**.** Suppose that all points of S are vertices of T. Let $\epsilon_0 > 0$ be such that $N_{\epsilon_0}(s)$ (the ϵ_0-neighborhood of s) satisfies $N_{\epsilon_0}(s) \cap T_0 = \{s\}$. Denoting by T_0 the set of all accumulation points of (a_{2k}), $S \subset T_0$ implies that every neighborhood of T_0 contains all but finitely many points of (a_{2k}). For every $\epsilon \in (0, \epsilon_0/4)$, we can then choose $k_0 \in \mathbb{N}$ so that

$$\|a_{2k_0} - s\| < \frac{\epsilon}{4}, \quad \inf \|a_{2k} - T_0\| < \frac{\epsilon}{4} \text{ for all } k \geq k_0,$$

and, using Lemma 8,

$$\sup_{b,x\in C} \|Q_{A,n}(b;x) - P_A(b)\| < \frac{\epsilon}{4}, \sup_{a,x\in C} \|Q_{B,n}(a;x) - P_B(a)\| < \frac{\epsilon}{4}, \text{ for all } n \geq n_{2k_0}. \tag{5}$$

We claim that for every $k \geq k_0$, if $\|a_{2k} - s\| < \epsilon/4$, then $\|a_{2k+2} - s\| < \epsilon_0/2$ (see Figure 3) and so $\|a_{2k+2} - s\| = \inf \|a_{2k+2} - T_0\|$, by the choice of ϵ_0. In fact, (4) implies

$$\|a_{2k+2} - s\| \leq \|a_{2k} - s\| + \|P_B(a_{2k}) - b_{2k+1}\| + \|P_A(b_{2k+1}) - a_{2k+2}\| < \epsilon/4 + \epsilon/4 + \epsilon/4 < \epsilon_0/2$$

Hence, $N_\epsilon(a_{2k+2}) \subset N_{\epsilon_0}(s)$, and so

$$N_\epsilon(a_{2k+2}) \cap T_0 \subset N_{\epsilon_0}(s) \cap T_0 = \{s\}\,.$$

This means that

$$\|a_{2k+2} - s\| = \inf \|a_{2k+2} - T_0\| < \epsilon/4.$$

By induction, we know that $\|a_{2k} - s\| < \epsilon/4$ for all $k \geq k_0$.

Case 2: $e = d-1$**.** Suppose s is in the relative interior of the $(d-1)$-dimensional face $T = T_e$. Let H be the supporting hyperplane of A that is perpendicular to v. Note that $T \subset A \cap H$. If ∂A_i goes through s, then $\partial A_i = H$ (otherwise, T would have dimension smaller than $d-1$.) Similarly, if ∂B_i goes through $P_B(s) = s - v$, then $\partial B_i = H - v$.

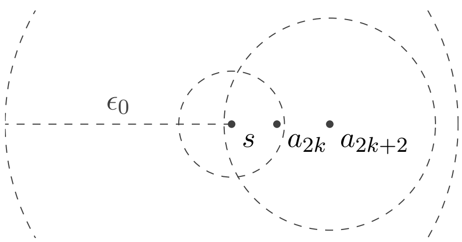

Figure 3

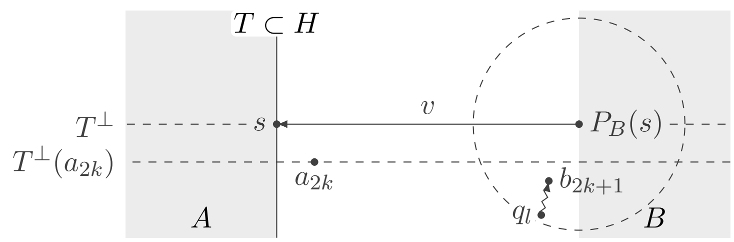

Figure 4

Thus, we can choose an ϵ_0-neighborhood $N_{\epsilon_0}(s)$ of s such that $N_{\epsilon_0}(s) \subset A_i$ for all $\partial A_i \neq H$ and such that $N_{\epsilon_0}(P_B(s)) \subset B_i$ for all $\partial B_i \neq H - v$. This, in particular, implies that for every $x \in N_{\epsilon_0}(P_B(s))$, $P_{B_i}(x) - x$ is always orthogonal to T.

For the rest of the proof of this case for $e = d-1$, we only highlight the key steps and leave the full-fledged proof to the general case below. For a fixed $k \geq k_0$, let $T^\perp$ and $T^\perp(a_{2k+2})$ be the lines orthogonal to T through s and through a_{2k}, respectively, and define $\ell^+ := \ell + n_{2k_0}$ and $q_\ell := Q_{B,\ell^+}(a_{2k}; a'_{2k})$ (see Figure 4). We claim that if $q_\ell \in N_{\epsilon_0}(P_B(s))$ for all $\ell \geq 0$, then

$$\inf \left\| b_{2k+1} - T^\perp(a_{2k}) \right\| = \inf \left\| q_0 - T^\perp(a_{2k}) \right\| \cdot \prod_{n > n_{2k_0}}^{n_{2k}} (1 - \lambda_n).$$

Indeed, recall that for $\ell \geq 1$, we have the recursion

$$q_\ell = \lambda_{\ell^+} a_{2k} + (1 - \lambda_{\ell^+}) P_{B_{\ell^+}}(q_{\ell-1}). \tag{6}$$

Since $q_{\ell-1} \in N_{\epsilon_0}(P_B(s))$, by the choice of ϵ_0, we know that $P_{B_{\ell^+}}(q_{\ell-1}) - q_{\ell-1}$ is orthogonal to T, hence

$$\inf \left\| P_{B_{\ell^+}}(q_{\ell-1}) - T^\perp \right\| = \inf \left\| q_{\ell-1} - T^\perp \right\|.$$

Therefore,

$$\inf \left\| q_\ell - T^\perp \right\| = (1 - \lambda_{\ell^+}) \inf \left\| q_{\ell-1} - T^\perp \right\|.$$

The claim follows from repeated application of this equality and the fact that $b_{2k+1} = q_{n_{2k} - n_{2k_0}}$.

Loosely speaking, the claim suggests that b_{2k+1} does not deviate far from $T^\perp(a_{2k})$, and, similarly, neither does a_{2k+2} deviate from $T^\perp(b_{2k+1})$. The accumulated deviations of a_{2k} from $T^\perp$ then can be bounded. This, together with the fact that a_{2k} converges to a point in T, implies that $\lim_{k\to\infty} a_{2k} = s$.

General case: $0 \leq e \leq d-1$. The proof of the general case is the juxtaposition of the ideas in the two previos cases above. Let T_e be the union of the e-dimensional faces of T. Since s is in the relative interior of F_e, we can choose $\epsilon_0 > 0$ so that the following hold:

1. $N_{\epsilon_0}(s) \cap T_e \subset F_e$;
2. for every $x \in N_{\epsilon_0}(s)$, $P_{A_m}(x) - x$ is orthogonal to F_e for all m;

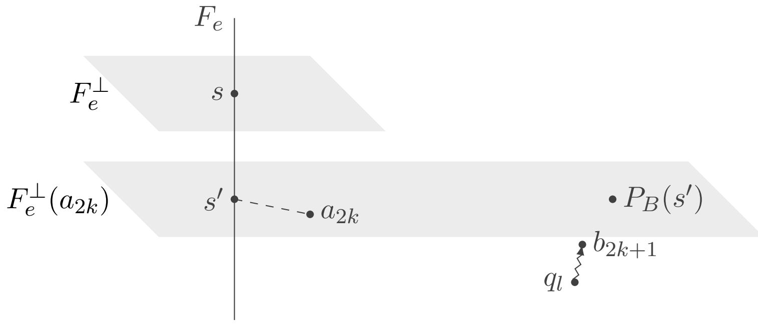

Figure 5

3. for every $x \in N_{\epsilon_0}(P_B(s))$, $P_{B_n}(x) - x$ is orthogonal to F_e for all n.

In view of (2), there is a constant $Z > 1$ such that

$$\sum_{l>k_0} \prod_{n>n_{2k_0}}^{n_{2l+1}} (1-\lambda_n) < Z \text{ for all } k_0 \in \mathbb{N}. \tag{7}$$

For every $\epsilon \in (0, \epsilon_0/4)$, we can choose $k_0 \in \mathbb{N}$ so that (5) and the following hold.

$$\|a_{2k_0} - s\| < \frac{\epsilon}{4Z}, \quad \inf \|a_{2k} - T_e\| < \frac{\epsilon}{4Z} \text{ for all } k \geq k_0.$$

Let $F_e^\perp$ be the $(d-e)$-dimensional affine subspace through s orthogonal to F_e. We prove by induction that for all $k \geq k_0$,

$$\inf \|a_{2k} - F_e\| < \frac{\epsilon}{4Z}, \tag{8a}$$

$$\inf \left\|a_{2k} - F_e^\perp\right\| < \frac{\epsilon}{4Z} + \frac{\epsilon}{2Z} \cdot \sum_{l>k_0}^{k} \prod_{n>n_{2k_0}}^{n_{2l+1}} (1-\lambda_m), \tag{8b}$$

$$\|a_{2k} - s\| < \epsilon. \tag{8c}$$

When $k = k_0$, it is obvious from the choice of k_0. Assume that (8) holds for k, we prove it for $k+1$. **Proof of** (8a)**.** We use (4) to obtain

$$\|s - a_{2k+2}\| \leq \|s - a_{2k}\| + \|P_B(a_{2k}) - b_{2k+1}\| + \|P_A(b_{2k+1}) - a_{2k+2}\| < \epsilon + \epsilon/4Z + \epsilon/4Z < \epsilon_0/2.$$

Hence, $N_{2\epsilon}(a_{2k+2}) \subset N_{\epsilon_0}(s)$, and so

$$N_{2\epsilon}(a_{2k+2}) \cap T_e \subset N_{\epsilon_0}(s) \cap T_e \subset F_e.$$

This means that

$$\inf \|a_{2k+2} - F_e\| = \inf \|a_{2k+2} - T_e\| < \epsilon/4Z.$$

Proof of (8b)**.** Again, define $\ell^+ := \ell + n_{2k_0}$ and $q_\ell := Q_{B,\ell^+}(a_{2k}; a'_{2k})$. For every $\ell \geq 0$, we have

$$\|P_B(s) - q_\ell\| \leq \|s - a_{2k}\| + \|P_B(a_{2k}) - q_\ell\| < \epsilon + \frac{\epsilon}{4Z} < \epsilon_0. \tag{9}$$

Let $F_e^\perp(a_{2k})$ be the $(d-k)$-dimensional subspace through a_{2k} orthogonal to F_e (see Figure 5). We have the recursion (6) for $\ell \geq 1$. Since $\|P_B(s) - q_{\ell-1}\| < \epsilon_0$, by the choice of ϵ_0, we know that $P_{B_{\ell^+}}(q_{\ell-1}) - q_{\ell-1}$ is orthogonal to F_e, hence,

$$\inf \left\| P_{B_{\ell^+}}(q_{\ell-1}) - F_e^\perp(a_{2k}) \right\| = \inf \left\| q_{\ell-1} - F_e^\perp(a_{2k}) \right\|.$$

Therefore,

$$\inf \left\| q_\ell - F_e^\perp(a_{2k}) \right\| = (1 - \lambda_{\ell^+}) \inf \left\| q_{\ell-1} - F_e^\perp(a_{2k}) \right\|. \tag{10}$$

Let s' be the intersection of F_e and $F_e^\perp(a_{2k})$, which is guaranteed to be nonempty by $\|a_{2k} - s\| < \epsilon < \epsilon_0$. Moreover, we have

$$\left\| a_{2k} - s' \right\| = \inf \|a_{2k} - F_e\| < \epsilon/4Z.$$

Since $s' \in T$, $P_B(s') = s' + v \in F_e^\perp(a_{2k})$. Notice that

$$\inf \left\| q_0 - F_e^\perp(a_{2k}) \right\| \leq \left\| q_0 - P_B(s') \right\| \leq \|q_0 - P_B(a_{2k})\| + \left\| a_{2k} - s' \right\| \leq \epsilon/4Z + \epsilon/4Z = \epsilon/2Z.$$

With repeated application of (10) and the fact that $b_{2k+1} = q_{n_{2k} - n_{2k_0}}$, we derive

$$\inf \left\| b_{2k+1} - F_e^\perp(a_{2k}) \right\| \leq \frac{\epsilon}{2Z} \cdot \prod_{n > n_{2k_0}}^{n_{2k}} (1 - \lambda_m). \tag{11}$$

Let $F_e^\perp(b_{2k+1})$ be the $(d-k)$-dimensional subspace through b_{2k+1} orthogonal to F_e. Notice that (9) implies that

$$\|b_{2k+1} - P_B(s)\| < \epsilon + \epsilon/4Z.$$

This would allow us to carry out a similar argument to conclude that

$$\inf \left\| a_{2k+2} - F_e^\perp(b_{2k+1}) \right\| < \frac{\epsilon}{2Z} \cdot \prod_{n > n_{2k_0}}^{n_{2k+1}} (1 - \lambda_m). \tag{12}$$

Combining (11) and (12), we get what is needed for the inductive step.

Proof of (8c)**.** It follows from

$$\|a_{2k+2} - s\| \leq \inf \|a_{2k+2} - F_e\| + \inf \left\| a_{2k+2} - F_e^\perp \right\|$$

and (7, 8a, 8b). □

Remark 1. *As mentioned in the proof, if the polyhedra A and B are known a priori to have only one best approximation pair, then the set S is automatically a singleton, hence (a_{2k}, b_{2k+1}) converges to the best approximation pair. In this case, the second inequality in* (2) *could be dropped from the assumptions in Theorem 1.*

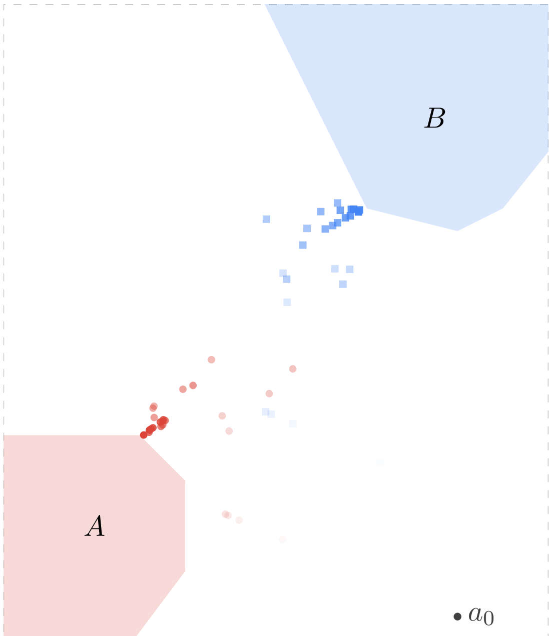

Figure 6

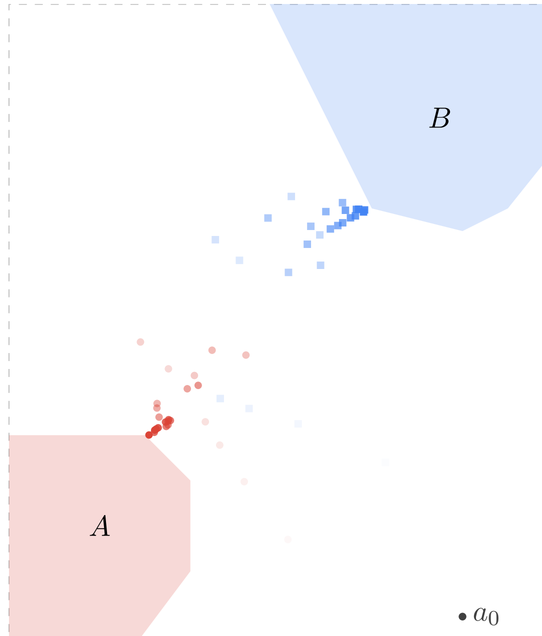

Figure 7

5 Discussion

One has the freedom to choose the auxiliary sequence (a'_{2k}, b'_{2k+1}) in the A-HLWB algorithm as long as it is bounded. The simplest way is to take $a'_{2k} = b'_{2k+1} = a_0$. In Figure 6, we run 50 iterations of the algorithm, and we plot the more recent points in darker color. The half-spaces are

$$A_1\colon 4x - 3y \le 17, \qquad A_2\colon x \le -4, \qquad A_3\colon x + y \le -11, \qquad A_4\colon y \le -5;$$
$$B_1\colon 5x - 4y \le 30, \qquad B_2\colon x - 2y \le 0, \qquad B_3\colon -x - 4y \le -24, \qquad B_4\colon -2x - y \le -13.$$

The choices of the parameters are

$$a_0 = (8, -13), \quad \lambda_n = \tfrac{1}{n+1}, \quad n_k = \lfloor 1.1^k \rfloor, \quad a'_{2k} = b'_{2k+1} = a_0.$$

The auxiliary point a'_{2k} can be seen as the starting point of the HLWB algorithm for polyhedron B. It makes sense to choose a'_{2k} close to B. Since b_{2k-1} is our best approximation to B so far, heuristically $a'_{2k} = b_{2k-1}$ might be a better choice. Similarly, it might be better to choose $b'_{2k+1} = a_{2k}$. One can use Lemma 3 to show that (a'_{2k}, b'_{2k+1}) is bounded. In Figure 7, we again run 50 iterations for the same half-spaces and the parameters, except that

$$a'_0 = a_0, \quad a'_{2k} = b_{2k-1} \text{ for all } k > 0, \quad b'_{2k+1} = a_{2k} \text{ for all } k \ge 0.$$

Comparing Figure 6 and Figure 7, interestingly we do not see significant difference in convergence. One possible explanation is the following. It is easily checked that

$$\left\| Q_{B,n_{2k}}(a;x) - Q_{B,n_{2k}}(a;x') \right\| \le \left\| x - x' \right\| \textstyle\prod_{n=1}^{n_{2k}} (1 - \lambda_n).$$

When $\prod_{n=1}^{n_{2k}} (1 - \lambda_n)$ is extremely small, the contribution of the auxiliary points is negligible.

Acknowledgements

We thank Yehuda Zur for Matlab programming work at the early stages of our research.

References

[ADW84] R. Aharoni, P. Duchet, and B. Wajnryb. Successive projections on hyperplanes. *J. Math. Anal. Appl.*, 103:134–138, 1984.

[Bau96] H. H. Bauschke. The approximation of fixed points of compositions of nonexpansive mappings in Hilbert space. *J. Math. Anal. Appl.*, 202:150–159, 1996.

[BC17] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics. Springer International Publishing, second edition, 2017.

[BCL04] H. H. Bauschke, P. L. Combettes, and D. R. Luke. Finding best approximation pairs relative to two closed convex sets in Hilbert spaces. *J. Approx. Theory*, 127:178–192, 2004.

[BK15] H. H. Bauschke and V. R. Koch. Projection methods: Swiss army knives for solving feasibility and best approximation problems with halfspaces. In *Infinite products of operators and their applications*, volume 636 of *Contemp. Math.*, pages 1–40. Amer. Math. Soc., Providence, RI, 2015.

[Byr08] C. L. Byrne. *Applied Iterative Methods*. A K Peters, Ltd., Wellesley, MA, 2008.

[CC15] Y. Censor and A. Cegielski. Projection methods: an annotated bibliography of books and reviews. *Optimization*, 64:2343–2358, 2015.

[CCC+12] Y. Censor, W. Chen, P. L. Combettes, R. Davidi, and G. T. Herman. On the effectiveness of projection methods for convex feasibility problems with linear inequality constraints. *Comput. Optim. Appl.*, 51:1065–1088, 2012.

[Ceg12] A. Cegielski. *Iterative Methods for Fixed Point Problems in Hilbert Spaces*, volume 2057 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, Heidelberg, 2012.

[Cen06] Y. Censor. Computational acceleration of projection algorithms for the linear best approximation problem. *Linear Algebra Appl.*, 416:111–123, 2006.

[CG59] W. Cheney and A. A. Goldstein. Proximity maps for convex sets. *Proc. Amer. Math. Soc.*, 10:448–450, 1959.

[CZ97] Y. Censor and S. A. Zenios. *Parallel Optimization: Theory, Algorithms, and Applications*. Oxford University Press, New York, NY, 1997.

[Dyk83] R. L. Dykstra. An algorithm for restricted least squares regression. *J. Amer. Statist. Assoc.*, 78:837–842, 1983.

[ER11] R. Escalante and M. Raydan. *Alternating Projection Methods*, volume 8 of *Fundamentals of Algorithms*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011.

[Gal04] A. Galántai. *Projectors and Projection Methods*, volume 6 of *Advances in Mathematics*. Kluwer Academic Publishers, Dordrecht, the Netherlands, 2004.

[Hal67] B. Halpern. Fixed points of nonexpanding maps. *Bull. Amer. Math. Soc.*, 73:957–961, 1967.

[Lio77] P.-L. Lions. Approximation de points fixes de contractions. *C. R. Acad. Sci. Paris Sér. A-B*, 284:A1357–A1359, 1977.

[Luk08] D. R. Luke. Finding best approximation pairs relative to a convex and prox-regular set in a Hilbert space. *SIAM J. Optim.*, 19:714–739, 2008.

[Mes96] R. Meshulam. On products of projections. *Discrete Math.*, 154:307–310, 1996.

[Wit92] R. Wittmann. Approximation of fixed points of nonexpansive mappings. *Arch. Math. (Basel)*, 58:486–491, 1992.