Pure and Applied Functional Analysis
Volume 8, Number 2, 2023, 519–532

Yokohama Publishers
ISSN 2189-3764 ONLINE JOURNAL

UNRESTRICTED (RANDOM) PRODUCTS OF PROJECTIONS IN HILBERT SPACE: REGULARITY, ABSOLUTE CONVERGENCE OF TRAJECTORIES AND STATISTICS OF DISPLACEMENTS

C. SINAN GÜNTÜRK AND NGUYEN T. THAO

Dedicated to Professor Ron DeVore on the occasion of his 80th birthday

Abstract. Given a finite list $\mathbf{V} := (V_1, \dots, V_N)$ of closed linear subspaces of a real Hilbert space H, let P_i denote the orthogonal projection operator onto V_i and $P_{i,\lambda} := (1-\lambda)I + \lambda P_i$ denote its relaxation with parameter $\lambda \in [0,2]$, $i = 1, \dots, N$. Under a mild regularity assumption on $\mathbf{V}$ known as "innate regularity" (which, for example, is always satisfied if each V_i has finite dimension or codimension), we show that all trajectories $(x_n)_0^\infty$ resulting from the iteration $x_{n+1} := P_{i_n,\lambda_n}(x_n)$, where the i_n and the λ_n are completely arbitrary other than the assumption that $\{\lambda_n : n \in \mathbb{N}\} \subset [\eta, 2-\eta]$ for some $\eta \in (0,1]$, possess uniformly bounded displacement moments of arbitrarily small orders. In particular, we show that

$$\sum_{n=0}^{\infty} \|x_{n+1} - x_n\|^\gamma \le C\|x_0\|^\gamma \quad \text{for all} \quad \gamma > 0,$$

where $C := C(\mathbf{V}, \eta, \gamma) < \infty$. This result strengthens prior results on norm convergence of these trajectories, known to hold under the same regularity assumption. For example, with $\gamma = 1$, it follows that the displacements series $\sum(x_{n+1} - x_n)$ converges absolutely in H.

Our proof produces an effective bound on the constant $C(\mathbf{V}, \eta, \gamma)$, in particular, as a function of $\gamma \in (0, \infty)$. Utilizing this bound as $\gamma \to 0$, we also derive an effective bound on the distribution function on normalized displacements, and show that their decreasing rearrangement obeys a root-exponential type decay uniformly for all trajectories.

1. Introduction

Starting with the Kaczmarz method [18] and its many variations that have followed, projection algorithms have been employed extensively in convex feasibility problems, in particular linear inverse problems. The literature is highly mature with excellent texts and review articles; see, for example, [7–11].

Consider a real Hilbert space H and a finite list $\mathbf{V} := (V_1, \dots, V_N)$ of closed linear subspaces. For each $i \in [N] := \{1, \dots, N\}$, let $P_i : H \to V_i$ be the orthogonal projection operator onto V_i, and for each $\lambda \in [0,2]$, let $P_{i,\lambda} : H \to H$ be its

2020 *Mathematics Subject Classification.* 41A25, 46C05, 65F10, 65J05.
Key words and phrases. Random products, projection algorithms, alternating projections, Kaczmarz method, chaotic control, innate regularity.

relaxation defined by

$$P_{i,\lambda}(x) := (1-\lambda)x + \lambda P_i(x),\ x \in H. \tag{1.1}$$

We will be concerned with iterations of relaxed projections chosen arbitrarily from the collection

$$\mathcal{P} := \mathcal{P}(\mathbf{V},\eta) := \Big\{ P_{i,\lambda} : 1 \le i \le N, \lambda \in [\eta, 2-\eta] \Big\},\ 0 < \eta \le 1. \tag{1.2}$$

Specifically, for each sequence $(P_{i_n,\lambda_n})_0^\infty$ in $\mathcal{P}$ and starting point $x_0 \in H$, we define a trajectory $(x_n)_0^\infty$ in H via the iteration

$$x_{n+1} := P_{i_n,\lambda_n}(x_n), \quad n \ge 0. \tag{1.3}$$

This setup easily extends to iterations of relaxed projections onto closed affine subspaces with nonempty intersection: Consider $\mathbf{A} := (A_1, \dots, A_N)$ where $A_i = V_i + a_i$ for some $a_i \in H$, $i \in [N]$. Let $P_{A_i,\lambda}$ denote the relaxed projection onto A_i defined analogously to (1.1). If $a \in A_1 \cap \cdots \cap A_N$, then the commutation relation

$$P_{A_i,\lambda}(x) - a = P_{V_i,\lambda}(x-a)$$

which holds uniformly for all i (and λ) implies that the "$\mathbf{A}$-trajectories" and the "$\mathbf{V}$-trajectories" are simply translates of each other. Since we will only be interested in properties of these trajectories (and not the actual feasibility problem), we will remain in the setting of linear subspaces.

The i_n define the so-called "control sequence" of the algorithm, and the λ_n are called relaxation coefficients. In practice the control sequence may be periodic (cyclic), quasi-periodic, stochastic, or greedily determined based on some criterion, such as maximization of $\|x_n - P_i(x_n)\|$, but historically there has also been significant interest in unrestricted (arbitrary) control sequences (also called *random* or *chaotic control*), which is the setting of this paper.

The best known special case of (1.3) involves alternating between two subspaces V_1 and V_2, with no relaxation (i.e., $\lambda_n = 1$ for all n). In this case, von Neumann's celebrated theorem [28] says that x_n converges (in norm) to the orthogonal projection of x_0 onto $V_1 \cap V_2$. This was extended to general N in [17] for cyclic control, and later in [27] for quasi-periodic control. See [22] for a simple geometric proof of von Neumann's theorem.

For unrestricted iterations the situation is more complicated. In [24] norm convergence was shown to hold in finite dimensional spaces. (It was generalized in [1] to include relaxation and convex combinations of projections.) In general Hilbert spaces, weak convergence was shown in [2] and norm convergence was proposed. This question remained unresolved for a long time (see, e.g. [12–15]), and was only answered recently, in the negative: One can find systems $\mathbf{V} = (V_1, V_2, V_3)$ such that for all nonzero initial points x_0, norm convergence fails for some control sequences; see [20, 21].

Nevertheless, norm convergence has been known to hold in general Hilbert spaces under mild regularity assumptions on $\mathbf{V}$ (also called angle criteria); see e.g. [4–6, 23, 25]. In this paper, we will work with the assumption of *innate regularity* which was introduced in [5]. This concept is defined for general convex subsets, but for linear subspaces it reduces to a rather simple form: A list $\mathbf{V} = (V_1, \dots, V_N)$ is innately regular if and only if the complementary angle between $\bigcap_{i \in I} V_i$ and $\bigcap_{i \in J} V_i$ is nonzero for all subsets $I, J \subset [N]$. As a special but important case, any $\mathbf{V}$ for which each V_i is either finite dimensional or finite codimensional is innately regular. (For these facts, see Section 2.1.)

Under the assumption of innate regularity, [5] showed norm convergence of unrestricted iterations of relaxed projections. In a sense, this is the best possible kind of result we can have because unlike cyclic control (or its variants where indices appear with some frequency), it is not possible to obtain any effective convergence rate guarantee for unrestricted iterations once $N \geq 3$ (even in finite dimensions), because one can adversarially slow down the speed of convergence by introducing arbitrarily long gaps between consecutive appearances of any chosen index i while cycling through the remaining indices.

Nevertheless, there is still room for qualitative (and even quantitative) improvements. We show in this paper that the displacements (increments) of the resulting trajectories have bounded moments of all orders. Our main result is the following:

Theorem 1.1. *For any innately regular list of closed linear subspaces* $\mathbf{V} = (V_1, \ldots, V_N)$ *in a real Hilbert space* H *and any* $\eta \in (0, 1]$, *let* $\mathcal{P} := \mathcal{P}(\mathbf{V}, \eta)$ *be defined as in* (1.2). *Then for any* $\gamma > 0$, *there exists a constant* $C = C(\mathbf{V}, \eta, \gamma) < \infty$ *such that for all* $x_0 \in H$ *and all sequences of relaxed projections* $(P_{i_n,\lambda_n})_0^\infty$ *in* $\mathcal{P}$, *the trajectory* $(x_n)_0^\infty$ *defined by* (1.3) *satisfies*

$$\sum_{n=0}^{\infty} \|x_{n+1} - x_n\|^\gamma \leq C \, \|x_0\|^\gamma. \tag{1.4}$$

The case $\gamma = 2$ is elementary and well-known (see, e.g. [7]); it is a fundamental ingredient of the asymptotic regularity property of the trajectories and it holds without the innate regularity assumption on the subspaces (but under the assumption that $\limsup \lambda_k < 2$). The strength of Theorem 1.1 starts with $\gamma = 1$ because it goes beyond the original norm convergence result known to hold for an innately regular $\mathbf{V}$ and shows, in addition, that all trajectories fall into a ball within a proper subspace of convergent sequences in H, namely the space

$$\mathrm{bv}(\mathbb{N}, H) := \left\{ f : \mathbb{N} \to H : \sum_{n=0}^{\infty} \|f(n+1) - f(n)\| < \infty \right\}$$

of bounded variation functions from $\mathbb{N}$ to H. This stronger sense of convergence is sometimes called *absolute convergence* (for sequences – see [19]); it simply means that the displacements series

$$x_0 + \sum_{n=0}^{\infty} (x_{n+1} - x_n)$$

converges *absolutely* (to $\lim x_n$).

As γ is decreased below 1, the strength of Theorem 1.1 goes beyond ensuring bounded total variation of the trajectories. Quantifying the constant $C(\mathbf{V}, \eta, \gamma)$ across all $\gamma > 0$, we also derive an effective bound on the distribution function of the norms of the displacements (see Proposition 5.1) and show that, despite the lack of possibility of establishing any effective convergence rate that holds uniformly for all trajectories, the nth largest displacement is bounded by $c \exp(-\rho n^{1/N})$ uniformly for all trajectories, i.e. the constants c and ρ only depend on $\mathbf{V}$ and η (see Theorem 5.2).

The paper is organized as follows: In Section 2, we review the notion of angle between subspaces and its connection to the notion of innate regularity. Section 3, which is at the heart of the paper, is devoted to geometric properties of successive relaxed projections for innately regular lists which will be needed in our proof of Theorem 1.1 given in Section 4. Section 5 concerns the distribution function of the displacements, and in particular, the derivation of the aforementioned decay bound on their decreasing rearrangements.

2. A review of tools: regularity

We start by recalling the notion of angle between two subspaces introduced in [16]; see [11] for a detailed discussion. Given two subspaces V and W of a Hilbert space H, the angle (also called complementary angle or Friedrichs angle) between V and W is defined to be the unique number $\varphi(V,W) \in [0, \frac{\pi}{2}]$ such that

$$\begin{aligned}(2.1)\qquad \cos\varphi(V,W) = \sup\Big\{&|\langle v,w\rangle| : \\ &v \in V \cap (V\cap W)^\perp,\ w \in W \cap (V\cap W)^\perp,\ \|v\| \le 1 \text{ and } \|w\| \le 1\Big\}.\end{aligned}$$

We note that there are some variations of this definition. Some authors restrict the test vectors v and w in (2.1) to be of unit norm which requires the exclusion of the case of nested subspaces. Meanwhile, some authors allow for nested subspaces, but in this case separately set the angle between them to be 0. Our choice for the definition of angle, as implied by (2.1), produces the value $\pi/2$ for nested subspaces (including the case $V = W$). This apparent discontinuity may seem counter-intuitive. However, there is also an intrinsic discontinuity in the problem we are considering in this paper: Both the limit of x_n defined by (1.3) and the associated total variation (the path length) $\sum \|x_{n+1} - x_n\|$ are discontinuous functions of $\mathbf{V}$. This is most easily seen by considering alternating projections between two lines ℓ_1 and ℓ_2 in $\mathbb{R}^2$ separated by an angle θ. As we let $\theta \to 0^+$, $\lim x_n$ remains fixed at the origin while the path length blows up, but when $\ell_1 = \ell_2$, $\lim x_n$ becomes the orthogonal projection of x_0 on ℓ_1 and the path length becomes finite.

It follows from the discussion in the preceding paragraph and finite dimensional linear algebra that the angle between finite dimensional subspaces is always nonzero. However, the angle between infinite dimensional subspaces could be zero. In general, we have the following characterization of positive angle (see [7, Proposition 5.16] and [11, Theorem 9.35]): For any two closed subspaces V and W in H,

$$(2.2)\qquad \varphi(V,W) > 0 \iff V^\perp + W^\perp \text{ is closed} \iff V + W \text{ is closed}.$$

2.1. **Innate regularity and its angular characterization.** When we have several subspaces in $\mathbf{V}$, a very useful notion of angular separation for convergence of random projections turns out to be *innate regularity*. There are various levels of regularity applicable to general convex sets (see, e.g., [5–7]) but for subspaces they all boil down to a single notion also known as bounded linear regularity, which we will simply call linear regularity in this paper. Following [5], a list of subspaces $\mathbf{V} = (V_1, \dots, V_n)$ is linearly regular if there exists a constant $\kappa < \infty$ such that

$$d(x, V_1 \cap \dots \cap V_N) \le \kappa \max_i d(x, V_i) \quad \text{for all } x \in H. \tag{2.3}$$

Here, $d(x, V)$ stands for the distance between $x \in H$ and the closed subspace V, also equal to $\|x - P_V x\|$ where P_V is the orthogonal projection onto V.

Linear regularity is not hereditary in the sense that a linearly regular list may not pass this property onto its sublists. (See, for example, [26].) A list $\mathbf{V}$ is said to be innately regular if all of its (non-void) sublists are linearly regular.

It is known that (see [7, Theorem 5.19]) $\mathbf{V}$ is linearly regular if and only if $V_1^\perp + \dots + V_N^\perp$ is closed. Therefore, as noted in [5, Fact 3.2]),

$$\mathbf{V} \text{ is innately regular} \iff \sum_{i \in I} V_i^\perp \text{ is closed for all } I \subset [N]. \tag{2.4}$$

Here we take the sum over the empty list to be the trivial (zero) subspace.

For any $I \subset [N]$, let us use the notation

$$V_I := \bigcap_{i \in I} V_i \tag{2.5}$$

where we take $V_\emptyset := H$. We identify V_i with $V_{\{i\}}$. Hence with (2.2) we have

$$\mathbf{V} \text{ is innately regular} \iff \varphi(V_I, V_J) > 0 \quad \text{for all } I, J \subset [N]. \tag{2.6}$$

As a special, but very important case, we note the following observation:

Proposition 2.1. *Suppose that for every $i \in [N]$, V_i has finite dimension or co-dimension. Then $\mathbf{V}$ is innately regular.*

Proof. Let $I, J \subset [N]$. If V_I and V_J both have finite dimension, then $V_I + V_J$, also having finite dimension, is closed. Otherwise, either V_I or V_J has finite co-dimension. Then $V_I^\perp + V_J^\perp$ is closed since the sum of a closed subspace and a finite dimensional subspace is always closed (see [11, Lemma 9.36]). In either case, (2.2) yields $\varphi(V_I, V_J) > 0$. □

2.2. **Quantifying regularity.** Consider two closed subspaces V and W of H. Since the list (V, W) is regular if and only if $\varphi(V, W) > 0$, it is natural to ask how the parameter κ in (2.3) is related to the angle $\varphi(V, W)$. While this specific relation will not be needed in this paper, the answer has a simple form which we note in the next proposition. (See also [3, Proposition 3.9] for an analogous result which involves more than two subspaces.)

Proposition 2.2. *For any two closed subspaces V and W of H,*

$$d(x, V \cap W)\ \sin \varphi(V, W) \le d(x, V) + d(x, W) \quad \textit{for all } x \in H. \tag{2.7}$$

In other words, for $N = 2$, the constant κ in (2.3) can be chosen to be $2/\sin \varphi(V_1, V_2)$.

Proof. Let P_U denote the orthogonal projection operator onto an arbitrary closed subspace U of H. For any $x \in H$, let $u := P_{(V \cap W)^\perp} x$. Noting the relation $P_V u =$

$P_V(x - P_{V\cap W}x) = P_V x - P_{V\cap W}x$, we observe that $P_V u \in V \cap (V\cap W)^\perp$. Similarly, we have $P_W u \in W \cap (V \cap W)^\perp$. Hence, as a consequence of (2.1), we have

$$\sin\varphi(V,W) \le \sin\varphi(P_V u, P_W u) \le \sin\varphi(u, P_V u) + \sin\varphi(u, P_W u),$$

where $\varphi(v,w) := \varphi(\mathbb{R}v, \mathbb{R}w)$ denotes the angle between the lines defined by v and w, and satisfies the triangle inequality. We multiply both sides of this inequality by $\|u\| = d(x, V\cap W)$. Observing that

$$\|u\| \sin\varphi(u, P_V u) = d(u,V) = \|P_{V^\perp} P_{(V\cap W)^\perp} x\| = \|P_{V^\perp} x\| = d(x,V)$$

(and similarly that $\|u\|\sin\varphi(u, P_W u) = d(x,W)$) yields the desired result. □

Remark 2.3. In fact, for distinct closed subspaces V and W, it can be shown that

$$\sin\varphi(V,W) = \inf_{\substack{x\in(V\cap W)^\perp \\ \|x\|=1}} \mathrm{d}(x,V) + \mathrm{d}(x,W). \tag{2.8}$$

3. Lemmas on successive relaxed projections

3.1. **Geometrical observations.** Let P_V be the orthogonal projection operator onto the closed subspace V of H. As before, for any $\lambda \in [0,2]$, we define the relaxed projection of $x \in H$ by $P_{V,\lambda}x := (1-\lambda)x + \lambda P_V x$. The following are elementary derivations:

- (E1) $x - P_{V,\lambda}x = \lambda(x - P_V x)$ so that $\|x - P_{V,\lambda}x\| = \lambda\|x - P_V x\|$,
- (E2) $P_{V,\lambda}x - P_V x = (1-\lambda)(x - P_V x) \perp V$ so that $\|P_{V,\lambda}x\|^2 = \|P_V x\|^2 + (1-\lambda)^2\|x - P_V x\|^2$ and
- (E3) $\|x\|^2 - \|P_{V,\lambda}x\|^2 = \lambda(2-\lambda)\|x - P_V x\|^2$.

It follows trivially from (E3) that $P_{V,\lambda}$ is non-expansive (i.e. $\|P_{V,\lambda}x\| \le \|x\|$ for all $x \in H$). But more is true: if $\lambda \in (0,2)$, then $P_{V,\lambda}$ is strictly contractive provided x is not near V. More precisely, defining the relative distance function $\theta_V : H \to [0,1]$ via

$$\theta_V(x) := \frac{d(x,V)}{\|x\|} = \frac{\|x - P_V x\|}{\|x\|}, \ x \neq 0, \ \text{and} \ \ \theta_V(0) := 0, \tag{3.1}$$

we have, for any $\varepsilon \in [0,1]$,

$$\theta_V(x) \ge \varepsilon \iff \|P_{V,\lambda}x\| \le (1 - \lambda(2-\lambda)\varepsilon^2)^{1/2}\,\|x\|. \tag{3.2}$$

Note that $\lambda(2-\lambda)\varepsilon^2 > 0$ if and only if $\lambda \in (0,2)$ and $\varepsilon > 0$.

The lemma below states that the relaxed projection with respect to W does not increase the relative distance with respect to any subspace V of W:

Lemma 3.1. *Let V and W be any two closed subspaces of H such that $V \subset W$. Then for all $\lambda \in [0,2]$ and $x \in H$,*

$$\theta_V(P_{W,\lambda}x) \le \theta_V(x).$$

Proof. Note that $y := P_{W,\lambda}x$ is a convex combination of x and $P_{W,2}x$. Since $P_{W,2}x$ is the mirror image of x with respect to W, we have $\|P_{W,2}x\| = \|x\|$. More generally,

$P_V P_{W,2} x = P_V x$ implies

$$\begin{aligned} d(P_{W,2}x, V) &= \|(P_W x - x) + (P_W x - P_V x)\| \\ &= \|(P_W x - x) - (P_W x - P_V x)\| \\ &= d(x, V). \end{aligned}$$

(The second equality above uses the fact that $P_W x - x$ is orthogonal to $P_W x - P_V x \in W$.) Hence, by convexity, we have $d(y, V) \le d(x, V)$. Since $P_V y = P_V x$, this implies $\tan \varphi(y, P_V y) \le \tan \varphi(x, P_V x)$ and therefore $\theta_V(y) = \sin \varphi(y, P_V y) \le \sin \varphi(x, P_V x) = \theta_V(x)$. □

Combining Proposition 2.2 and Lemma 3.1 (where (V, W) is replaced by $(V \cap W, W)$) yields the following corollary:

Corollary 3.2. *Let V and W be any two closed subspaces of H such that $\varphi(V, W) > 0$. Then for all $\lambda \in [0, 2]$ and $x \in H$,*

$$\theta_{V\cap W}(P_{W,\lambda}x) \le \theta_{V\cap W}(x) \le \kappa(V, W) \max\Big(\theta_V(x), \theta_W(x)\Big),$$

where $\kappa(V, W) := 2/\sin \varphi(V, W)$.

3.2. **Dynamics of successive relaxed projections.** For any innately regular list $\mathbf{V} = (V_1, \dots, V_N)$, let us define

$$\kappa_* := \kappa_*(\mathbf{V}) := \max_{I,J\subset[N]} \kappa(V_I, V_J) \tag{3.3}$$

where $\kappa(V, W)$ is defined in Corollary 3.2. Note that $2 \le \kappa_* < \infty$.

Now consider any sequence $(x_n)_0^\infty$ of iterates defined by (1.3), i.e. $x_{n+1} := P_{i_n,\lambda_n}(x_n)$, $n \ge 0$. Let

$$W_n := \bigcap_{k=0}^{n} V_{i_k} \qquad \text{and} \qquad N_n := \#\big\{i_k : 0 \le k \le n\big\}, \quad n \ge 0 \tag{3.4}$$

with $W_{-1} := H$. The following lemma will be useful in our analysis.

Lemma 3.3. *We have*

$$\theta_{W_n}(x_{n+1}) \le \kappa_*^{N_n} \max_{0\le k\le n} \theta_{V_{i_k}}(x_k). \tag{3.5}$$

Proof. We begin by applying Corollary 3.2 for $V = V_{i_n}$, $W = W_{n-1}$, $\lambda = \lambda_n$, $x = x_n$. Note that $V_{i_n} \cap W_{n-1} = W_n$. Hence,

$$\theta_{W_n}(x_{n+1}) \le \kappa_* \max\big(\theta_{W_{n-1}}(x_n), \theta_{V_{i_n}}(x_n)\big). \tag{3.6}$$

We can now prove (3.5) by induction. Since $\theta_{W_{-1}}(x_0) = \theta_H(x_0) = 0$, the bound (3.6) yields $\theta_{W_0}(x_1) \le \kappa_* \theta_{V_{i_0}}(x_0)$. With $N_0 = 1$, the statement (3.5) for $n = 0$ follows.

For the induction step, we assume

$$\theta_{W_{n-1}}(x_n) \le \kappa_*^{N_{n-1}} \max_{0\le k\le n-1} \theta_{V_{i_k}}(x_k).$$

There are two cases:

(i) If $N_n = N_{n-1}$, then $W_n = W_{n-1}$. Using Lemma 3.1 (with $V = W_n$ and $W = V_{i_n}$), we have

$$\begin{aligned}\theta_{W_n}(x_{n+1}) &\leq \theta_{W_n}(x_n)\\ &= \theta_{W_{n-1}}(x_n)\\ &\leq \kappa_*^{N_{n-1}} \max_{0\leq k\leq n-1} \theta_{V_{i_k}}(x_k)\\ &\leq \kappa_*^{N_n} \max_{0\leq k\leq n} \theta_{V_{i_k}}(x_k).\end{aligned}$$

(ii) If $N_n = N_{n-1} + 1$, we get from (3.6), along with $\kappa_* \geq 2$, that

$$\begin{aligned}\theta_{W_n}(x_{n+1}) &\leq \kappa_* \max\Big(\kappa_*^{N_{n-1}} \max_{0\leq k\leq n-1} \theta_{V_{i_k}}(x_k), \kappa_*^{N_{n-1}}\theta_{V_{i_n}}(x_n)\Big)\\ &\leq \kappa_*^{N_n} \max_{0\leq k\leq n} \theta_{V_{i_k}}(x_k).\end{aligned}$$

This completes the induction step and the proof. □

Let us make two observations:

Observation 3.4. $\|x_k - x_{k+1}\| = \|x_k - P_{i_k,\lambda_k} x_k\| = \lambda_k \|x_k - P_{i_k} x_k\| = \lambda_k \theta_{V_{i_k}}(x_k)\|x_k\|$.

Observation 3.5. For all $0 \leq m \leq n$,

$$x_{m+1} - x_0 = \sum_{k=0}^{m}(x_{k+1} - x_k) \in V_{i_0}^{\perp} + \cdots + V_{i_m}^{\perp} \subset (V_{i_0} \cap \cdots \cap V_{i_m})^{\perp} \subset W_n^{\perp}.$$

Proposition 3.6. *Let $n \geq 0$. If*

$$\theta_{V_{i_k}}(x_k) < \kappa_*^{-N_n} \quad \textit{for all} \quad 0 \leq k \leq n, \tag{3.7}$$

then either $x_0 = 0$ or $x_0 \notin W_n^{\perp}$.

Proof. Suppose (3.7) holds and $x_0 \in W_n^{\perp}$. We will show that $x_0 = 0$. Using Observation 3.5 for $m = n$, we have $x_{n+1} \in W_n^{\perp}$ so that $d(x_{n+1}, W_n) = \|x_{n+1}\|$ which means either $\theta_{W_n}(x_{n+1}) = 1$, or else $x_{n+1} = 0$. The former is ruled out since Lemma 3.3 combined with the assumption (3.7) implies that $\theta_{W_n}(x_{n+1}) < 1$. Hence we conclude that $x_{n+1} = 0$.

Now, Observation 3.4 combined with $\lambda_k \leq 2-\eta$ and $\theta_{V_{i_k}}(x_k) < \kappa_*^{-N_n} \leq \kappa_*^{-1} \leq \frac{1}{2}$, yields

$$\|x_k - x_{k+1}\| \leq (1 - \tfrac{\eta}{2})\|x_k\| \quad \text{for all } 0 \leq k \leq n. \tag{3.8}$$

Since $1 - \frac{\eta}{2} < 1$ and $x_{n+1} = 0$, (3.8) implies $x_n = 0$. Iterating this process, we obtain $x_n = x_{n-1} = \cdots = x_0 = 0$. □

4. Proof of Theorem 1.1

For any innately regular list $\mathbf{V} = (V_1, \ldots, V_N)$ of closed subspaces in H, $\eta \in (0, 1]$, and $\gamma > 0$, let us define $C(\mathbf{V}, \eta, \gamma)$ to be the infimum of $C \in [0, \infty]$ for which

$$\sum_{k=p}^{q} \|x_{k+1} - x_k\|^{\gamma} \leq C\|x_p\|^{\gamma} \tag{4.1}$$

for all $0 \le p \le q < \infty$ and all trajectories $(x_k)_0^\infty$ satisfying $x_{k+1} = P_{i_k,\lambda_k}(x_k)$ for some sequence of projections $(P_{i_k,\lambda_k})_0^\infty$ in $\mathcal{P}(\mathbf{V},\eta)$. We will show that $C(\mathbf{V},\eta,\gamma) < \infty$ by induction on the length N of the list $\mathbf{V}$.

For the base case $N = 1$, let us consider any trajectory $(x_k)_0^\infty$ associated to a sequence $(P_{1,\lambda_k})_0^\infty$ of projections in $\mathcal{P}((V_1),\eta)$, and define $y_k := x_k - P_1x_0$. Noting that $P_1x_k = P_1x_0$ for all k, as implied by Observation 3.5, we have

$$y_{k+1} = x_{k+1} - P_1x_0 = P_{1,\lambda_k}x_k - P_1x_0 = (1-\lambda_k)(x_k - P_1x_0) = (1-\lambda_k)y_k.$$

Since $|1-\lambda_k| \le 1-\eta$, it then follows that

$$\|y_k\| \le (1-\eta)^k\|y_0\| \le (1-\eta)^k\|x_0\|$$

so that

$$\|x_{k+1} - x_k\| = \|y_{k+1} - y_k\| \le (2-\eta)\|y_k\| \le (2-\eta)(1-\eta)^k\|x_0\|$$

for all k. This gives

$$\sum_{n=0}^{\infty}\|x_{n+1} - x_n\|^\gamma \le \frac{(2-\eta)^\gamma}{1-(1-\eta)^\gamma}\|x_0\|^\gamma;$$

thus we have

$$C((V_1),\eta,\gamma) \le \frac{(2-\eta)^\gamma}{1-(1-\eta)^\gamma} < \infty.$$

For the induction step, assume $C(\mathbf{V},\eta,\gamma) < \infty$ whenever $N < \ell$ and let $\mathbf{V} = (V_1,\dots,V_N)$ be a given innately regular list of closed subspaces in H with $N = \ell$. Given any sequence $(x_k)_0^\infty$ of iterates associated to a sequence of projections $(P_{i_k,\lambda_k})_0^\infty$ in $\mathcal{P}(\mathbf{V},\eta)$, and any choice of integers $0 \le p \le q < \infty$, let

$$W_{p,q} := \bigcap_{k=p}^{q} V_{i_k} \quad \text{and} \quad N_{p,q} := \#\{i_k : p \le k \le q\}.$$

Due to Observation 3.5, we have for any integer $k \in [p, q{+}1]$ that $x_k \in W_{p,q}^\perp + x_p$ and hence $P_{W_{p,q}}x_k = P_{W_{p,q}}x_p$. Therefore, x_k yields the orthogonal decomposition

$$x_k = P_{W_{p,q}}x_p + y_k, \quad \text{where} \quad y_k := P_{W_{p,q}^\perp}x_k, \quad k \in [p, q{+}1]. \tag{4.2}$$

We then have

$$y_{k+1} - y_k = x_{k+1} - x_k \quad \text{and} \quad y_{k+1} = P_{i_k,\lambda_k}y_k, \quad k \in [p,q], \tag{4.3}$$

the second relation being implied by the fact that $P_{W_{p,q}^\perp} = I - P_{W_{p,q}}$ commutes with P_{i_k} (since $W_{p,q} \subset V_{i_k}$) and hence with P_{i_k,λ_k} for each $k \in [p,q]$.

If $y_p = 0$, then $y_{k+1} = 0$ for all $k \in [p,q]$ and therefore (4.1) holds trivially (with $C = 0$), so let us consider the case $y_p \ne 0$. Note that we also have $y_p \in W_{p,q}^\perp$. Hence Proposition 3.6 implies (by contrapositive) that for some $k \in [p,q]$ we must have $\theta_{V_{i_k}}(y_k) \ge \kappa_*^{-N_{p,q}} > \varepsilon_*$, where

$$\varepsilon_* = \varepsilon_*(\mathbf{V}) := \frac{1}{2}\kappa_*^{-\ell}.$$

Let us enumerate the set $\{k \in [p,q] : \theta_{V_{i_k}}(y_k) > \varepsilon_*\}$ as an increasing sequence $r_1 < \dots < r_L$. This results in a segmentation of $[p,q]$ in the form

$$[p_0,q_0] \cup \{r_1\} \cup [p_1,q_1] \cup \dots \cup \{r_L\} \cup [p_L,q_L]$$

where $p_0 := p$, $q_L := q$, and for all $j = 1, \dots, L$ we have $p_j := r_j+1$ and $q_{j-1} := r_j-1$. (It is certainly possible that a given segment $[p_j, q_j]$ is empty; this happens precisely when $q_j = p_j - 1$.) Let us also define $r_0 := p$. Note that r_0 need not be distinct from r_1.

We can safely ignore any j for which $[p_j, q_j] = \emptyset$. For $[p_j, q_j] \neq \emptyset$, we have $\theta_{V_{i_k}}(y_k) \leq \varepsilon_* < \kappa_*^{-N_{p_j,q_j}}$ for all $k \in [p_j, q_j]$, so Proposition 3.6 implies that either $y_{p_j} = 0$ (in which case $y_k = 0$ for all $k \geq p_j$ and we are done), or else $y_{p_j} \notin W^\perp_{p_j,q_j}$. Noting that $y_{p_j} \in W^\perp_{p,q} \supset W^\perp_{p_j,q_j}$, we conclude in this second case that $W_{p,q}$ must be a *proper* subspace of W_{p_j,q_j} so that $N_{p_j,q_j} < N_{p,q} \leq \ell$. In other words, the segment $(y_k)_{p_j}^{q_j}$ can be associated with a proper sublist of $\mathbf{V}$. (By a proper sublist, we mean any $(U_1, \dots, U_M)$ where $M < N$ and every U_i is equal to some V_j.) Since $N = \ell$ and therefore any proper sublist $\mathbf{U}$ of $\mathbf{V}$ has fewer than ℓ subspaces, we can now employ our induction hypothesis. Let us denote by $D := D(\mathbf{V}, \eta, \gamma)$ the maximum value of $C(\mathbf{U}, \eta, \gamma)$ over all proper sublists $\mathbf{U}$ of $\mathbf{V}$. Clearly D is finite, and we have

$$\sum_{k=p_j}^{q_j} \|y_{k+1} - y_k\|^\gamma \leq D\|y_{p_j}\|^\gamma. \tag{4.4}$$

For the terms corresponding to $k = r_j$, we use the generic bound $\|y_{r_j+1} - y_{r_j}\| \leq (2-\eta)\|y_{r_j}\|$ which follows from $y_{k+1} = P_{i_k,\lambda_k} y_k$ and (E1) in Section 3.1. Combining this with (4.4), we obtain

$$\sum_{k=p}^{q} \|y_{k+1} - y_k\|^\gamma \leq D\sum_{j=0}^{L} \|y_{p_j}\|^\gamma + (2-\eta)^\gamma \sum_{j=1}^{L} \|y_{r_j}\|^\gamma. \tag{4.5}$$

The instances $k = r_j$, $j = 1, \dots, L$, are characterized by the event $\theta_{V_{i_k}}(y_k) > \varepsilon_*$ which yield strict geometric decay of the norms via (3.2). Defining

$$\beta_* := \beta_*(\mathbf{V}, \eta) := \left(1 - \eta(2-\eta)\varepsilon_*^2\right)^{1/2}, \tag{4.6}$$

we have

$$\|y_{p_j}\| = \|y_{r_j+1}\| \leq \beta_*\|y_{r_j}\|, \; j = 1, \dots, L. \tag{4.7}$$

Combining this with the monotonicity of $(\|y_k\|)$, we obtain

$$\|y_{r_j}\| \leq \beta_*^{j-1}\|y_{r_1}\| \leq \beta_*^{j-1}\|y_p\|, \; j = 1, \dots, L. \tag{4.8}$$

Injecting (4.7) and (4.8) into (4.5), we obtain

$$\begin{aligned}\sum_{k=p}^{q} \|y_{k+1} - y_k\|^\gamma &\leq D\|y_p\|^\gamma + (D\beta_*^\gamma + (2-\eta)^\gamma)\sum_{j=1}^{L} \|y_{r_j}\|^\gamma \\ &\leq \left(D + \frac{D\beta_*^\gamma + (2-\eta)^\gamma}{1-\beta_*^\gamma}\right)\|y_p\|^\gamma \;=\; C\|y_p\|^\gamma\end{aligned} \tag{4.9}$$

where $C := \frac{D+(2-\eta)^\gamma}{1-\beta_*^\gamma}$. Since $\|x_{k+1} - x_k\| = \|y_{k+1} - y_k\|$ and $\|y_k\| \leq \|x_k\|$ for all k, the desired conclusion (4.1) follows, and with the same constant. Notice, in particular, that

$$C(\mathbf{V},\eta,\gamma) \le \frac{D(\mathbf{V},\eta,\gamma) + (2-\eta)^\gamma}{1 - [\beta_*(\mathbf{V},\eta)]^\gamma} < \infty. \quad \square$$

An effective bound for $C(\mathbf{V},\eta,\gamma)$. Given an innately regular list $\mathbf{V} = (V_1,\dots,V_N)$, and for any $m = 1,\dots,N$, let us define C_m to be the maximum of $C(\mathbf{U},\eta,\gamma)$ over all sublists $\mathbf{U}$ of $\mathbf{V}$ of length at most m. If C_{m+1} is realized by the list $\mathbf{U}$, then by the same reasoning given in the proof of Theorem 1.1 and also noting that $\beta_*(\mathbf{U},\eta) \le \beta_*(\mathbf{V},\eta)$, we have

$$C_{m+1} = C(\mathbf{U},\eta,\gamma) \le \frac{D(\mathbf{U},\eta,\gamma) + (2-\eta)^\gamma}{1 - [\beta_*(\mathbf{U},\eta)]^\gamma} \le \frac{C_m + (2-\eta)^\gamma}{1-\beta_*^\gamma},$$

where $\beta_* := \beta_*(\mathbf{V},\eta)$. Iterating this recursive inequality together with the bound $C_1 \le \frac{(2-\eta)^\gamma}{1-(1-\eta)^\gamma}$ as given in the proof of Theorem 1.1, we obtain

$$\begin{aligned} C(\mathbf{V},\eta,\gamma) = C_N &\le \left(\frac{1}{1-\beta_*^\gamma}\right)^{N-1} C_1 + \left[\left(\frac{1}{1-\beta_*^\gamma}\right)^{N-1} - 1\right] \frac{(2-\eta)^\gamma}{\beta_*^\gamma} \\ &< \left(\frac{1}{1-\beta_*^\gamma}\right)^{N-1} \frac{(2-\eta)^\gamma}{1-(1-\eta)^\gamma} + \left(\frac{1}{1-\beta_*^\gamma}\right)^{N-1} \frac{(2-\eta)^\gamma}{\beta_*^\gamma} \\ &< \left(\frac{1}{1-\beta_*^\gamma}\right)^{N-1} (2-\eta)^\gamma \left(\frac{1}{1-\beta_*^\gamma} + \frac{1}{\beta_*^\gamma}\right) \\ &= \left(\frac{1}{1-\beta_*^\gamma}\right)^{N} \frac{(2-\eta)^\gamma}{\beta_*^\gamma}, \end{aligned} \tag{4.10}$$

where in the last inequality we also made use of the fact that $1 > \beta_* > (1-\eta(2-\eta))^{1/2} = 1-\eta$.

5. Statistics of displacements via moment bounds

While it is possible to arrange control sequences that result in arbitrarily slow convergence of (x_n), the moment bounds of Theorem 1.1 place strong restrictions on the number of displacements exceeding any given value. In this section we will quantify this proposition.

Let us fix $\mathbf{V}$ and $\eta \in (0,1]$ according to Theorem 1.1 and consider any trajectory $(x_n)_0^\infty$ where $x_0 \ne 0$. Since $\|x_{n+1} - x_n\| \le (2-\eta)\|x_n\| \le (2-\eta)\|x_0\|$, let us define

$$\delta_n := \frac{\|x_{n+1}-x_n\|}{(2-\eta)\|x_0\|}, \ n \in \mathbb{N},$$

as a normalized measure of the displacements. For any $\tau \in [0,1]$, let us also define

$$S(\tau) := \#\Lambda_\tau, \quad \text{where} \quad \Lambda_\tau := \Big\{ n \in \mathbb{N} : \delta_n \ge \tau \Big\}. \tag{5.1}$$

The next proposition shows that $S(\tau) = O(|\log\tau|^N)$ as $\tau \to 0$.

Proposition 5.1. *Assume the hypothesis of Theorem 1.1. Let β_* be defined as in (4.6) and $S(\tau)$ as above where $x_0 \neq 0$. Then for all $\tau \in (0,1]$ we have*

$$S(\tau) \leq e^N \left(1 + \frac{\log(\beta_* \tau)}{N \log \beta_*}\right)^N. \tag{5.2}$$

Proof. Let $\tau \in (0,1]$ be arbitrary. With Theorem 1.1 we have

$$(2-\eta)^{-\gamma} C(\mathbf{V}, \eta, \gamma) \geq \sum_{n=0}^{\infty} \frac{\|x_{n+1} - x_n\|^\gamma}{(2-\eta)^\gamma \|x_0\|^\gamma} = \sum_{n=0}^{\infty} \delta_n^\gamma \geq \sum_{n \in \Lambda_\tau} \delta_n^\gamma \geq \tau^\gamma S(\tau).$$

This inequality holds for all $0 < \gamma < \infty$, so

$$\begin{aligned} S(\tau) &\leq \inf_{0<\gamma<\infty} \tau^{-\gamma} (2-\eta)^{-\gamma} C(\mathbf{V}, \eta, \gamma) \\ &\leq \inf_{0<\gamma<\infty} \tau^{-\gamma} \left(\frac{1}{1 - \beta_*^\gamma}\right)^N \frac{1}{\beta_*^\gamma}, \end{aligned} \tag{5.3}$$

where in the last step we have used the explicit bound derived in (4.10).

To ease our computation, we slightly relax the upper bound. Note that, for any $0 < r < 1$,

$$\frac{1}{1 - r^\gamma} = 1 + \frac{1}{r^{-\gamma} - 1} \leq 1 + \frac{1}{\log r^{-\gamma}} = 1 + \frac{1}{\gamma \log r^{-1}},$$

so that

$$S(\tau) \leq \inf_{0<\gamma<\infty} (\tau \beta_*)^{-\gamma} \left(1 + \frac{1}{\gamma \log \beta_*^{-1}}\right)^N. \tag{5.4}$$

For any $t \in (0,1)$, noting that $t^{-\gamma}\gamma^{-N}$ is minimized at $\gamma_t := N/\log t^{-1}$, we set $\gamma = \gamma_{\tau\beta_*}$ to produce a convenient upper bound for the right hand side of (5.4). The desired bound of (5.2) follows immediately once we observe $(\tau\beta_*)^{-N/\log(\tau\beta_*)^{-1}} = e^N$. ☐

As an immediate application of this proposition, we will derive an explicit decay estimate for the decreasing rearrangement of $(\delta_n)_0^\infty$ which we denote by $(\delta_n^*)_0^\infty$. Recall that this is the (unique) sequence

$$\delta_0^* \geq \delta_1^* \geq \cdots$$

satisfying $\delta_n^* = \delta_{\pi(n)}$ for some bijection $\pi : \mathbb{N} \to \mathbb{N}$.

Theorem 5.2. *Assume the hypothesis of Proposition 5.1. Then*

$$\delta_n^* < c_* \exp(-\rho_* n^{1/N}) \quad \textit{for all} \quad n \geq 0, \tag{5.5}$$

where $\rho_* := \frac{N}{e} \log \beta_*^{-1} > 0$ *and* $c_* := \beta_*^{-N-1}$.

Proof. The result holds trivially when $\delta_n^* = 0$, so it suffices to consider the nonzero values only. Note that

$$S(\delta_n^*) = \#\left\{k \in \mathbb{N} : \delta_k \geq \delta_n^*\right\} = \#\left\{k \in \mathbb{N} : \delta_k^* \geq \delta_n^*\right\} \geq n + 1$$

so that $n < S(\delta_n^*)$ which implies, when combined with Proposition 5.1,

$$n < e^N \left(1 + \frac{\log(\beta_* \delta_n^*)}{N \log \beta_*}\right)^N.$$

The desired bound (5.5) then easily follows from this inequality by solving for δ_n^*. $\square$

Acknowledgment. The authors thank Heinz Bauschke for a very helpful correspondence and Halyun Jeong for an insightful inquiry which triggered a notable strengthening of our result. The authors would also like to thank the referees for their helpful feedback and additional references.

References

[1] R. Aharoni and Y. Censor, *Block-iterative projection methods for parallel computation of solutions to convex feasibility problems*, Linear Algebra Appl. **120** (1989), 165–175.

[2] I. Amemiya and T. Ando, *Convergence of random products of contractions in Hilbert space*, Acta Sci. Math. (Szeged) **26** (1965), 239–244.

[3] C. Badea, S. Grivaux and V. Müller, *The rate of convergence in the method of alternating projections*, St. Petersburg Math. J. **23** (2012), 413–434.

[4] C. Bargetz, J. Klemenc, S. Reich and N. Skorokhod, *On angles, projections and iterations*, Linear Algebra Appl. **603** (2020), 41–56.

[5] H. H. Bauschke, *A norm convergence result on random products of relaxed projections in Hilbert space*, Trans. Amer. Math. Soc. **347** (1995),1365–1373.

[6] H. H. Bauschke, *Projection algorithms: results and open problems*, Studies in Computational Mathematics, vol. 8, Elsevier, 2001, pp. 11–22.

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

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

[9] A. Cegielski, *Iterative Methods for Fixed Point Problems in Hilbert Spaces*, vol. 2057, Springer, 2012.

[10] P. L. Combettes, *The convex feasibility problem in image recovery*, Advances in imaging and electron physics, vol. 95, Elsevier, 1996, pp. 155–270.

[11] F. Deutsch, *Best Approximation in Inner Product Spaces*, CMS Books in Mathematics, vol. 7, Springer, New York, 2001.

[12] J. Dye, M. A. Khamsi and S. Reich, *Random products of contractions in Banach spaces*, Trans. Amer. Math. Soc. **325** (1991), 87–99.

[13] J. Dye, *A generalization of a theorem of Amemiya and Ando on the convergence of random products of contractions in Hilbert space*, Integral Equations Operator Theory **12** (1989), 155–162.

[14] J. M. Dye and S. Reich, *On the unrestricted iteration of projections in Hilbert space*, J. Math. Anal. Appl. **156** (1991), 101–119.

[15] J. M. Dye and S. Reich, *Unrestricted iterations of nonexpansive mappings in Hilbert space*, Nonlinear Anal. **18** (1992), 199–207.

[16] K. Friedrichs, *On certain inequalities and characteristic value problems for analytic functions and for functions of two variables*, Trans. Amer. Math. Soc. **41** (1937), 321–364.

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

[18] S. Kaczmarz, *Angenäherte Auflösung von Systemen linearer Gleichungen*, Bull. Pol. Acad. Sci. Math. **A35** (1937), 355–357.

[19] K. Knopp, *Infinite Sequences and Series*, Courier Corporation, 1956.

[20] E. Kopecká and V. Müller, *A product of three projections*, Studia Mathematica **2** (2014), 175–186.

[21] E. Kopecká and A. Paszkiewicz, *Strange products of projections*, Israel J. Math. **219** (2017), 271–286.

[22] E. Kopecká and S. Reich, *A note on the von Neumann alternating projections algorithm*, J. Nonlinear Convex Anal. **5** (2004), 379–386.

[23] I. Oppenheim, *Angle criteria for uniform convergence of averaged projections and cyclic or random products of projections*, Israel J. Math. **223** (2018), 343–362.

[24] M. Práger, *Uber ein Konvergenzprinzip im Hilbertschen Raum*, Czechoslovak Math. J **10** (1960), 271–272.

[25] E. Pustylnik, S. Reich and A. J. Zaslavski, *Convergence of non-periodic infinite products of orthogonal projections and nonexpansive operators in Hilbert space*, J. Approx. Theory **164** (2012), 611–624.

[26] S. Reich and A. J. Zaslavski, *An example concerning bounded linear regularity of subspaces in Hilbert space*, Bull. Aust. Math. Soc. **89** (2014), 217–226.

[27] M. Sakai, *Strong convergence of infinite products of orthogonal projections in Hilbert space*, Appl. Anal. **59** (1995), 109–120.

[28] J. von Neumann, *On rings of operators. Reduction theory*, Ann. of Math. (1949), 401–485.

Manuscript received January 18 2022
revised June 10 2022

C. S. Güntürk
Courant Institute, New York University, NY, USA
E-mail address: gunturk@cims.nyu.edu

N. T. Thao
City College, City University of New York, NY, USA
E-mail address: tnguyen@ccny.cuny.edu