

Reading OpenAI's fixed-point argument in reflexive Banach spaces

Ordered trees, finite switching, and an RNP extension

Version 1

Prepared: 16 October 2026

Publication responsibility: Ryotaro Tanaka

This note reconstructs the argument in the cited OpenAI preprint and records an RNP extension; it is not an announcement of an independent solution.

Publication responsibility covers release, versioning and corrections; it does not claim authorship of the original proof.

Abstract

This note gives a detailed reconstruction of the ordered-tree and finite-switching argument in OpenAI's preprint on nonexpansive fixed points in reflexive Banach spaces. The proof is organized through root responses of a completed tree: local changes of leaf data are propagated to the root, where exact parent-child additivity produces a bounded convex tree after normalization. We make explicit the fixed-node weak compactness used in the limiting argument and record that the final step can be completed by a small-slice argument. In this formulation, the same mechanism yields the weak fixed point property for Banach spaces with the Radon–Nikodým property. No priority claim is made for this observation.

Contents

1	The theorem and the root-response mechanism	2
1.1	The geometric obstruction supplied by the RNP	3
1.2	The exact convex tree we want at the end	3
1.3	Why the proof keeps looking at the root	4
1.4	How the remaining sections build the two trees	5
2	Classical preparation and the normalized counterexample	5
2.1	The Goebel–Karlovitze diameter lemma	5
2.2	One functional for a compact set of comparison points	7
2.3	Reduction without assuming reflexivity	7
3	An anchor for convex combinations of resolvent images	8
3.1	Resolvents and the two estimates that will be reused	8
3.2	Finite simultaneous approximation	9
3.3	A compact semigroup and the anchor point	10
4	One fixed tree, evaluated from the leaves upwards	12
4.1	The parent relation and the two kinds of weight	12
4.2	Prescribed values at one level; determined values above it	13
4.3	Fixed labels and two useful one-node estimates	15

5	Constructing a tree before its new coefficients are known	16
5.1	The comparison output and the compactness used here	16
5.2	The construction and all its choices	17
6	Finite switching and the differences of all blocks	22
6.1	One common input for each fixed depth	23
6.2	The switching states and what their root values mean	23
6.3	The errors in a uniformly assigned subtree	24
6.4	Bounds and exact addition for the block differences	26
7	One exact convex tree from the finite block differences	29
7.1	A weakly compact set for each fixed node	29
7.2	One subsequence for all nodes	29
8	A small slice gives the contradiction	31
8.1	Vanishing on fixed vectors and the weighted deficit	31
8.2	The closed convex hull cannot have a small slice	32
8.3	The commuting-family consequence	34

1. The theorem and the root-response mechanism

OpenAI’s preprint [10] proves that every real reflexive Banach space has the fixed point property for nonexpansive mappings in its original norm. The purpose of this note is to reconstruct the mechanism of that proof in a form that exposes the roles of the ordered tree, finite switching, and the root-response calculation. We first state the slightly more general formulation obtained by carrying the same mechanism through a fixed-node compactness argument and a final small-slice contradiction.

A map $F : C \rightarrow C$ is *nonexpansive* if $\|F(a) - F(b)\| \leq \|a - b\|$ for all $a, b \in C$. The weak fixed point property concerns such maps on nonempty weakly compact convex sets; the fixed point property concerns all nonempty closed bounded convex sets. These two domain conditions coincide in reflexive spaces, but not in general.

Theorem 1.1. *Let X be a real Banach space with the Radon–Nikodým property, and let $C \subset X$ be nonempty, weakly compact, and convex. Every nonexpansive selfmap $F : C \rightarrow C$ has a fixed point.*

The map and all distances are measured in the given norm. There is no renorming and no weak-continuity assumption on F .

Corollary 1.2 (The reflexive-space conclusion). *Let X be a real reflexive Banach space. Every nonexpansive selfmap of a nonempty closed bounded convex subset of X has a fixed point.*

Derivation from Theorem 1.1. A reflexive space has the Radon–Nikodým property [9, Corollary 2.11], and each of its closed bounded convex subsets is weakly compact. Thus Theorem 1.1 applies. $\square$

The expository target is Corollary 1.2. The anchor, ordered tree, and finite-switching construction in Sections 2 to 6 reconstruct the argument in the cited OpenAI preprint. The fixed-node compactness step in Section 7 and the slice argument in Section 8 record the same mechanism

in the RNP setting, replacing the source proof's final use of weak compactness of the entire unit ball. This origin boundary is kept visible throughout the note.

Kirk's classical theorem derives fixed points from normal structure on weakly compact convex sets [6]. Related RNP-based fixed-point conclusions are known in more specialized settings, including preduals associated with locally compact groups [7]. Wiśnicki's nonlinear Ryll–Nardzewski theorem treats nonexpansive, norm-distal dynamical systems on RNP sets [11]; the additional distality hypothesis makes that result different from the arbitrary single nonexpansive selfmap considered here. We record the present argument in this slightly more general form and make no priority claim for the RNP observation.

Reading guide. Section 1 states the target and explains the root-response viewpoint. Sections 2 and 3 collect the fixed-point reduction and construct the anchor. Sections 4 and 5 define and build the ordered tree, while Section 6 contains the common-endpoint switching calculation. Readers following the extension should then continue through the fixed-node diagonal limit in Section 7 and the exact slice contradiction in Section 8. None of those mechanisms is abbreviated in this presentation.

1.1 The geometric obstruction supplied by the RNP

For a nonempty bounded set $D \subset X$, a functional $g \in X^*$, and $\varepsilon > 0$, write

$$S(D, g, \varepsilon) = \left\{ d \in D : g(d) > \sup_{z \in D} g(z) - \varepsilon \right\}. \quad (1.1)$$

We use the standard dentability characterization of the Radon–Nikodým property: every nonempty closed bounded convex set in an RNP space has slices of arbitrarily small norm diameter [9, Remark 2.2 and Theorem 2.5]. In the form needed below,

$$\begin{aligned} &X \text{ has the RNP, } D \subset X \text{ is nonempty, closed, bounded, and convex} \\ \implies &\text{for every } \eta > 0 \text{ there is a nonempty slice } S \subset D \text{ with } \text{diam } S < \eta. \end{aligned} \quad (1.2)$$

Starting from a fixed-point-free map, we shall construct a particular closed bounded convex set all of whose nonempty slices have diameter at least one. That will contradict (1.2).

1.2 The exact convex tree we want at the end

The final object is not the original tree of nonexpansive maps. It is a bounded family of vectors (d_v) on all nodes and functionals (f_v) on the nonroot nodes of the same tree, in the normalized counterexample space Y introduced in Subsection 2.3. Let L_k be the finite level of depth k . If v is a child of u , write $v^\uparrow = u$. Each edge carries a local coefficient $\alpha_v \geq 0$ with

$$\sum_{v^\uparrow=u} \alpha_v = 1.$$

The path weights are

$$W_r = 1, \quad W_v = W_{v^\uparrow} \alpha_v, \quad \sum_{v \in L_k} W_v = 1.$$

We shall obtain

$$d_u = \sum_{v^\uparrow=u} \alpha_v d_v, \quad \|d_v\| \leq 1, \quad \|f_v\| \leq 1 \quad (v \neq r). \quad (1.3)$$

The same functionals will have two apparently opposing properties. For every fixed $d \in Y$,

$$\max_{v \in L_k} |f_v(d)| \longrightarrow 0, \quad (1.4)$$

whereas on the vectors moving with the level,

$$\sum_{v \in L_k} W_v (1 - f_v(d_v)) \longrightarrow 0. \quad (1.5)$$

Thus every fixed vector eventually becomes invisible to all functionals at level k , while the weighted average of the values $f_v(d_v)$ tends to one.

The last two properties alone are not contradictory: an orthonormal sequence and its coordinate functionals already exhibit a similar “moving vector versus fixed vector” phenomenon. The decisive extra condition is the exact convex decomposition (1.3). If a slice of the closed convex hull contains one d_u , that identity forces a fixed positive proportion of the deeper descendant mass to remain in the same slice. On a slice of diameter less than one, (1.4) applied to the fixed vector d_u then keeps those descendants’ values $f_v(d_v)$ uniformly below one. This contradicts (1.5). The complete calculation appears in Section 8.

1.3 Why the proof keeps looking at the root

Most tree constructions are first imagined from the root towards the leaves. The decisive calculation here runs in the opposite direction. Once one tree has been completed, a finite level is treated as an input layer. At every internal node we take the convex average of its child values and apply a fixed nonexpansive map:

$$H_u = G_u \left(\sum_{v \uparrow = u} \alpha_v H_v \right). \quad (1.6)$$

Prescribing values at the leaves therefore determines one value H_r at the root. The root is the output port of this nonlinear finite computation.

Now fix a depth N and a node v . We do not alter the nodes, edges, maps, or coefficients. We change only the values assigned to the depth- N leaves below v : their labels are replaced by one common point b_N . All internal values are then recomputed upwards by the same rule (1.6). Denote the root value before the whole block is changed by U_v and the root value after it is changed by V_v . The datum attached to v is the root response

$$\Delta_v^{(N)} := U_v - V_v. \quad (1.7)$$

Thus a node specifies a local block of inputs, while $\Delta_v^{(N)}$ records the global effect of changing that block.

All blocks at the same depth are read from one right-to-left switching sequence and use the same point b_N . A parent’s leaf block is the consecutive union of the blocks below its children. Hence the intermediate root values cancel in a telescoping sum and give the exact identity

$$\Delta_u^{(N)} = \sum_{v \uparrow = u} \Delta_v^{(N)}. \quad (1.8)$$

This is not linearity of the maps G_u ; it is additivity of differences of actual root outputs in one common switching experiment. Nonexpansiveness also gives

$$\|\Delta_v^{(N)}\| \leq W_v. \quad (1.9)$$

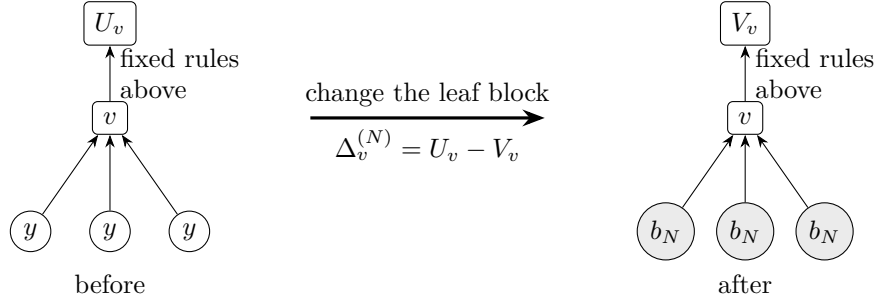


Figure 1: Schematic root-response experiment. The omitted ancestors, the tree, and all upward rules are fixed. Only the depth- N leaf values below v are changed. The vector assigned to v is the resulting change in the root output, not the value at the internal node v .

For $W_v > 0$, the normalized response $d_v^{(N)} = \Delta_v^{(N)}/W_v$ lies in the unit ball, and (1.8) becomes

$$d_u^{(N)} = \sum_{v^\dagger=u} \alpha_v d_v^{(N)}.$$

Weak compactness at each fixed node, followed by one diagonal subsequence, then makes the finite-depth vectors simultaneous over all nodes and produces (1.3).

The approximation machinery has a different role. It does not create the exact addition law: that law already follows from the common switching sequence. The error estimates are used to arrange

$$f_v(\Delta_v^{(N)}) \geq W_v - 5e_v$$

with a summable error budget. After normalization and passage to the fixed- node limits, this yields the average detection property (1.5) without ever requiring e_v/W_v to be small.

1.4 How the remaining sections build the two trees

Section 2 reduces a counterexample to a separable minimal invariant set of diameter one and proves the separation lemma used later. Section 3 constructs one anchor point and arbitrarily accurate images returning to it in convex combinations. Section 4 defines the upward evaluation of a completed tree and proves its weighted sensitivity estimate. Section 5 chooses the maps, labels, coefficients, and functionals; its parameterized comparison family handles coefficients that are not yet known. Section 6 performs the one common switching experiment at each finite depth and records every block response. Section 7 turns those responses into the exact convex tree, and Section 8 gives the slice contradiction.

2. Classical preparation and the normalized counterexample

We use Hahn–Banach separation, weak lower semicontinuity of the norm, and the Eberlein–Šmulian theorem in their usual Banach-space forms [8]. Write $\overline{\text{co}}A$ for the norm-closed convex hull and $\overline{A}^w$ for weak closure. Weak convergence is denoted by $u_n \rightharpoonup u$, meaning $f(u_n) \rightarrow f(u)$ for every continuous linear functional f . The closed unit ball of a space Z is B_Z .

2.1 The Goebel–Karlóvitz diameter lemma

Lemma 2.1 (Goebel–Karlóvitz). *Let K be a nonempty weakly compact convex subset of a real Banach space, and let $T : K \rightarrow K$ be nonexpansive. Suppose K has no proper nonempty*

closed convex T -invariant subset and $\text{diam } K = d > 0$. For every sequence $(u_n) \subset K$ with $\|Tu_n - u_n\| \rightarrow 0$,

$$\|u_n - w\| \rightarrow d \quad (w \in K).$$

Proof. This is the diameter lemma of [4, 5]. Put $H = \overline{\text{co}}T(K)$. The inclusions

$$T(K) \subset K, \quad H \subset K, \quad T(H) \subset T(K) \subset H$$

follow from invariance and closed convexity of K . Thus H is a nonempty closed convex invariant subset, and minimality gives

$$K = \overline{\text{co}}T(K). \quad (2.1)$$

For $z \in K$ define

$$s(z) = \sup_{w \in K} \|z - w\|, \quad r(z) = \limsup_n \|z - u_n\|.$$

The triangle inequality gives

$$\|z - w\| \leq \|z - z'\| + \|z' - w\|.$$

Taking the supremum over w , or setting $w = u_n$ and taking the limsup, then interchanging z, z' , yields

$$|s(z) - s(z')| \leq \|z - z'\|, \quad |r(z) - r(z')| \leq \|z - z'\|.$$

Both functions are also convex. For r , take the limsup in

$$\|tz + (1-t)z' - u_n\| \leq t\|z - u_n\| + (1-t)\|z' - u_n\| \quad (0 \leq t \leq 1);$$

for s , take the supremum in the same inequality with u_n replaced by w .

To use (2.1), recall the elementary convex-hull calculation

$$\left\| a - \sum_i \lambda_i w_i \right\| \leq \sum_i \lambda_i \|a - w_i\| \leq \sup_{w \in W} \|a - w\| \quad \left(w_i \in W, \lambda_i \geq 0, \sum_i \lambda_i = 1 \right).$$

Continuity in the second argument then gives $\sup_{w \in \overline{\text{co}}W} \|a - w\| = \sup_{w \in W} \|a - w\|$. Consequently,

$$\begin{aligned} s(Tz) &= \sup_{w \in K} \|Tz - w\| = \sup_{w \in \overline{\text{co}}T(K)} \|Tz - w\| \\ &= \sup_{w \in T(K)} \|Tz - w\| = \sup_{u \in K} \|Tz - Tu\| \leq \sup_{u \in K} \|z - u\| = s(z), \\ r(Tz) &= \limsup_n \|Tz - u_n\| \\ &\leq \limsup_n (\|Tz - Tu_n\| + \|Tu_n - u_n\|) \leq \limsup_n \|z - u_n\| = r(z). \end{aligned}$$

Here the last line uses both nonexpansiveness and $\|Tu_n - u_n\| \rightarrow 0$. Every nonempty sublevel set of s or r is therefore closed, convex, and invariant. A value strictly between two values of either function would define a proper such subset. Minimality forces both functions to be constant. Since $\sup_K s = d$, write

$$s \equiv d, \quad r \equiv R \leq d.$$

By Eberlein–Šmulian, $u_{n_j} \rightharpoonup a \in K$ along a subsequence. For every $w \in K$, weak lower semicontinuity gives

$$\|w - a\| \leq \liminf_j \|w - u_{n_j}\| \leq \limsup_n \|w - u_n\| = R.$$

Taking the supremum over w gives $d = s(a) \leq R \leq d$, so $R = d$. This reasoning applies to any subsequence, since its fixed point defects also tend to zero. If the asserted limit failed at w , some subsequence would satisfy $\|u_{n_j} - w\| \leq d - \eta$ for a fixed $\eta > 0$, contradicting its limsup d . $\square$

2.2 One functional for a compact set of comparison points

Lemma 2.2 (Simultaneous separation). *Let K, T satisfy the hypotheses of Lemma 2.1, with $\text{diam } K = 1$, and let Z be their ambient Banach space. Fix a nonempty norm-compact set $E \subset K$ and $\varepsilon > 0$. There is $\delta > 0$ such that*

$$\begin{aligned} y \in K, \quad \|Ty - y\| \leq \delta &\implies \\ \exists f_y \in B_{Z^*} \quad f_y(y - z) \geq 1 - \varepsilon &\text{ for every } z \in E. \end{aligned} \quad (2.2)$$

Thus δ works for every eligible y . The functional is chosen after y , but the same f_y works for all $z \in E$.

Proof. Choose an $\varepsilon/2$ -net $a_1, \dots, a_m$ in E and set $a_* = m^{-1} \sum_{i=1}^m a_i \in K$. By Lemma 2.1, some $\delta > 0$ satisfies

$$\|Ty - y\| \leq \delta \implies \|y - a_*\| \geq 1 - \frac{\varepsilon}{2m}. \quad (2.3)$$

Indeed, otherwise for each n we could choose $y_n \in K$ with $\|Ty_n - y_n\| \leq 1/n$ and $\|y_n - a_*\| < 1 - \varepsilon/(2m)$, contradicting Lemma 2.1 with $w = a_*$.

Fix y satisfying (2.3). Take a norming functional $f_y \in B_{Z^*}$ for $y - a_*$. Since $y, a_i \in K$ and $\text{diam } K = 1$,

$$1 - f_y(y - a_i) \geq 0.$$

Their sum is controlled by (2.3):

$$\begin{aligned} \sum_{i=1}^m (1 - f_y(y - a_i)) &= m - m f_y(y - a_*) \\ &= m(1 - \|y - a_*\|) \leq \frac{\varepsilon}{2}. \end{aligned}$$

Each nonnegative summand is at most $\varepsilon/2$, so $f_y(y - a_i) \geq 1 - \varepsilon/2$ for every i . For $z \in E$, take a_i with $\|z - a_i\| < \varepsilon/2$. Then

$$f_y(y - z) = f_y(y - a_i) + f_y(a_i - z) \geq 1 - \frac{\varepsilon}{2} - \|a_i - z\| \geq 1 - \varepsilon. \quad \square$$

2.3 Reduction without assuming reflexivity

Suppose Theorem 1.1 fails, with a fixed-point-free nonexpansive map $F : C \rightarrow C$ on a nonempty weakly compact convex set in an RNP space X . Starting at $a \in C$, define

$$C_0 = \{a\}, \quad C_{n+1} = \overline{\text{co}}(C_n \cup F(C_n)), \quad C' = \overline{\bigcup_{n \geq 0} C_n}.$$

These are increasing subsets of C . Each C_n is separable: this is true of C_0 , continuous images preserve separability, and finite rational convex combinations of a countable dense set give a countable dense subset of its convex hull. Hence C' is separable, closed, bounded, and convex. Since $F(C_n) \subset C_{n+1}$ and F is norm-continuous, $F(C') \subset C'$. Closed convexity makes C' weakly closed; as $C' \subset C$, it is weakly compact. Only the weak compactness of C , not reflexivity of X , is used.

A decreasing chain of nonempty closed convex invariant subsets of C' has nonempty intersection, by compactness and the finite intersection property. Zorn's lemma gives a minimal such subset K_0 . Its diameter d is positive, since a singleton invariant set is a fixed point. Choose $a_0 \in K_0$ and set

$$Y = \overline{\text{span}}(K_0 - K_0), \quad K = \frac{K_0 - a_0}{d}, \quad T(y) = \frac{F(a_0 + dy) - a_0}{d} \quad (y \in K). \quad (2.4)$$

Now $K \subset Y$ because $a_0 \in K_0$, and for $y, y' \in K$,

$$\|Ty - Ty'\| = \frac{\|F(a_0 + dy) - F(a_0 + dy')\|}{d} \leq \|y - y'\|.$$

Also $Ty = y$ would mean $F(a_0 + dy) = a_0 + dy$, which is impossible. Translation and multiplication by $1/d$ preserve weak compactness, convexity, and minimal invariance. We henceforth use the following setting:

Y is a separable real Banach space with the RNP,

$$K \subset Y \text{ is nonempty, weakly compact, and convex, } \text{diam } K = 1, \quad (2.5)$$

$Y = \overline{\text{span}}(K - K)$, $T : K \rightarrow K$ is nonexpansive and K is minimal invariant.

The RNP passes to the closed subspace Y : a nonempty closed bounded convex subset of Y is also such a subset of X , and a small slice supplied by (1.2) is unchanged when its functional is restricted to Y .

Fix once and for all a norm-dense sequence $(z_j)_{j \geq 1}$ in K . Then

$$\overline{\text{span}}\{z_i - z_j : i, j \geq 1\} = Y. \quad (2.6)$$

All functionals used below belong to Y^* . We do not assume that Y^* is separable or that B_Y is weakly compact. The construction through Section 7 will use the weak compactness of K ; the RNP is used at the end in Section 8.

3. An anchor for convex combinations of resolvent images

Throughout this section K, T, Y satisfy (2.5); in particular $\text{diam } K = 1$. We seek one point $x \in K$ that is approximable by convex combinations of images with arbitrarily small fixed point defect. The same x must work when the required accuracy is increased after earlier choices.

3.1 Resolvents and the two estimates that will be reused

For $n \geq 1$ and $a \in K$, the map

$$u \mapsto \frac{a}{n} + \left(1 - \frac{1}{n}\right) Tu$$

is a contraction of the complete set K into itself. Banach's contraction principle supplies its unique fixed point $R_n(a)$:

$$R_n(a) = \frac{a}{n} + \left(1 - \frac{1}{n}\right) TR_n(a). \quad (3.1)$$

For $n = 1$ this reads $R_1(a) = a$. Comparing (3.1) at a, a' and using nonexpansiveness of T gives

$$\|R_n(a) - R_n(a')\| \leq \frac{1}{n} \|a - a'\| + \left(1 - \frac{1}{n}\right) \|R_n(a) - R_n(a')\|.$$

Subtract the last term and multiply by n to obtain $\|R_n(a) - R_n(a')\| \leq \|a - a'\|$. Thus every R_n is nonexpansive.

For $u = R_n(a)$, (3.1) also gives

$$Tu - u = \frac{Tu - a}{n}, \quad u - a = \left(1 - \frac{1}{n}\right) (Tu - a).$$

Since $Tu, a \in K$, the first identity yields $\|Tu - u\| \leq 1/n$. For the second one,

$$\begin{aligned} \|u - a\| &\leq \left(1 - \frac{1}{n}\right) (\|Tu - Ta\| + \|Ta - a\|) \\ &\leq \left(1 - \frac{1}{n}\right) (\|u - a\| + \|Ta - a\|). \end{aligned}$$

Rearranging gives the two estimates

$$\|TR_n(a) - R_n(a)\| \leq \frac{1}{n}, \quad \|R_n(a) - a\| \leq (n-1)\|Ta - a\| \quad (a \in K). \quad (3.2)$$

The first makes an approximate fixed point; the second says how little R_n moves a point that already has small defect.

Let $\mathcal{G}$ be the family generated by Id_K and all R_n under finite composition and pointwise finite convex combinations. These operations preserve nonexpansiveness and the range K . For every $G \in \mathcal{G}$ there is a finite constant $c_G \geq 0$ such that

$$\|G(a) - a\| \leq c_G \|Ta - a\| \quad (a \in K). \quad (3.3)$$

Here is the verification for the generating operations. For the generators, (3.2) permits $c_{\text{Id}} = 0$, $c_{R_n} = n-1$. For a convex combination $G = \sum_i \lambda_i G_i$,

$$\|G(a) - a\| = \left\| \sum_i \lambda_i (G_i(a) - a) \right\| \leq \sum_i \lambda_i c_{G_i} \|Ta - a\|.$$

For a composition $G \circ H$,

$$\begin{aligned} \|G(H(a)) - a\| &\leq \|G(H(a)) - G(a)\| + \|G(a) - a\| \\ &\leq \|H(a) - a\| + \|G(a) - a\| \leq (c_H + c_G) \|Ta - a\|. \end{aligned}$$

This proves (3.3). Fix one such c_G for each G . Related semigroups of asymptotically invariant nonexpansive maps occur in [2].

For $m \geq 1$ define

$$\mathcal{H}_m = \{R_n \circ G : n \geq m, G \in \mathcal{G}\} \subset \mathcal{G}. \quad (3.4)$$

These nonempty families decrease as m increases. If $h = R_n \circ G \in \mathcal{H}_m$, apply (3.2) at $G(a)$ to obtain

$$\|Th(a) - h(a)\| \leq \frac{1}{n} \leq \frac{1}{m} \quad (h \in \mathcal{H}_m, a \in K). \quad (3.5)$$

Also, for $H \in \mathcal{G}$,

$$(R_n \circ G) \circ H = R_n \circ (G \circ H) \in \mathcal{H}_m, \quad \mathcal{H}_m \circ \mathcal{G} \subset \mathcal{H}_m. \quad (3.6)$$

The defect estimate (3.5) is asserted for members of $\mathcal{H}_m$, not for their convex combinations or weak limits.

3.2 Finite simultaneous approximation

Give K^K the topology of pointwise weak convergence:

$$h_\gamma \xrightarrow{pw} h \iff h_\gamma(a) \rightarrow h(a) \quad (a \in K).$$

As a product of copies of weakly compact K , this is a compact Hausdorff space. Write $\overline{\mathcal{B}}^{pw}$ for closure in this product topology.

Lemma 3.1. *Let $\mathcal{B}$ be a nonempty convex family of nonexpansive selfmaps of K , and let $h \in \overline{\mathcal{B}}^{pw}$. Then h is nonexpansive. Moreover, for $a_1, \dots, a_q \in K$ and $\eta > 0$ there is a single $g \in \mathcal{B}$ such that*

$$\max_{1 \leq i \leq q} \|g(a_i) - h(a_i)\| < \eta.$$

Proof. Take a net $h_\gamma \in \mathcal{B}$ converging pointwise weakly to h . For fixed a, b , the differences converge weakly, and the norm is weakly lower semicontinuous. Hence

$$\|h(a) - h(b)\| \leq \liminf_\gamma \|h_\gamma(a) - h_\gamma(b)\| \leq \|a - b\|.$$

For simultaneous approximation use the convex set

$$V = \{(g(a_1), \dots, g(a_q)) : g \in \mathcal{B}\} \subset Y^q, \quad \|(y_1, \dots, y_q)\|_\infty = \max_i \|y_i\|.$$

Pointwise weak convergence puts $(h(a_1), \dots, h(a_q))$ in the weak closure of V . On a finite product the weak topology is the product of the weak topologies, and the weak and norm closures of a convex set agree. Thus

$$(h(a_1), \dots, h(a_q)) \in \overline{V}^w = \overline{V}^{\|\cdot\|_\infty}.$$

A point of V at distance less than η gives the required one map g . □

3.3 A compact semigroup and the anchor point

Lemma 3.2 (Ellis). *A nonempty compact Hausdorff semigroup in which every right translation $h \mapsto hp$ is continuous contains an idempotent.*

Proof. This is [3, Lemma 1], with the right-translation convention. Compactness and Zorn's lemma give a minimal nonempty compact subsemigroup S_0 . Fix $p \in S_0$. The set S_0p is compact by continuity, and

$$(ap)(bp) = (apb)p \in S_0p \quad (a, b \in S_0).$$

It is therefore a nonempty compact subsemigroup, so $S_0p = S_0$. In particular $hp = p$ for some $h \in S_0$. Consequently $E = \{h \in S_0 : hp = p\}$ is nonempty and compact. If $h, k \in E$, then $(hk)p = h(kp) = hp = p$, so E is a subsemigroup. Minimality gives $E = S_0$; thus $p \in E$ and $p^2 = p$. □

Proposition 3.3 (The anchor point). *There is $x \in K$ such that*

$$D_x = \bigcap_{m \geq 1} \overline{\{h(x) : h \in \mathcal{H}_m\}}^w \tag{3.7}$$

is nonempty and weakly compact, and

$$x \in \overline{\text{co}} D_x. \tag{3.8}$$

Proof. Put

$$A_m = \overline{\text{co}} \mathcal{H}_m^{pw}, \quad A = \bigcap_{m \geq 1} A_m.$$

The A_m are nonempty compact sets and decrease with m , so their intersection A is nonempty. By Lemma 3.1, their members are nonexpansive. We next check closure under composition.

Since $\text{co } \mathcal{H}_m \subset \mathcal{G}$, the right-composition rule (3.6) gives the algebraic inclusions

$$(\text{co } \mathcal{H}_m) \circ (\text{co } \mathcal{H}_m) \subset (\text{co } \mathcal{H}_m) \circ \mathcal{G} \subset \text{co}(\mathcal{H}_m \circ \mathcal{G}) \subset \text{co } \mathcal{H}_m. \quad (3.9)$$

The middle inclusion uses the exact identity

$$\left(\sum_i \lambda_i h_i \right) \circ k = \sum_i \lambda_i (h_i \circ k).$$

Only precomposition is distributed over the outer convex combination.

To pass to A_m , fix $h, k \in A_m$, arguments $a_1, \dots, a_q \in K$, and $\eta > 0$. By Lemma 3.1 choose $k_0, h_0 \in \text{co } \mathcal{H}_m$ such that

$$\max_i \|k_0(a_i) - k(a_i)\| < \eta, \quad \max_i \|h_0(k(a_i)) - h(k(a_i))\| < \eta.$$

The second application uses the fixed arguments $k(a_i)$. Nonexpansiveness of h_0 gives

$$\begin{aligned} \|h_0(k_0(a_i)) - h(k(a_i))\| &\leq \|h_0(k_0(a_i)) - h_0(k(a_i))\| + \|h_0(k(a_i)) - h(k(a_i))\| \\ &\leq \|k_0(a_i) - k(a_i)\| + \|h_0(k(a_i)) - h(k(a_i))\| < 2\eta. \end{aligned}$$

The composition $h_0 \circ k_0$ belongs to $\text{co } \mathcal{H}_m$ by (3.9). Since the arguments and η were arbitrary, these finite simultaneous norm approximations give $h \circ k \in A_m$. Thus A is a compact semigroup.

For a fixed $k \in A$, right translation is continuous. Indeed, if $h_\gamma \rightarrow h$ pointwise weakly, then for $a \in K$ and $f \in Y^*$,

$$f((h_\gamma \circ k)(a)) = f(h_\gamma(k(a))) \longrightarrow f(h(k(a))) = f((h \circ k)(a)).$$

The argument $k(a)$ is fixed; no weak continuity of k is assumed. By Lemma 3.2 there is $P \in A$ with $P \circ P = P$. Choose $a \in K$ and set $x = P(a)$. Then

$$P(x) = P(P(a)) = P(a) = x.$$

This is a fixed point of P , not yet a fixed point of T .

For this x let $B_m = \overline{\{h(x) : h \in \mathcal{H}_m\}}^w$. The sets B_m are nonempty, weakly compact, and decreasing, so $D_x = \bigcap_m B_m$ is nonempty and weakly compact. Evaluation at x is continuous for pointwise weak convergence. Using $P \in A_m$ and equality of weak and norm closures of a convex set gives

$$\begin{aligned} x = P(x) &\in \overline{\{g(x) : g \in \text{co } \mathcal{H}_m\}}^w \\ &= \overline{\text{co}\{h(x) : h \in \mathcal{H}_m\}}^w = \overline{\text{co}\{h(x) : h \in \mathcal{H}_m\}} \subset \overline{\text{co}} B_m. \end{aligned} \quad (3.10)$$

Fix $\ell \in Y^*$. The inclusion $x \in \overline{\text{co}} B_m$ in (3.10) gives $\ell(x) \leq \max_{B_m} \ell$ for every m . Let $t = \inf_m \max_{B_m} \ell$. The nonempty compact sets $B_m \cap \{z : \ell(z) \geq t\}$ decrease, so their intersection contains a point of D_x . Hence

$$\max_{D_x} \ell \geq t = \inf_m \max_{B_m} \ell \geq \ell(x).$$

If $x \notin \overline{\text{co}} D_x$, Hahn–Banach separation would give an ℓ with $\ell(x) > \sup_{D_x} \ell$, a contradiction. This proves (3.8). $\square$

Fix the x of Proposition 3.3 for the rest of the proof. The definition (3.7) gives the following useful selection rule:

$$\begin{aligned} d \in D_x, \quad m \geq 1, \quad U \text{ a weak neighborhood of } d \text{ in } K \\ \implies \exists G \in \mathcal{H}_m \text{ such that } G(x) \in U. \end{aligned} \quad (3.11)$$

The concrete recursive choices using this rule are made in Section 5, after their comparison sets are defined.

We also fix one common sequence. By (3.2) and (3.3),

$$b_j = R_j(x), \quad \|Tb_j - b_j\| \leq \frac{1}{j}, \quad \|G(b_j) - b_j\| \leq c_G \|Tb_j - b_j\| \leq \frac{c_G}{j}. \quad (3.12)$$

The last bound holds for every fixed $G \in \mathcal{G}$. Thus any fixed finite collection of these maps nearly fixes the same sufficiently late b_j .

4. One fixed tree, evaluated from the leaves upwards

There are two directions to distinguish. We construct the nodes, their maps, and their coefficients from the root towards the later levels. Once these are fixed, we compute values in the opposite direction: prescribe a level, then work upwards to the root. The latter finite calculation is what we call a *working tree*. It does not change the underlying tree.

4.1 The parent relation and the two kinds of weight

The root is r , and L_k is the finite, nonempty level of depth k , with $L_0 = \{r\}$. Each node u has a finite, nonempty ordered list of children, which we may write as $u_1, \dots, u_{m(u)}$. Our notation for the parent is

$$v^\uparrow = u \iff v \text{ is a child of } u.$$

Thus $\sum_{v^\uparrow=u}$ is a sum over the children of u ; no separate symbol for the set of children is needed. The lists of children are placed consecutively in the order of their parents. Write $h <_k v$ for the resulting left-to-right order in L_k , and write $u \prec v$ for proper ancestry. We write $u \preceq v$ when $u = v$ is also allowed.

At the edge leading to each nonroot node v , put a coefficient α_v . The *local coefficients* and the *path weights* satisfy

$$\alpha_v \geq 0, \quad \sum_{v^\uparrow=u} \alpha_v = 1; \quad W_r = 1, \quad W_v = W_{v^\uparrow} \alpha_v. \quad (4.1)$$

For a binary tree with equal local coefficients, $W_v = 2^{-k}$ on L_k . We allow any finite number of children, including one, and allow zero coefficients. A zero coefficient means that a node is retained in the list but contributes no weight. It does not require a second record of deleted candidates.

A parent's weight is exactly the sum of its children's weights:

$$\sum_{v^\uparrow=u} W_v = W_u \sum_{v^\uparrow=u} \alpha_v = W_u.$$

Consequently, passing from one level to the next preserves total weight:

$$\sum_{v \in L_{j+1}} W_v = \sum_{u \in L_j} \sum_{v^\uparrow=u} W_v = \sum_{u \in L_j} W_u = 1.$$

Applying the same identity only to the descendants of v gives

$$\sum_{s \in L_N, v \preceq s} W_s = W_v \quad (v \in L_k, N \geq k). \quad (4.2)$$

These descendants form one consecutive block in the ordered level L_N . This is immediate for children; replacing each node of an interval by its consecutive child list preserves the interval property at the next level.

4.2 Prescribed values at one level; determined values above it

Each node carries a fixed nonexpansive map $G_u : K \rightarrow K$, with $G_r = \text{Id}_K$. Fix a depth k . Use the single symbol H_u for the value at node u in this finite calculation:

$$\begin{aligned} u \in L_k : & \quad H_u \in K \text{ is prescribed;} \\ u \in L_j, j < k : & \quad H_u = G_u \left(\sum_{v \uparrow = u} \alpha_v H_v \right) \text{ is then determined.} \end{aligned} \quad (4.3)$$

The rule is applied at levels $k-1, k-2, \dots, 0$, in that order. The cutoff values are not renamed. In particular, G_v is not applied to H_v at the cutoff level L_k . Convexity and $G_u(K) \subset K$ ensure that every computed value also belongs to K .

The notation Φ_k records only the resulting root value:

$$\begin{aligned} \Phi_k((H_v)_{v \in L_k}) &= H_r, \quad \Phi_0(H_r) = H_r, \\ \Phi_k((H_v)_{v \in L_k}) &= \Phi_{k-1} \left(\left(G_u \left(\sum_{v \uparrow = u} \alpha_v H_v \right) \right)_{u \in L_{k-1}} \right) \quad (k \geq 1). \end{aligned} \quad (4.4)$$

Thus Φ_k uses maps strictly above L_k and coefficients down to L_k . It does not use maps at L_k or any data below it. If a working tree starts at a deeper level L_N , its calculated values on L_k can instead be treated as the inputs for the remaining calculation. The rule (4.3) above L_k has not changed. Consequently,

$$\Phi_N((H_s)_{s \in L_N}) = H_r = \Phi_k((H_h)_{h \in L_k}) \quad (0 \leq k \leq N). \quad (4.5)$$

Here the values H_h on L_k are the ones actually calculated from the original depth- N inputs. This identity is the justification for comparing a deeper working tree at an intermediate cutoff. For example, let $L_1 = \{u_1, u_2\}$ and let their respective children be v_{11}, v_{12} and v_{21}, v_{22} . Set

parent	r	u_1	u_2
child coefficients	$(\lambda, 1 - \lambda)$	$(\mu, 1 - \mu)$	$(\nu, 1 - \nu)$

The full calculation from L_2 to the root consists of three equations:

$$\begin{aligned} H_{u_1} &= G_{u_1}(\mu H_{v_{11}} + (1 - \mu) H_{v_{12}}), \\ H_{u_2} &= G_{u_2}(\nu H_{v_{21}} + (1 - \nu) H_{v_{22}}), \\ H_r &= \lambda H_{u_1} + (1 - \lambda) H_{u_2}. \end{aligned}$$

Equivalently,

$$\begin{aligned} \Phi_2((H_v)_{v \in L_2}) &= \lambda G_{u_1}(\mu H_{v_{11}} + (1 - \mu) H_{v_{12}}) \\ &\quad + (1 - \lambda) G_{u_2}(\nu H_{v_{21}} + (1 - \nu) H_{v_{22}}). \end{aligned} \quad (4.6)$$

For nonlinear maps this expression cannot be flattened into a convex combination of the four cutoff values.

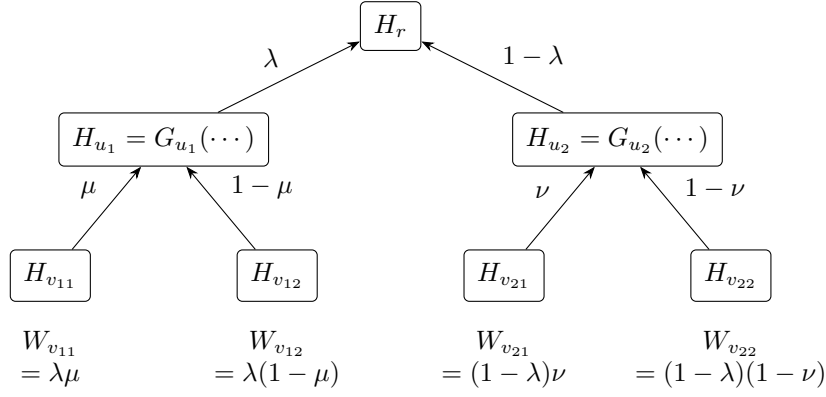


Figure 2: The values at L_2 are prescribed; the arrows show the order of computation. Local coefficients belong to edges. Path weights are their products and will control the effect of a cutoff change on the root.

Lemma 4.1 (Propagation of cutoff errors). *Let H, H' be two assignments on L_k , extended upwards by (4.3). Then*

$$\|\Phi_k((H_v)_{v \in L_k}) - \Phi_k((H'_v)_{v \in L_k})\| \leq \sum_{v \in L_k} W_v \|H_v - H'_v\|. \quad (4.7)$$

Thus changing only the value at v changes the root by at most $W_v \|H_v - H'_v\|$. If every cutoff error is at most η , then the root error is at most η .

Proof. For $k = 0$ the assertion is equality. Assume $k \geq 1$. At a node u one level above the cutoff, the calculation is

$$\begin{aligned} \|H_u - H'_u\| &= \left\| G_u \left(\sum_{v \uparrow = u} \alpha_v H_v \right) - G_u \left(\sum_{v \uparrow = u} \alpha_v H'_v \right) \right\| \\ &\leq \left\| \sum_{v \uparrow = u} \alpha_v (H_v - H'_v) \right\| \leq \sum_{v \uparrow = u} \alpha_v \|H_v - H'_v\|. \end{aligned}$$

When $k \geq 2$, at $p \in L_{k-2}$ the next step is substitution into this same inequality, now at p :

$$\begin{aligned} \|H_p - H'_p\| &\leq \sum_{u \uparrow = p} \alpha_u \|H_u - H'_u\| \\ &\leq \sum_{u \uparrow = p} \sum_{v \uparrow = u} \alpha_u \alpha_v \|H_v - H'_v\|. \end{aligned}$$

This displays the two successive factors on a path. To carry the calculation through all levels without writing an increasingly nested sum, multiply the one-node inequality by W_u and sum over $u \in L_j$:

$$\begin{aligned} \sum_{u \in L_j} W_u \|H_u - H'_u\| &\leq \sum_{u \in L_j} \sum_{v \uparrow = u} W_u \alpha_v \|H_v - H'_v\| \\ &= \sum_{v \in L_{j+1}} W_v \|H_v - H'_v\| \quad (0 \leq j < k). \end{aligned}$$

Here each v occurs under exactly one parent, and $W_u \alpha_v = W_v$. Using $W_r = 1$ from (4.1) and applying the preceding inequality at $j = 0, 1, \dots, k-1$ gives the chain

$$\|H_r - H'_r\| \leq \sum_{u \in L_1} W_u \|H_u - H'_u\| \leq \dots \leq \sum_{v \in L_k} W_v \|H_v - H'_v\|.$$

For uniform cutoff error, (4.2) with $v = r$ gives $\sum_{v \in L_k} W_v = 1$, so the last sum is at most η . $\square$

4.3 Fixed labels and two useful one-node estimates

The maps to be constructed lie in $\mathcal{G}$ from Section 3, and $y_u = G_u(x)$ is a fixed label. In contrast, H_u is a value calculated by (4.3). Suppose for the moment that the child labels satisfy

$$\left\| \sum_{v \uparrow = u} \alpha_v y_v - x \right\| \leq \tau_k \quad (u \in L_k). \quad (4.8)$$

The numbers τ_k and these child choices will be supplied by Proposition 5.1. The average in (4.8) is near x , not near y_u . After applying the parent map we get

$$\left\| G_u \left(\sum_{v \uparrow = u} \alpha_v y_v \right) - y_u \right\| = \left\| G_u \left(\sum_{v \uparrow = u} \alpha_v y_v \right) - G_u(x) \right\| \leq \tau_k.$$

For values calculated from a deeper cutoff, recall both definitions at this node:

$$H_u = G_u \left(\sum_{v \uparrow = u} \alpha_v H_v \right), \quad y_u = G_u(x).$$

Their difference is therefore bounded by

$$\begin{aligned} \|H_u - y_u\| &= \left\| G_u \left(\sum_{v \uparrow = u} \alpha_v H_v \right) - G_u(x) \right\| \\ &\leq \left\| \sum_{v \uparrow = u} \alpha_v H_v - x \right\| \\ &= \left\| \sum_{v \uparrow = u} \alpha_v (H_v - y_v) + \sum_{v \uparrow = u} \alpha_v y_v - x \right\| \\ &\leq \sum_{v \uparrow = u} \alpha_v \|H_v - y_v\| + \tau_k. \end{aligned} \quad (4.9)$$

Nonexpansiveness is used in the second line; the last line uses the triangle inequality and (4.8).

For comparison with one common point $b \in K$, the same node rule gives

$$\begin{aligned} \|H_u - b\| &= \left\| G_u \left(\sum_{v \uparrow = u} \alpha_v H_v \right) - b \right\| \\ &\leq \left\| G_u \left(\sum_{v \uparrow = u} \alpha_v H_v \right) - G_u(b) \right\| + \|G_u(b) - b\| \\ &\leq \left\| \sum_{v \uparrow = u} \alpha_v (H_v - b) \right\| + \|G_u(b) - b\| \\ &\leq \sum_{v \uparrow = u} \alpha_v \|H_v - b\| + \|G_u(b) - b\|. \end{aligned} \quad (4.10)$$

The third line uses $\sum_{v \uparrow = u} \alpha_v = 1$ from (4.1). Thus a parent error is bounded by the convex average of the child errors, plus its own error. The finite inductions in Subsection 6.3 will sum these one-node estimates over the whole subtree, with the leaf values stated explicitly.

5. Constructing a tree before its new coefficients are known

We now construct the maps, labels, coefficients, and functionals. Once chosen, they will not be changed. What depends on the future coefficients is a calculated root output. We must control that output before choosing the current functional. The solution is to collect all possible coefficient choices in one norm-compact set of comparison limits and apply Lemma 2.2 to that set.

5.1 The comparison output and the compactness used here

For a completed level L_k and $v \in L_k$, let $I_v = \{h \in L_k : h <_k v\}$. For $b \in K$ define a particular working tree:

$$H_h = \begin{cases} y_h, & h \in I_v, \\ b, & h \in L_k \setminus I_v, \end{cases} \quad \Psi_v(b) = \Phi_k((H_h)_{h \in L_k}). \quad (5.1)$$

The values are prescribed at level k and then computed upwards by (4.3). Thus v and all nodes to its right have value b , whereas the earlier nodes have their labels. A deeper working tree will approximate this assignment; the approximation is not part of its definition.

This is why even the children of one parent must be selected in order. For three children v_1, v_2, v_3 of the root, switching from right to left changes their values as follows:

$$(y_{v_1}, y_{v_2}, y_{v_3}) \longrightarrow (y_{v_1}, y_{v_2}, b) \longrightarrow (y_{v_1}, b, b) \longrightarrow (b, b, b).$$

For example,

$$\Psi_{v_2}(b) = \alpha_{v_1} y_{v_1} + (1 - \alpha_{v_1})b.$$

The earlier sibling label y_{v_1} is needed to form this comparison; its coefficient may still be unknown. We therefore select labels one at a time, parameterize their coefficients, and choose a finite list of coefficients only after the needed labels and functionals have been obtained.

Two compactness facts will be used in this construction. First, the weak topology on K is metrizable. To see this without assuming Y^* separable, take a norm-dense sequence (u_j) in the unit sphere of the separable space Y , and choose $g_j \in B_{Y^*}$ with $g_j(u_j) = 1$. These functionals separate points of Y : if $\|u\| = 1$ and $\|u - u_j\| < 1/2$, then

$$g_j(u) \geq g_j(u_j) - \|u - u_j\| > 1/2.$$

The map $K \rightarrow \mathbb{R}^{\mathbb{N}}$, $z \mapsto (g_j(z))_j$, is therefore a continuous injection from weakly compact K into a Hausdorff space, hence a homeomorphism onto its image. We may fix the compatible metric

$$\rho(a, b) = \sum_{j \geq 1} 2^{-j} \min\{1, |g_j(a - b)|\} \quad (a, b \in K). \quad (5.2)$$

Only the weak topology on K , not on every bounded subset of Y , is being metrized.

Second, fix a free ultrafilter $\mathcal{U}$ on $\mathbb{N}$. For bounded scalars,

$$\lim_{j \rightarrow \mathcal{U}} t_j = t \iff \{j : |t_j - t| < \varepsilon\} \in \mathcal{U} \quad (\varepsilon > 0).$$

These limits exist and preserve finite linear combinations. The filter contains each cofinite set and is closed under finite intersections; none of its members is empty. For $(w_j) \subset K$, weak compactness gives a unique weak $\mathcal{U}$ -limit $w \in K$, characterized by

$$w = w\text{-}\lim_{j \rightarrow \mathcal{U}} w_j \iff f(w) = \lim_{j \rightarrow \mathcal{U}} f(w_j) \quad (f \in Y^*). \quad (5.3)$$

In bidual terms, $L(f) = \lim_{j \rightarrow \mathcal{U}} f(w_j)$ is linear and $|L(f)| \leq \sup_K \|z\| \|f\|$, so $L \in Y^{**}$. The canonical image $J_Y(K)$ is weak-star compact: J_Y identifies the weak topology of K with the weak-star topology on that image. For any finite list of functionals, the corresponding scalar limit conditions hold on a finite intersection of members of $\mathcal{U}$. Thus L belongs to the weak-star closure of $J_Y(K)$, which is $J_Y(K)$ itself. Consequently $L = J_Y(w)$ for some $w \in K$, proving (5.3). This does not assert surjectivity of $J_Y : Y \rightarrow Y^{**}$; reflexivity is not used. The notation denotes a limit in the ordinary weak topology.

5.2 The construction and all its choices

Recall precisely the inputs from Proposition 3.3, (3.7), and (3.12):

$$\begin{aligned} x \in \overline{\text{co}}D_x, \quad D_x \subset \overline{\{G(x) : G \in \mathcal{H}_m\}}^w \quad (m \geq 1), \\ b_j = R_j(x), \quad \|Tb_j - b_j\| \leq j^{-1}, \quad \|G(b_j) - b_j\| \leq c_G/j \quad (G \in \mathcal{G}). \end{aligned} \quad (5.4)$$

The point x , the set D_x , and the sequence (b_j) are fixed for the entire construction.

Proposition 5.1. *There is one infinite ordered finitely branching tree with $G_r = \text{Id}_K$, $G_v \in \mathcal{G}$, $y_v = G_v(x)$, and the coefficients and weights of (4.1). For every nonroot v there are $e_v > 0$ and $f_v \in B_{Y^*}$ such that*

$$\sum_{v \neq r} e_v \leq \frac{1}{128}. \quad (5.5)$$

Define

$$\mu_k = \min_{v \in L_1 \cup \dots \cup L_k} e_v, \quad \tau_k = 2^{-k-1} \mu_k \quad (k \geq 1), \quad \tau_0 = \frac{1}{2}. \quad (5.6)$$

The following estimates hold:

$$\left\| \sum_{v^\dagger=u} \alpha_v y_v - x \right\| \leq \tau_k \quad (u \in L_k), \quad (5.7)$$

$$\sum_{p \prec v} \|G_p(y_v) - y_v\| \leq e_v \quad (v \neq r), \quad (5.8)$$

$$f_v(y_v - z_i) \geq 1 - e_v \quad (v \in L_k, 1 \leq i \leq k), \quad (5.9)$$

$$\lim_{j \rightarrow \mathcal{U}} f_v(y_v - \Psi_v(b_j)) \geq 1 - e_v \quad (v \neq r). \quad (5.10)$$

Here (z_j) is the norm-dense sequence fixed in (2.6), and Ψ_v is the comparison output defined in (5.1).

Proof. We construct L_k after the levels $L_0, \dots, L_{k-1}$ have been fixed. Within L_k we visit the parents in their left-to-right order. For one parent we first construct candidate labels and functionals, and then take a finite initial list and its coefficients. All members of that finite list are retained, including ones assigned coefficient zero.

1. Fix a parent and its approximation targets.

Let $u \in L_{k-1}$ be the current parent. Its maps and those above it are fixed, as are all coefficients down to level $k-1$. Thus

$$\Phi_{k-1}, \quad (G_p)_{p \in L_{k-1}}, \quad (W_p)_{p \in L_{k-1}}, \quad \tau_{k-1} > 0$$

are known. The children of all parents preceding u have already been chosen, together with their coefficients. Using $x \in \overline{\text{co}}D_x$ in (5.4), choose

$$\begin{aligned} d_1, \dots, d_q &\in D_x, \quad \lambda_t \geq 0, \quad \sum_{t=1}^q \lambda_t = 1, \\ d_* &:= \sum_{t=1}^q \lambda_t d_t, \quad \|d_* - x\| < \frac{\tau_{k-1}}{2}. \end{aligned} \tag{5.11}$$

The point d_* and the coefficients λ_t are now fixed for this parent; they will not be reselected. They are not the eventual child coefficients α_v .

With the metric ρ in (5.2), we shall choose candidate children $v_1, v_2, \dots$ so that

$$t(i) = 1 + ((i-1) \bmod q), \quad \rho(y_{v_i}, d_{t(i)}) < 1/i. \tag{5.12}$$

The next three steps specify every additional requirement at stage i . The finite length M is not specified in advance; it is obtained in Step 5.

2. Form the comparison family using the earlier labels.

Suppose the candidates $v_1, \dots, v_{i-1}$ of u have been chosen and $v = v_i$ is next. The earlier list is

$$I_v = \bigcup_{p <_{k-1} u} \{h : h^\uparrow = p\} \cup \{v_1, \dots, v_{i-1}\}. \tag{5.13}$$

It is finite and every y_h with $h \in I_v$ is known. This is the list of earlier nodes that will occur in (5.1). The current y_v and its sibling coefficients are not known. Later children are placed to the right, so they do not change I_v .

For any eventual completion of the level, the assignment (5.1) sets $H_h = y_h$ on I_v and $H_h = b$ off I_v . Since each parent's coefficients sum to one, its input satisfies

$$\begin{aligned} \sum_{h^\uparrow=p} \alpha_h H_h &= \sum_{\substack{h \in I_v \\ h^\uparrow=p}} \alpha_h y_h + \sum_{\substack{h \notin I_v \\ h^\uparrow=p}} \alpha_h b \\ &= \sum_{\substack{h \in I_v \\ h^\uparrow=p}} \alpha_h y_h + \left(1 - \sum_{\substack{h \in I_v \\ h^\uparrow=p}} \alpha_h\right) b. \end{aligned} \tag{5.14}$$

Thus the individual later labels and coefficients have disappeared. Only coefficients on the fixed finite list I_v remain. We allow these to vary over

$$\Theta_v = \left\{ \theta \in [0, 1]^{I_v} : \sum_{\substack{h \in I_v \\ h^\uparrow=p}} \theta_h \leq 1 \text{ for every } p \in L_{k-1} \right\}. \tag{5.15}$$

This is a compact subset of a finite cube. If I_v is empty, the cube consists of one empty tuple. Allowing even the earlier completed groups' coefficients to vary only enlarges the family and includes the eventual choice.

For $p \in L_{k-1}$ define

$$\begin{aligned} a_{v,p}(b; \theta) &= \sum_{\substack{h \in I_v \\ h^\uparrow = p}} \theta_h y_h + \left(1 - \sum_{\substack{h \in I_v \\ h^\uparrow = p}} \theta_h \right) b \\ &= b + \sum_{\substack{h \in I_v \\ h^\uparrow = p}} \theta_h (y_h - b) \in K. \end{aligned} \quad (5.16)$$

The first line is a convex combination, by (5.15). The root output of this parameterized working tree is

$$\Psi_v(b; \theta) = \Phi_{k-1} \left((G_p(a_{v,p}(b; \theta)))_{p \in L_{k-1}} \right). \quad (5.17)$$

The inputs here are at level $k-1$, after applying the parent maps. The evaluation Φ_{k-1} uses only the fixed maps above that level, as specified in (4.4). Neither G_v nor y_v occurs.

In the depth-two example (4.6), consider the fourth candidate $v = v_{22}$, so $I_v = \{v_{11}, v_{12}, v_{21}\}$. The formulas are

$$\begin{aligned} H_{u_1} &= G_{u_1}(\theta_{v_{11}} y_{v_{11}} + \theta_{v_{12}} y_{v_{12}} + (1 - \theta_{v_{11}} - \theta_{v_{12}}) b), \\ H_{u_2} &= G_{u_2}(\theta_{v_{21}} y_{v_{21}} + (1 - \theta_{v_{21}}) b), \\ \Psi_v(b; \theta) &= \lambda H_{u_1} + (1 - \lambda) H_{u_2}. \end{aligned}$$

The constraints are $\theta_{v_{11}} + \theta_{v_{12}} \leq 1$ and $\theta_{v_{21}} \leq 1$. For deeper levels, the next upward step is

$$H_q = G_q \left(\sum_{p^\uparrow = q} \alpha_p G_p(a_{v,p}(b; \theta)) \right) \quad (q \in L_{k-2}),$$

followed by the remaining fixed rules above L_{k-2} .

3. Obtain a norm-compact set of comparison limits.

Subtract (5.16) for $\theta, \eta \in \Theta_v$:

$$\begin{aligned} a_{v,p}(b; \theta) - a_{v,p}(b; \eta) &= \sum_{\substack{h \in I_v \\ h^\uparrow = p}} (\theta_h - \eta_h) (y_h - b), \\ \|a_{v,p}(b; \theta) - a_{v,p}(b; \eta)\| &\leq \sum_{\substack{h \in I_v \\ h^\uparrow = p}} |\theta_h - \eta_h| \|y_h - b\| \leq \sum_{\substack{h \in I_v \\ h^\uparrow = p}} |\theta_h - \eta_h|. \end{aligned}$$

The last inequality uses $y_h, b \in K$ and $\text{diam } K = 1$. Apply Lemma 4.1, namely (4.7), to the parent outputs in (5.17), and then use nonexpansiveness:

$$\begin{aligned} \|\Psi_v(b; \theta) - \Psi_v(b; \eta)\| &\leq \sum_{p \in L_{k-1}} W_p \|G_p(a_{v,p}(b; \theta)) - G_p(a_{v,p}(b; \eta))\| \\ &\leq \sum_{p \in L_{k-1}} W_p \|a_{v,p}(b; \theta) - a_{v,p}(b; \eta)\| \\ &\leq \sum_{p \in L_{k-1}} \sum_{\substack{h \in I_v \\ h^\uparrow = p}} W_p |\theta_h - \eta_h| \\ &= \sum_{h \in I_v} W_{h^\uparrow} |\theta_h - \eta_h| \leq \sum_{h \in I_v} |\theta_h - \eta_h|. \end{aligned} \quad (5.18)$$

Each $h \in I_v$ has exactly one parent, and $0 \leq W_{h\uparrow} \leq 1$ by (4.1). If no child of p lies in I_v , its sums are empty, $a_{v,p}(b; \theta) = b = a_{v,p}(b; \eta)$, and its contribution to the output difference is $G_p(b) - G_p(b) = 0$.

Recall from (3.12) that the one fixed sequence is

$$b_j = R_j(x), \quad \|Tb_j - b_j\| \leq 1/j.$$

All values $\Psi_v(b_j; \theta)$ lie in weakly compact K , so (5.3) defines

$$Z_v(\theta) = w\text{-}\lim_{j \rightarrow \mathcal{U}} \Psi_v(b_j; \theta) \in K. \quad (5.19)$$

For $f \in B_{Y^*}$, linearity of scalar ultralimits and (5.18) give

$$\begin{aligned} |f(Z_v(\theta) - Z_v(\eta))| &= \left| \lim_{j \rightarrow \mathcal{U}} f(\Psi_v(b_j; \theta) - \Psi_v(b_j; \eta)) \right| \\ &\leq \sum_{h \in I_v} |\theta_h - \eta_h|. \end{aligned}$$

Taking the supremum over $f \in B_{Y^*}$ yields

$$\|Z_v(\theta) - Z_v(\eta)\| \leq \sum_{h \in I_v} |\theta_h - \eta_h|. \quad (5.20)$$

Thus Z_v is norm-continuous on the compact finite-dimensional set Θ_v . The points $z_1, \dots, z_k$ in the next formula are the first k terms of the norm-dense sequence fixed in (2.6); they are not chosen during this induction. We may now set

$$E_v = \{z_1, \dots, z_k\} \cup Z_v(\Theta_v) \subset K. \quad (5.21)$$

It is nonempty and norm-compact, so it satisfies the set hypothesis of Lemma 2.2. It was defined before G_v, y_v, f_v and covers all future coefficient choices. The weak limit was taken after evaluating the entire output in (5.17); no limit was moved through a nonlinear G_p .

4. Choose the map and then the functional.

Let n be the number of nonroot nodes retained before starting the current parent's list. For its i th candidate v assign

$$e_v = 2^{-(n+i)-8}.$$

Apply Lemma 2.2 with ambient space Y , $E = E_v$ from (5.21), and $\varepsilon = e_v$. It gives $\delta_v > 0$ such that

$$\begin{aligned} y \in K, \quad \|Ty - y\| \leq \delta_v &\implies \\ \exists f \in B_{Y^*} \quad f(y - z) \geq 1 - e_v \quad (z \in E_v). \end{aligned} \quad (5.22)$$

All proper ancestor maps are already fixed. Recall their estimates from (3.3):

$$\|G_p(a) - a\| \leq c_{G_p} \|Ta - a\| \quad (a \in K, p \prec v).$$

There are finitely many ancestors, so their constants have a finite sum. Choose $m \geq 1$ with

$$\frac{1}{m} \leq \delta_v, \quad \frac{1}{m} \sum_{p \prec v} c_{G_p} \leq e_v. \quad (5.23)$$

Since $d_{t(i)} \in D_x$, (3.11) applied to its ρ -ball of radius $1/i$ gives

$$G_v \in \mathcal{H}_m, \quad y_v = G_v(x), \quad \rho(y_v, d_{t(i)}) < 1/i. \quad (5.24)$$

Membership in $\mathcal{H}_m$ has a separate role from the weak approximation: by (3.4), $G_v = R_\ell \circ G$ for some $\ell \geq m$ and $G \in \mathcal{G}$. Consequently, by (3.2),

$$\|Ty_v - y_v\| = \|TR_\ell(G(x)) - R_\ell(G(x))\| \leq \frac{1}{\ell} \leq \frac{1}{m} \leq \delta_v. \quad (5.25)$$

Apply (3.3) to each ancestor at $a = y_v$, then use (5.25) and (5.23):

$$\begin{aligned} \sum_{p \prec v} \|G_p(y_v) - y_v\| &\leq \sum_{p \prec v} c_{G_p} \|Ty_v - y_v\| \\ &\leq \frac{1}{m} \sum_{p \prec v} c_{G_p} \leq e_v. \end{aligned} \quad (5.26)$$

This proves the required ancestor bound (5.8) for this candidate. Its own unknown constant c_{G_v} was not used in choosing m .

By (5.25), the hypothesis of (5.22) holds at $y = y_v$. We can now choose

$$f_v \in BY^*, \quad f_v(y_v - z) \geq 1 - e_v \quad (z \in E_v). \quad (5.27)$$

In particular this proves (5.9), since $z_1, \dots, z_k \in E_v$. The finite comparison list and all requirements for the next candidate are now defined. The order is

$$I_v \longrightarrow \Psi_v(\cdot; \theta) \longrightarrow E_v \longrightarrow \delta_v, m \longrightarrow G_v, y_v \longrightarrow f_v.$$

No sibling coefficients have yet been selected.

5. Extract a finite list and choose its coefficients.

Steps 2–4 give a sequence of possible children of u . For each fixed $t \in \{1, \dots, q\}$, the target condition (5.24) gives

$$\rho(y_{v_{qj+t}}, d_t) < \frac{1}{qj+t} \longrightarrow 0 \quad (j \geq 0), \quad y_{v_{qj+t}} \rightharpoonup d_t.$$

Here metric convergence is weak convergence by (5.2). The set $C_u = \overline{\text{co}}\{y_{v_i} : i \geq 1\}$ is weakly closed, being norm closed and convex, and contains each term of those subsequences. Hence

$$d_t \in C_u \quad (1 \leq t \leq q), \quad d_* = \sum_{t=1}^q \lambda_t d_t \in C_u.$$

This is the *same* point d_* fixed in (5.11); recall its distance bound $\|d_* - x\| < \tau_{k-1}/2$. Since $d_* \in C_u$, choose a finite convex combination of the candidates at distance less than $\tau_{k-1}/2$ from d_* . Let M be its largest index and set unused coefficients to zero. Thus

$$\alpha_{v_i} \geq 0, \quad \sum_{i=1}^M \alpha_{v_i} = 1, \quad \left\| \sum_{i=1}^M \alpha_{v_i} y_{v_i} - d_* \right\| < \frac{\tau_{k-1}}{2}. \quad (5.28)$$

The two errors now have their stated sources:

$$\begin{aligned} \left\| \sum_{i=1}^M \alpha_{v_i} y_{v_i} - x \right\| &\leq \left\| \sum_{i=1}^M \alpha_{v_i} y_{v_i} - d_* \right\| + \|d_* - x\| \\ &< \frac{\tau_{k-1}}{2} + \frac{\tau_{k-1}}{2} = \tau_{k-1} \quad \text{by (5.28) and (5.11).} \end{aligned}$$

This is (5.7) for the parent u . Retain all $v_1, \dots, v_M$, including zero-weight ones. For a retained v_i , (5.13) involves only preceding parents and $v_1, \dots, v_{i-1}$, so ignoring the tail $i > M$ changes none of its comparison sets or chosen maps and functionals.

Proceed to the next parent, increasing n by M . There are only finitely many parents in L_{k-1} , and every retained list is finite and nonempty. Thus L_k is finite and nonempty. Define μ_k, τ_k by (5.6); their minimum is over finitely many positive errors. This completes the induction to the next level. The retained nonroot nodes have successive indices $\ell = 1, 2, \dots$, so

$$\sum_{v \neq r} e_v = \sum_{\ell \geq 1} 2^{-\ell-8} = \frac{1}{256} \leq \frac{1}{128}.$$

We have proved (5.5), (5.7), (5.8), and (5.9).

6. Insert the actual coefficients and verify the remaining estimate.

Fix a completed level L_k and $v \in L_k$. Its list is still $I_v = \{h \in L_k : h <_k v\}$. The restriction $\alpha|_{I_v} = (\alpha_h)_{h \in I_v}$ belongs to Θ_v , since

$$\alpha_h \geq 0, \quad \sum_{\substack{h \in I_v \\ h^\dagger = p}} \alpha_h \leq \sum_{h^\dagger = p} \alpha_h = 1 \quad (p \in L_{k-1}).$$

For the ideal values $H_h = y_h$ on I_v and $H_h = b$ elsewhere, (5.14) and (5.16) give at each parent

$$\sum_{h^\dagger = p} \alpha_h H_h = b + \sum_{\substack{h \in I_v \\ h^\dagger = p}} \alpha_h (y_h - b) = a_{v,p}(b; \alpha|_{I_v}).$$

The definition (5.1), the recursion (4.4), and then the family definition (5.17) therefore yield

$$\begin{aligned} \Psi_v(b) &= \Phi_k((H_h)_{h \in L_k}) \\ &= \Phi_{k-1} \left(\left(G_p \left(\sum_{h^\dagger = p} \alpha_h H_h \right) \right)_{p \in L_{k-1}} \right) \\ &= \Phi_{k-1} \left((G_p(a_{v,p}(b; \alpha|_{I_v})))_{p \in L_{k-1}} \right) = \Psi_v(b; \alpha|_{I_v}). \end{aligned} \quad (5.29)$$

This is an equality of finite calculations, not an approximation. Use (5.29) at the fixed sequence $b_j = R_j(x)$, and then the weak-limit definition (5.19):

$$\begin{aligned} \lim_{j \rightarrow \mathcal{U}} f_v(y_v - \Psi_v(b_j)) &= f_v(y_v) - \lim_{j \rightarrow \mathcal{U}} f_v(\Psi_v(b_j; \alpha|_{I_v})) \\ &= f_v(y_v - Z_v(\alpha|_{I_v})) \geq 1 - e_v. \end{aligned}$$

The last inequality is (5.27), because $Z_v(\alpha|_{I_v}) \in E_v$ by (5.21). This proves (5.10) and completes the construction. $\square$

All maps, labels, coefficients, errors, and functionals are now fixed. The parameterized families were used only to establish Proposition 5.1. The finite calculations below use its four estimates (5.7)–(5.10) and the one-argument comparison output (5.1).

6. Finite switching and the differences of all blocks

Fix the single infinite tree of Proposition 5.1. Its maps, labels, coefficients, functionals, and errors remain unchanged from now on. For an integer $N \geq 1$, write

$$\mathcal{T}_N = L_0 \cup \dots \cup L_N, \quad B_N(v) = \{s \in L_N : v \preceq s\} \quad (v \in \mathcal{T}_N). \quad (6.1)$$

The set $B_N(v)$ is the consecutive block of depth- N leaves below v . By (4.2), its total weight is W_v . We retain the change in the root value for *every* block. This will preserve a parent–child addition identity for use in Section 7.

6.1 One common input for each fixed depth

Fix N . Recall the single sequence $b_j = R_j(x)$ in (3.12). For every nonroot $v \in \mathcal{T}_N$, the estimate (5.10) of Proposition 5.1 says

$$\lim_{j \rightarrow \mathcal{U}} f_v(y_v - \Psi_v(b_j)) \geq 1 - e_v.$$

Since $e_v > 0$, the open interval $(1 - 2e_v, \infty)$ contains that limit. The definition of a scalar ultralimit in Subsection 5.1 and closure of a filter under supersets imply

$$A_v := \{j : f_v(y_v - \Psi_v(b_j)) \geq 1 - 2e_v\} \in \mathcal{U}.$$

There are *only finitely many* such conditions: each L_k is finite, N is fixed, and

$$\#(\mathcal{T}_N \setminus \{r\}) = \sum_{k=1}^N \#L_k < \infty.$$

The individual A_v are subsets of $\mathbb{N}$; it is their number, not their cardinality, that is finite here. Also, (3.12) gives the finite sum estimate

$$\sum_{u \in \mathcal{T}_N} \|G_u(b_j) - b_j\| \leq \frac{1}{j} \sum_{u \in \mathcal{T}_N} c_{G_u} \rightarrow 0.$$

The number μ_N from (5.6) is positive. Choose J so that the last sum is at most μ_N whenever $j \geq J$. A free ultrafilter contains $\{j : j \geq J\}$ and is closed under finite intersections, so

$$\emptyset \neq \{j : j \geq J\} \cap \bigcap_{v \in \mathcal{T}_N \setminus \{r\}} A_v \in \mathcal{U}.$$

Choose j_N in this intersection and put $b_N = b_{j_N}$. For the rest of this section abbreviate $b = b_N$. We have proved, for this one b , the simultaneous estimates

$$f_v(y_v - \Psi_v(b)) \geq 1 - 2e_v \quad (v \in \mathcal{T}_N \setminus \{r\}), \quad (6.2)$$

$$\gamma_u := \|G_u(b) - b\|, \quad D_N := \sum_{u \in \mathcal{T}_N} \gamma_u \leq \mu_N. \quad (6.3)$$

The choice b_N may change with N ; for a fixed N it is shared by all blocks and all switching states.

6.2 The switching states and what their root values mean

List $L_N = \{s_1, \dots, s_M\}$ from left to right. For $0 \leq i \leq M$, prescribe the leaf values and calculate the values above them by (4.3):

$$\begin{aligned} H_{s_j}^{[i]} &= \begin{cases} y_{s_j}, & j \leq i, \\ b, & j > i, \end{cases} \\ H_u^{[i]} &= G_u \left(\sum_{w \uparrow = u} \alpha_w H_w^{[i]} \right) \quad (u \in \mathcal{T}_N \setminus L_N), \\ S_i &= \Phi_N((H_s^{[i]})_{s \in L_N}) = H_r^{[i]}. \end{aligned} \quad (6.4)$$

The superscript $[i]$ specifies one assignment, not a level of the tree. The first i leaves have their labels; all later leaves have the common value b . In particular, switching from right to left produces

$$S_M, S_{M-1}, \dots, S_0.$$

Every value in (6.4) lies in K , by convexity and $G_u(K) \subset K$.

If $B_N(v) = \{s_a, \dots, s_c\}$, define

$$U_v := S_c = H_r^{[c]}, \quad V_v := S_{a-1} = H_r^{[a-1]}, \quad \Delta_v^{(N)} := U_v - V_v. \quad (6.5)$$

Thus the subscript v on U_v, V_v names the *block*; these are values at the root r , not values at v . Their assignments can be read directly from (6.4):

	$s_j, j < a$	$s_j, a \leq j \leq c$	$s_j, j > c$	root
$[c]$	y_{s_j}	y_{s_j}	b	$U_v = H_r^{[c]}$
$[a-1]$	y_{s_j}	b	b	$V_v = H_r^{[a-1]}$

Only the block below v changes. At the root, $B_N(r) = L_N$, so $\Delta_r^{(N)} = S_M - S_0$. For every block the consecutive differences telescope:

$$\begin{aligned} \Delta_v^{(N)} &= S_c - S_{a-1} \\ &= (S_c - S_{c-1}) + (S_{c-1} - S_{c-2}) + \dots + (S_a - S_{a-1}) \\ &= \sum_{i=a}^c (S_i - S_{i-1}). \end{aligned} \quad (6.6)$$

These are differences of actual points of K . No nonlinear map is being distributed over a sum.

6.3 The errors in a uniformly assigned subtree

The point $b = b_N$ and the numbers γ_u, D_N remain the ones fixed in (6.3). We need the following two estimates at intermediate levels before comparing the block endpoints.

Lemma 6.1. *Let $h \in L_k$, $0 \leq k \leq N$, and compute values H_u from depth- N leaf values by (4.3). If every leaf $s \in B_N(h)$ is assigned $H_s = y_s$, then*

$$\|H_h - y_h\| \leq \varepsilon_{k,N} := \sum_{d=k}^{N-1} \tau_d. \quad (6.7)$$

For $k \geq 1$ this error satisfies

$$\varepsilon_{k,N} \leq 2^{-k} \mu_k \leq e_v \quad \text{for every } v \in L_k. \quad (6.8)$$

If instead every leaf $s \in B_N(h)$ is assigned $H_s = b$, then

$$\|H_h - b\| \leq \sum_{\substack{u \in T_N \setminus L_N \\ h \preceq u}} \gamma_u \leq D_N \leq \mu_N. \quad (6.9)$$

All these estimates also hold when some path weights are zero.

Proof. 1. All leaf values are labels.

At depth N the prescribed value is $H_s = y_s$, so its error is zero. At $u \in L_{N-1}$, using $y_u = G_u(x)$ and the return estimate (5.7) of Proposition 5.1,

$$\|H_u - y_u\| = \left\| G_u \left(\sum_{s^\dagger=u} \alpha_s y_s \right) - G_u(x) \right\| \leq \left\| \sum_{s^\dagger=u} \alpha_s y_s - x \right\| \leq \tau_{N-1}.$$

If $N \geq 2$, at the next level $p \in L_{N-2}$ the definitions give

$$\begin{aligned}
\|H_p - y_p\| &= \left\| G_p \left(\sum_{u \uparrow = p} \alpha_u H_u \right) - G_p(x) \right\| \\
&\leq \left\| \sum_{u \uparrow = p} \alpha_u H_u - x \right\| \\
&\leq \sum_{u \uparrow = p} \alpha_u \|H_u - y_u\| + \left\| \sum_{u \uparrow = p} \alpha_u y_u - x \right\| \\
&\leq \tau_{N-1} \sum_{u \uparrow = p} \alpha_u + \tau_{N-2} = \tau_{N-1} + \tau_{N-2}.
\end{aligned}$$

The last line uses the first-level bound just proved, (5.7) at p , and $\sum_{u \uparrow = p} \alpha_u = 1$ from (4.1).

For the general induction, assume every child w of $h \in L_k$ has error at most $\sum_{d=k+1}^{N-1} \tau_d$. The one-node calculation (4.9), with (5.7), gives

$$\begin{aligned}
\|H_h - y_h\| &\leq \sum_{w \uparrow = h} \alpha_w \|H_w - y_w\| + \tau_k \\
&\leq \sum_{w \uparrow = h} \alpha_w \left(\sum_{d=k+1}^{N-1} \tau_d \right) + \tau_k = \sum_{d=k}^{N-1} \tau_d.
\end{aligned}$$

The induction starts at the leaves with an empty sum equal to zero, proving (6.7). For $k \geq 1$, the schedule (5.6) gives $\tau_d = 2^{-d-1} \mu_d$ and $\mu_d \leq \mu_k \leq e_v$ for $d \geq k$ and $v \in L_k$. Hence

$$\varepsilon_{k,N} = \sum_{d=k}^{N-1} 2^{-d-1} \mu_d \leq \mu_k \sum_{d=k}^{\infty} 2^{-d-1} = 2^{-k} \mu_k \leq e_v.$$

This proves (6.8). In particular, the budget e_v of one node bounds the untouched error at *any* node h of the same level.

2. All leaf values are the common point.

Now $H_s = b$ at every leaf below h , so $\|H_s - b\| = 0$ there. Recall $\gamma_u = \|G_u(b) - b\|$ from (6.3). The one-node calculation (4.10) reads

$$\begin{aligned}
\|H_u - b\| &\leq \left\| G_u \left(\sum_{w \uparrow = u} \alpha_w H_w \right) - G_u(b) \right\| + \gamma_u \\
&\leq \sum_{w \uparrow = u} \alpha_w \|H_w - b\| + \gamma_u.
\end{aligned} \tag{6.10}$$

At depth $N - 1$ its sum is zero because $H_s = b$ at each child leaf; thus $\|H_u - b\| \leq \gamma_u$. One level higher it gives

$$\|H_p - b\| \leq \gamma_p + \sum_{u \uparrow = p} \alpha_u \gamma_u \leq \gamma_p + \sum_{u \uparrow = p} \gamma_u.$$

For all depths the induction statement is precisely

$$\|H_h - b\| \leq \sum_{\substack{u \in \mathcal{T}_N \setminus L_N \\ h \preceq u}} \gamma_u.$$

At a leaf both sides are zero. If it holds at the children w of an internal node h , substitution into (6.10) yields

$$\begin{aligned} \|H_h - b\| &\leq \gamma_h + \sum_{w \uparrow = h} \alpha_w \|H_w - b\| \\ &\leq \gamma_h + \sum_{w \uparrow = h} \alpha_w \sum_{\substack{u \in \mathcal{T}_N \setminus L_N \\ w \preceq u}} \gamma_u \\ &\leq \gamma_h + \sum_{w \uparrow = h} \sum_{\substack{u \in \mathcal{T}_N \setminus L_N \\ w \preceq u}} \gamma_u = \sum_{\substack{u \in \mathcal{T}_N \setminus L_N \\ h \preceq u}} \gamma_u. \end{aligned}$$

The second inequality is the induction hypothesis. The third uses $0 \leq \alpha_w \leq 1$ and $\gamma_u \geq 0$. The last equality uses the disjoint partition

$$\{u \in \mathcal{T}_N \setminus L_N : h \preceq u\} = \{h\} \sqcup \bigsqcup_{w \uparrow = h} \{u \in \mathcal{T}_N \setminus L_N : w \preceq u\}.$$

Each proper descendant of h belongs to exactly one child subtree. This proves the induction. The sum obtained is over a subset of $\mathcal{T}_N$, and every summand is nonnegative. Using the definition and bound in (6.3),

$$\|H_h - b\| \leq \sum_{\substack{u \in \mathcal{T}_N \setminus L_N \\ h \preceq u}} \gamma_u \leq \sum_{u \in \mathcal{T}_N} \gamma_u = D_N \leq \mu_N.$$

No path weight has been divided out, so zero weights cause no exception. $\square$

6.4 Bounds and exact addition for the block differences

Proposition 6.2. *For every $N \geq 1$, the differences $\Delta_v^{(N)}$ defined by (6.5), using the one common b_N of Subsection 6.1, satisfy*

$$\Delta_v^{(N)} \in K - K, \quad \|\Delta_v^{(N)}\| \leq W_v \quad (v \in \mathcal{T}_N), \quad (6.11)$$

$$f_v(\Delta_v^{(N)}) \geq W_v - 5e_v \quad (v \in \mathcal{T}_N \setminus \{r\}), \quad (6.12)$$

$$\Delta_u^{(N)} = \sum_{v \uparrow = u} \Delta_v^{(N)} \quad (u \in \mathcal{T}_N \setminus L_N). \quad (6.13)$$

Proof. Fix N and $b = b_N$ as in Subsection 6.1. For $v \in L_k$, write $B_N(v) = \{s_a, \dots, s_c\}$. The root values before and after the block are, by (6.4) and (6.5),

$$U_v = S_c = H_r^{[c]}, \quad V_v = S_{a-1} = H_r^{[a-1]}.$$

We first compare each of these root values with a point used by f_v .

1. The root value before the block.

Assume $v \neq r$, so $k \geq 1$. All leaves below v have their labels in assignment $[c]$. Thus Lemma 6.1, specifically (6.7) and (6.8), gives

$$\|H_v^{[c]} - y_v\| \leq \varepsilon_{k,N} \leq e_v. \quad (6.14)$$

Let $p_0 = r, p_1, \dots, p_k = v$ be the path to v . At p_i , separate the on-path child p_{i+1} from the other children in the node rule (6.4):

$$\begin{aligned} \|H_{p_i}^{[c]} - y_v\| &\leq \left\| G_{p_i} \left(\sum_{h^\dagger = p_i} \alpha_h H_h^{[c]} \right) - G_{p_i}(y_v) \right\| + \|G_{p_i}(y_v) - y_v\| \\ &\leq \left\| \alpha_{p_{i+1}} (H_{p_{i+1}}^{[c]} - y_v) + \sum_{\substack{h^\dagger = p_i \\ h \neq p_{i+1}}} \alpha_h (H_h^{[c]} - y_v) \right\| + \|G_{p_i}(y_v) - y_v\| \\ &\leq \alpha_{p_{i+1}} \|H_{p_{i+1}}^{[c]} - y_v\| + (1 - \alpha_{p_{i+1}}) + \|G_{p_i}(y_v) - y_v\|. \end{aligned}$$

The second line uses nonexpansiveness and the coefficient sum one in (4.1). For the third line, each off-path value $H_h^{[c]}$ and y_v lies in the diameter-one set K , so $\|H_h^{[c]} - y_v\| \leq 1$.

Multiply by W_{p_i} and use $W_{p_i} \alpha_{p_{i+1}} = W_{p_{i+1}}$ from (4.1):

$$W_{p_i} \|H_{p_i}^{[c]} - y_v\| - W_{p_{i+1}} \|H_{p_{i+1}}^{[c]} - y_v\| \leq W_{p_i} - W_{p_{i+1}} + W_{p_i} \|G_{p_i}(y_v) - y_v\|.$$

Sum over $i = 0, \dots, k-1$. Both differences telescope. Since $W_r = 1$ and $H_r^{[c]} = U_v$ by (6.5), we obtain

$$\|U_v - y_v\| \leq W_v \|H_v^{[c]} - y_v\| + 1 - W_v + \sum_{i=0}^{k-1} W_{p_i} \|G_{p_i}(y_v) - y_v\|. \quad (6.15)$$

Here are the two previously established bounds being substituted:

$$\begin{aligned} W_v \|H_v^{[c]} - y_v\| &\leq W_v e_v \leq e_v && \text{by (6.14) and } W_v \leq 1, \\ \sum_{i=0}^{k-1} W_{p_i} \|G_{p_i}(y_v) - y_v\| &\leq \sum_{p \prec v} \|G_p(y_v) - y_v\| \leq e_v && \text{by (5.8).} \end{aligned}$$

Substitution into (6.15) gives

$$\|U_v - y_v\| \leq e_v + 1 - W_v + e_v = 1 - W_v + 2e_v. \quad (6.16)$$

Thus $1 - W_v$ comes from the off-path weights; the two e_v terms have different sources, displayed just above.

2. The root value after the block.

Continue to assume $v \neq r$. The assignment after this block is $[a-1]$, not a newly chosen working tree. Its root is $H_r^{[a-1]} = S_{a-1} = V_v$ by (6.4) and (6.5). At depth k , the list $I_v = \{h \in L_k : h <_k v\}$ identifies the untouched subtrees; all other depth- k subtrees have switched completely. For comparison with the ideal input in (5.1),

$h \in L_k$	input for $\Psi_v(b)$	computed error in $[a-1]$
$h \in I_v$	y_h	$\ H_h^{[a-1]} - y_h\ \leq \varepsilon_{k,N} \leq e_v$
$h \notin I_v$	b	$\ H_h^{[a-1]} - b\ \leq \mu_N \leq \mu_k \leq e_v.$

The first bound is (6.7)–(6.8) of Lemma 6.1. The second is (6.9), followed by $\mu_N \leq \mu_k \leq e_v$ from (5.6). By the cutoff consistency identity (4.5), the actual root value after the block can be evaluated from the calculated values on L_k :

$$V_v = \Phi_N((H_s^{[a-1]})_{s \in L_N}) = \Phi_k((H_h^{[a-1]})_{h \in L_k}).$$

The ideal comparison output in (5.1) is

$$\Psi_v(b) = \Phi_k((H_h^{\text{ideal}})_{h \in L_k}), \quad H_h^{\text{ideal}} = \begin{cases} y_h, & h \in I_v, \\ b, & h \in L_k \setminus I_v. \end{cases}$$

Both root values are therefore obtained by applying the same map Φ_k to two level- k assignments. Apply the propagation estimate (4.7) of Lemma 4.1, then split the sum according to I_v :

$$\begin{aligned} \|V_v - \Psi_v(b)\| &= \left\| \Phi_k((H_h^{[a-1]})_{h \in L_k}) - \Phi_k((H_h^{\text{ideal}})_{h \in L_k}) \right\| \\ &\leq \sum_{h \in L_k} W_h \|H_h^{[a-1]} - H_h^{\text{ideal}}\| \\ &= \sum_{h \in I_v} W_h \|H_h^{[a-1]} - y_h\| + \sum_{h \in L_k \setminus I_v} W_h \|H_h^{[a-1]} - b\| \\ &\leq \varepsilon_{k,N} \sum_{h \in I_v} W_h + \mu_N \sum_{h \in L_k \setminus I_v} W_h \\ &\leq e_v \left(\sum_{h \in I_v} W_h + \sum_{h \in L_k \setminus I_v} W_h \right) \\ &= e_v \sum_{h \in L_k} W_h = e_v. \end{aligned} \tag{6.17}$$

The two error bounds are (6.7)–(6.8) and (6.9); the penultimate inequality also uses $\mu_N \leq \mu_k \leq e_v$ from (5.6). The last equality is total mass one, (4.2) with the root as ancestor.

3. The functional detects the block difference.

The definition (6.5) and the identity

$$U_v - V_v = (y_v - \Psi_v(b)) + (U_v - y_v) - (V_v - \Psi_v(b))$$

allow us to use all three estimates at once. Since $\|f_v\| \leq 1$ in Proposition 5.1,

$$\begin{aligned} f_v(\Delta_v^{(N)}) &\geq f_v(y_v - \Psi_v(b)) - \|U_v - y_v\| - \|V_v - \Psi_v(b)\| \\ &\geq (1 - 2e_v) - (1 - W_v + 2e_v) - e_v = W_v - 5e_v. \end{aligned}$$

The three substituted bounds are respectively (6.2), (6.16), and (6.17). This proves (6.12).

4. The norm bound and the parent–child identity.

Both U_v, V_v belong to K , so $\Delta_v^{(N)} \in K - K$. The two leaf assignments in (6.5) agree outside $B_N(v)$ and change each leaf in that block from y_s to b . Recall from (6.5) that $\Delta_v^{(N)} = U_v - V_v$. By Lemma 4.1 and (4.2),

$$\|\Delta_v^{(N)}\| = \|U_v - V_v\| \leq \sum_{s \in B_N(v)} W_s \|y_s - b\| \leq \sum_{s \in B_N(v)} W_s = W_v.$$

Here $y_s, b \in K$ gives $\|y_s - b\| \leq 1$. This proves (6.11), including $v = r$ and $W_v = 0$.

For u strictly above level N , let its children be $v_1, \dots, v_m$ in order. Write $B_N(v_j) = \{s_{a_j}, \dots, s_{c_j}\}$. These blocks partition $B_N(u) = \{s_a, \dots, s_c\}$ consecutively:

$$a_1 = a, \quad a_{j+1} = c_j + 1 \ (j < m), \quad c_m = c.$$

Using the same output sequence S_i for all blocks,

$$\begin{aligned} \sum_{v \uparrow = u} \Delta_v^{(N)} &= \sum_{j=1}^m (S_{c_j} - S_{a_{j-1}}) \\ &= (S_{c_1} - S_{a_{-1}}) + (S_{c_2} - S_{c_1}) + \cdots + (S_c - S_{c_{m-1}}) \\ &= S_c - S_{a_{-1}} = \Delta_u^{(N)}. \end{aligned}$$

This proves (6.13). It is precisely why all blocks at a fixed depth use one b_N and one switching experiment. $\square$

7. One exact convex tree from the finite block differences

For every depth N , Proposition 6.2 supplies the differences $\Delta_v^{(N)} = U_v - V_v$ of actual root values, with

$$\Delta_v^{(N)} \in K - K, \quad \|\Delta_v^{(N)}\| \leq W_v, \quad \Delta_u^{(N)} = \sum_{v \uparrow = u} \Delta_v^{(N)}.$$

The first two assertions are (6.11); the last is (6.13) for nodes strictly above level N . The tree and the weights are the same for all N . We now take limits while keeping each node fixed. This is the additional compactness step that replaces the use of reflexivity in the source proof.

7.1 A weakly compact set for each fixed node

For $W_v > 0$ and N at least the depth of v , define

$$d_v^{(N)} = \frac{\Delta_v^{(N)}}{W_v}, \quad C_v = B_Y \cap \frac{K - K}{W_v}. \quad (7.1)$$

By (6.11), $d_v^{(N)} \in C_v$. Here C_v is weakly compact: $K - K$ is the continuous image of weakly compact $K \times K$ under $(a, b) \mapsto a - b$; multiplication by the fixed scalar $1/W_v$ preserves weak compactness; and the unit ball B_Y is weakly closed. No weak compactness of B_Y itself is required.

For N below the depth of v , set $d_v^{(N)} = 0$. This remains in C_v because $0 \in K - K$. If $W_v = 0$, set $d_v^{(N)} = 0$ for every N . In that case (6.11) also gives $\Delta_v^{(N)} = 0$ whenever it is defined. Thus for every node of depth at most N we have

$$W_v d_v^{(N)} = \Delta_v^{(N)}, \quad \|d_v^{(N)}\| \leq 1. \quad (7.2)$$

Although the sets C_v depend on v and $1/W_v$ may be large, each v has its own fixed weakly compact set for the entire sequence in N .

7.2 One subsequence for all nodes

The union of the finite levels is countable. Enumerate the nodes of positive weight as $v(1), v(2), \dots$. By Eberlein–Šmulian applied to $C_{v(1)}$, choose an infinite index set $\Lambda_1 \subset \mathbb{N}$ along which $d_{v(1)}^{(N)}$ converges weakly. Within it choose an infinite $\Lambda_2 \subset \Lambda_1$ giving weak convergence at $v(2)$, and continue. At each stage the earlier convergences are preserved because a subsequence of a convergent sequence has the same limit.

Choose successively $N_j \in \Lambda_j$ with $N_j > N_{j-1}$ and $N_j \geq j$. For fixed m , all N_j with $j \geq m$ belong to Λ_m . Thus this one diagonal subsequence gives

$$d_v^{(N_j)} \rightharpoonup d_v \in C_v \quad \text{for every fixed node } v \text{ with } W_v > 0. \quad (7.3)$$

At zero-weight nodes put $d_v = 0$. Consequently $\|d_v\| \leq 1$ for all nodes. All limits below use the same subsequence (7.3).

Proposition 7.1. *The limit vectors satisfy*

$$d_u = \sum_{v^\uparrow=u} \alpha_v d_v, \quad \|d_u\| \leq 1 \quad \text{for every node } u, \quad (7.4)$$

$$0 \leq W_v(1 - f_v(d_v)) \leq 5e_v \quad (v \neq r). \quad (7.5)$$

For a positive-weight node $u \in L_k$ and any $n > k$,

$$\sum_{\substack{v \in L_n \\ u \preceq v}} \frac{W_v}{W_u} = 1, \quad d_u = \sum_{\substack{v \in L_n \\ u \preceq v}} \frac{W_v}{W_u} d_v. \quad (7.6)$$

Proof. 1. Pass the exact parent–child identity to the limit.

Fix u with $W_u > 0$. For N greater than its depth, (6.13) and (7.2) give

$$W_u d_u^{(N)} = \sum_{v^\uparrow=u} W_v d_v^{(N)}.$$

Since $W_v = W_u \alpha_v$ by (4.1), division by the fixed positive W_u gives

$$d_u^{(N)} = \sum_{v^\uparrow=u} \alpha_v d_v^{(N)}.$$

The child list $\{v : v^\uparrow = u\}$ is finite, all coefficients α_v are fixed, and (7.3) uses one common subsequence for the parent and all its children. Finite linear combinations are weakly continuous, so we may take the weak limit directly:

$$\begin{aligned} d_u &= w\text{-}\lim_{j \rightarrow \infty} d_u^{(N_j)} = w\text{-}\lim_{j \rightarrow \infty} \sum_{v^\uparrow=u} \alpha_v d_v^{(N_j)} \\ &= \sum_{v^\uparrow=u} \alpha_v \left(w\text{-}\lim_{j \rightarrow \infty} d_v^{(N_j)} \right) = \sum_{v^\uparrow=u} \alpha_v d_v. \end{aligned}$$

This proves (7.4) at u . No weak limit is being moved through a nonlinear map here; only a fixed finite linear combination is involved. If $W_u = 0$, every child also has weight zero and all the corresponding d_u, d_v were defined as zero. The same equality then holds without division.

2. Pass the detection estimate to the limit.

For a nonroot v of positive weight, (6.12) and (7.2) yield, for all sufficiently large N ,

$$0 \leq W_v(1 - f_v(d_v^{(N)})) = W_v - f_v(\Delta_v^{(N)}) \leq 5e_v.$$

The first inequality uses $\|f_v\| \leq 1$ and $\|d_v^{(N)}\| \leq 1$. The functional f_v is fixed, so (7.3) lets us pass to its scalar limit, proving (7.5). At weight zero the left expression is zero, so that estimate holds as well.

3. Iterate the convex decomposition to deeper levels.

For (7.6), the mass assertion is (4.2) divided by $W_u > 0$. The vector assertion can be checked one level at a time without changing any weights. Multiplying (7.4) by W_u first gives

$$W_u d_u = \sum_{v^\uparrow=u} W_v d_v.$$

At two levels it gives

$$W_u d_u = \sum_{v \uparrow = u} \sum_{w \uparrow = v} W_w d_w = \sum_{\substack{w \in L_{k+2} \\ u \preceq w}} W_w d_w.$$

More generally, if the expression at level n equals $W_u d_u$, then

$$\begin{aligned} \sum_{\substack{v \in L_n \\ u \preceq v}} W_v d_v &= \sum_{\substack{v \in L_n \\ u \preceq v}} \sum_{w \uparrow = v} W_w d_w \\ &= \sum_{\substack{w \in L_{n+1} \\ u \preceq w}} W_w d_w. \end{aligned}$$

The child lists partition the descendants at the next level. Induction and division by W_u prove (7.6). $\square$

The original labels $y_v = G_v(x)$ did not form an exact convex tree: their return estimate (5.7) was only approximate and was centered at x . The new vectors d_v do form one, by (7.4). Exactness comes from the addition of block differences (6.13), followed by fixed-node weak limits.

8. A small slice gives the contradiction

We now combine the fixed functionals from Proposition 5.1 with the vectors from Proposition 7.1. The functionals become small on each fixed vector, but their average deficit on the vectors at their own level also tends to zero. The exact convex decomposition will make these two properties incompatible with a small slice.

8.1 Vanishing on fixed vectors and the weighted deficit

Lemma 8.1. *For every fixed $d \in Y$,*

$$\max_{v \in L_k} |f_v(d)| \longrightarrow 0 \quad (k \rightarrow \infty). \quad (8.1)$$

For the limit vectors d_v in (7.3) and each $k \geq 1$, the nonnegative numbers

$$q_k := \sum_{v \in L_k} W_v (1 - f_v(d_v)) \leq 5 \sum_{v \in L_k} e_v \quad (8.2)$$

satisfy $q_k \rightarrow 0$.

Proof. Recall the norm-dense sequence (z_j) in K fixed in (2.6). For fixed i, j and $k \geq \max\{i, j\}$, (5.9) of Proposition 5.1 gives, for every $v \in L_k$,

$$1 - e_v \leq f_v(y_v - z_i) \leq 1, \quad 1 - e_v \leq f_v(y_v - z_j) \leq 1.$$

The upper bounds use $\|f_v\| \leq 1$, $y_v, z_i, z_j \in K$, and $\text{diam } K = 1$. Subtract the two scalar expressions:

$$f_v(z_i - z_j) = f_v(y_v - z_j) - f_v(y_v - z_i), \quad |f_v(z_i - z_j)| \leq e_v.$$

The budget (5.5) is summable across the levels, so

$$\sum_{k \geq 1} \sum_{v \in L_k} e_v < \infty, \quad \max_{v \in L_k} e_v \leq \sum_{v \in L_k} e_v \longrightarrow 0. \quad (8.3)$$

If $a = \sum_{t=1}^m c_t(z_{it} - z_{jt})$, choose k beyond all the indices occurring in this finite sum. Then

$$\max_{v \in L_k} |f_v(a)| \leq \left(\sum_{t=1}^m |c_t| \right) \max_{v \in L_k} e_v \rightarrow 0.$$

Such vectors a are dense in Y by (2.6). For arbitrary fixed $d \in Y$ and $\eta > 0$, choose one with $\|d - a\| < \eta$. Since $\|f_v\| \leq 1$,

$$\max_{v \in L_k} |f_v(d)| \leq \|d - a\| + \max_{v \in L_k} |f_v(a)|.$$

The limsup is at most η ; let $\eta \downarrow 0$ to prove (8.1). Finally, summing (7.5) over the finite level L_k gives (8.2), and (8.3) gives $q_k \rightarrow 0$. $\square$

8.2 The closed convex hull cannot have a small slice

Proof of Theorem 1.1. Assume a counterexample. The reduction (2.5) gives the RNP space Y . Proposition 5.1 supplies the maps and functionals; Proposition 6.2 supplies the finite block differences; Proposition 7.1 gives the limit vectors and their exact convex decomposition. Define

$$D = \overline{\text{co}}\{d_v : W_v > 0\} \subset B_Y. \quad (8.4)$$

This set is nonempty because $W_r = 1$. It is bounded because $\|d_v\| \leq 1$, and is closed and convex by definition. Zero-weight nodes are excluded from the generating set; they never contribute to a convex decomposition in (7.6).

By the RNP slice property (1.2), choose $g \in Y^*$ and $\varepsilon > 0$ such that the slice

$$\sigma = \sup_{d \in D} g(d), \quad a = \sigma - \varepsilon, \quad S = \{d \in D : g(d) > a\}, \quad \delta := \text{diam } S < 1 \quad (8.5)$$

is nonempty. We will show that a positive fraction of the descendants of one node must stay in S . On this fraction their functional deficit cannot tend to zero, contradicting (8.2).

1. Choose a node in the slice and fix it.

Since D is the closed convex hull in (8.4),

$$\sup_{d \in D} g(d) = \sup_{W_v > 0} g(d_v).$$

Indeed, a linear functional on any finite convex combination is at most the supremum of its values on the generating points, and continuity preserves the inequality under norm closure. As $a < \sigma$, there is a node u with

$$W_u > 0, \quad d_u \in S, \quad \kappa := \frac{g(d_u) - a}{\varepsilon} > 0. \quad (8.6)$$

Fix this u , as well as $g, \varepsilon, a, \delta, \kappa$, for all that follows. In particular W_u is one fixed positive number, even if it is small.

2. A fixed positive mass of descendants remains in the slice.

Let $u \in L_k$ and $n > k$. The exact descendant identities (7.6) say

$$\sum_{\substack{v \in L_n \\ u \preceq v}} \frac{W_v}{W_u} = 1, \quad d_u = \sum_{\substack{v \in L_n \\ u \preceq v}} \frac{W_v}{W_u} d_v.$$

For the descendants lying in S , put

$$p_n = \sum_{\substack{v \in L_n \\ u \preceq v, d_v \in S}} \frac{W_v}{W_u}. \quad (8.7)$$

Terms of weight zero contribute nothing. Every positive-weight d_v lies in D . By (8.5), its value under g is at most a outside S and at most $\sigma = a + \varepsilon$ inside S . Thus

$$\begin{aligned} g(d_u) &= \sum_{\substack{v \in L_n \\ u \preceq v}} \frac{W_v}{W_u} g(d_v) \\ &= \sum_{\substack{v \in L_n \\ u \preceq v, d_v \in S}} \frac{W_v}{W_u} g(d_v) + \sum_{\substack{v \in L_n \\ u \preceq v, d_v \notin S}} \frac{W_v}{W_u} g(d_v) \\ &\leq (a + \varepsilon)p_n + a(1 - p_n) = a + \varepsilon p_n. \end{aligned}$$

Rearrange and use the definition of κ in (8.6):

$$p_n \geq \frac{g(d_u) - a}{\varepsilon} = \kappa > 0 \quad (n > k). \quad (8.8)$$

The lower bound is independent of n .

3. A small slice forces a positive deficit on those descendants.

Apply (8.1) of Lemma 8.1 to the fixed vector d_u . Since $\delta < 1$, there is n_0 such that

$$\max_{v \in L_n} |f_v(d_u)| \leq \frac{1 - \delta}{2} \quad (n \geq n_0). \quad (8.9)$$

If $d_v \in S$ and $n \geq n_0$, both d_v and d_u belong to the set S of diameter δ , so $\|d_v - d_u\| \leq \delta$. Together with (8.9) and $\|f_v\| \leq 1$, this gives

$$\begin{aligned} f_v(d_v) &= f_v(d_u) + f_v(d_v - d_u) \\ &\leq |f_v(d_u)| + \|d_v - d_u\| \leq \frac{1 - \delta}{2} + \delta = \frac{1 + \delta}{2}. \end{aligned}$$

Consequently,

$$1 - f_v(d_v) \geq \frac{1 - \delta}{2} > 0 \quad (v \in L_n, d_v \in S, n \geq n_0). \quad (8.10)$$

4. Compare with the deficit of the whole level.

Recall $q_n = \sum_{v \in L_n} W_v(1 - f_v(d_v)) \rightarrow 0$ from (8.2). All summands are nonnegative, since $\|f_v\|, \|d_v\| \leq 1$. For $n > k$ and $n \geq n_0$, retain only the terms belonging to descendants of u inside S :

$$\begin{aligned} \frac{q_n}{W_u} &\geq \sum_{\substack{v \in L_n \\ u \preceq v, d_v \in S}} \frac{W_v}{W_u} (1 - f_v(d_v)) \\ &\geq \frac{1 - \delta}{2} \sum_{\substack{v \in L_n \\ u \preceq v, d_v \in S}} \frac{W_v}{W_u} = \frac{1 - \delta}{2} p_n \geq \frac{1 - \delta}{2} \kappa > 0. \end{aligned}$$

The second inequality is (8.10); the last is (8.8). But $q_n/W_u \rightarrow 0$, because $q_n \rightarrow 0$ and $W_u > 0$ was fixed in (8.6). This contradiction proves Theorem 1.1. $\square$

The same argument shows that every nonempty slice of the particular set (8.4) has diameter at least one. The RNP is used only to exclude such a set. No martingale convergence theorem or weak compactness of the entire ball B_Y is needed. The role of weak compactness is localized in the sets C_v of (7.1).

8.3 The commuting-family consequence

Corollary 8.2. *Let X be a real Banach space with the RNP, let $C \subset X$ be nonempty, weakly compact, and convex, and let $\mathcal{S}$ be a pairwise commuting family of nonexpansive selfmaps of C . Then*

$$\text{Fix}(\mathcal{S}) = \{x \in C : T(x) = x \text{ for every } T \in \mathcal{S}\}$$

is nonempty and is a nonexpansive retract of C .

Proof. If $\mathcal{S}$ is empty, the identity is the required retraction. For the nonempty case we apply [1, Theorem 1]. Its hypotheses are the fixed point property of C , the conditional fixed point property of C , and weak compactness of C . The first follows from Theorem 1.1. To check the second, let $T : C \rightarrow C$ be nonexpansive and $E \subset C$ a nonempty norm-closed bounded convex T -invariant subset. Being norm closed and convex, E is weakly closed; as a weakly closed subset of weakly compact C , it is weakly compact. Thus Theorem 1.1 applied to $T|_E$ gives a fixed point in E . This proves the conditional property, in fact without needing to assume that T already has a fixed point in C . Bruck's theorem gives the nonempty common fixed set and a nonexpansive retraction onto it. $\square$

Attribution

This note reconstructs the argument in the cited OpenAI preprint and records an RNP extension; it is not an announcement of an independent solution. The minimal invariant-set reduction, anchor, ordered-tree construction, finite-switching mechanism, and weighted block detection are attributed to [10]. The fixed-node compactness formulation, simultaneous diagonal limit, exact bounded convex tree, and small-slice conclusion are the additional organization recorded in this note.

Use of ChatGPT

The entire manuscript was drafted and revised using ChatGPT.

References

- [1] R. E. Bruck, Jr., *A common fixed point theorem for a commuting family of nonexpansive mappings*, Pacific J. Math. **53** (1974), 59–71. doi:10.2140/pjm.1974.53.59.
- [2] R. E. Bruck, Jr., *A common fixed-point theorem for compact convex semigroups of nonexpansive mappings*, Proc. Amer. Math. Soc. **53** (1975), 113–116. doi:10.1090/S0002-9939-1975-0383164-3.
- [3] R. Ellis, *Distal transformation groups*, Pacific J. Math. **8** (1958), 401–405. doi:10.2140/pjm.1958.8.401.
- [4] K. Goebel, *On the structure of minimal invariant sets for nonexpansive mappings*, Ann. Univ. Mariae Curie-Skłodowska Sect. A **29** (1975), 73–77 (published 1977).
- [5] L. A. Karlovitz, *Existence of fixed points of nonexpansive mappings in a space without normal structure*, Pacific J. Math. **66** (1976), 153–159. doi:10.2140/pjm.1976.66.153.
- [6] W. A. Kirk, *A fixed point theorem for mappings which do not increase distances*, Amer. Math. Monthly **72** (1965), no. 9, 1004–1006. doi:10.2307/2313345.
- [7] A. T.-M. Lau, P. F. Mah, and A. Ülger, *Fixed point property and normal structure for Banach spaces associated to locally compact groups*, Proc. Amer. Math. Soc. **125** (1997), no. 7, 2021–2027. doi:10.1090/S0002-9939-97-03773-8.
- [8] R. E. Megginson, *An Introduction to Banach Space Theory*, Graduate Texts in Mathematics, vol. 183, Springer-Verlag, New York, 1998. doi:10.1007/978-1-4612-0603-3.
- [9] G. Pisier, *Martingales in Banach Spaces (in connection with Type and Cotype)*, IHP course notes, 2–8 February 2011, version dated 9 February 2011. Remark 2.2, Theorem 2.5, and Corollary 2.11. Author’s course notes.
- [10] OpenAI, *Fixed Points of Nonexpansive Maps in Reflexive Banach Spaces*, preprint dated 24 September 2026, OpenAI Math Release, family 328. The source used for this reconstruction is pinned to repository commit 1581d4cbdd69a0faf34b2089614fe2a814316a6a. Official preprint at the pinned commit.
- [11] A. Wiśnicki, *Around the nonlinear Ryll–Nardzewski theorem*, Math. Ann. **377** (2020), 267–279. doi:10.1007/s00208-020-01985-8.