> **This is a recreation of the original chat by Codex using the data export supplied by the user.**

Prove Delta Density Conjecture

Conversation ID: 6aa8b59d-d12c-83ea-8930-950e1d7a818f

54 recorded user and assistant messages • Dates shown in Asia/Shanghai (UTC+08:00), with source timestamps.

Source: ChatGPT data export

Source: supplied ChatGPT data export dated 2026-10-25. Conversation URL: https://chatgpt.com/c/6aa8b59d-d12c-83ea-8930-950e1d7a818f
This PDF reproduces all 54 saved user questions and assistant text responses: 49 on the saved active conversation path and 5 saved alternate message versions in a separate appendix.
Messages retain their original wording. Mathematics and Markdown are typeset; citation chips are reconstructed from exported reference metadata.
Message dates use the original creation timestamps, displayed in Asia/Shanghai (UTC+08:00); UTC source timestamps are also shown.
The exact source message text is embedded in this PDF as a UTF-8 transcript and JSON record. Original uploaded files and available cited file assets are also embedded; use a PDF reader with an attachments panel.
This is a record of the conversation, including any errors and later corrections in the original answers.
Message date range: 2026-09-15 11:03:57 (Asia/Shanghai, UTC+08:00) to 2026-09-17 11:06:32 (Asia/Shanghai, UTC+08:00)

Main conversation

1. User

2026-09-15 11:03:57 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T03:03:57.239000Z

in another chat I defined delta-density largeness in relation to the 2-large is large conjecture and my work in progress with ryan alweiss. A natural analogue of the 2-large is large conjecture is the following conjecture: If D subseteq mathbb{N} is delta-density large for some delta in (0,1), then D is delta-density large for every delta in (0,1). I have a feeling that this might be easier to prove than the original 2-large is large conjecture, because we can use density increment/decrement arguments. Please try to prove my conjecture.

---

2. Assistant

2026-09-15 11:04:12 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T03:04:12.691363Z

I'll recover your precise definition and relevant earlier results, then test whether a density-increment argument can lower the density threshold and yield a proof.

---

3. Assistant

2026-09-15 11:07:14 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T03:07:14.724896Z

Upward monotonicity follows by thinning; the challenge is lowering density. I'm testing amplification by translates, while controlling the obstruction that forbidden differences may become multiples of the original difference.

---

4. Assistant

2026-09-15 11:16:17 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T03:16:17.202156Z

A useful reformulation tracks the maximal density of sets avoiding k-term progressions with differences in D; the conjecture asserts its large- k limit is always 0 or 1.

---

5. Assistant

2026-09-15 11:33:16 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T03:33:16.343916Z

I found a useful part of the density-increment strategy: **the critical density is unchanged by restricting the allowable differences to a Bohr neighborhood, or by passing from D to $\{n : qn \in D\}$.** There is also an unconditional partition-regularity result for the family of sets that are density large at *some* density.

I have not proved the remaining density-amplification step, so I am not claiming a solution to your conjecture. Below are the precise results and proofs, followed by the obstruction that remains. In particular, the difficulty is not an upper-density versus upper-Banach-density issue.

1. An exact extremal formulation of your conjecture

I use your definition: $D \subseteq \mathbb{N}$ is ** δ-density large** if every $A \subseteq \mathbb{N}$ with

$$\overline{d}(A) = \delta$$

contains arbitrarily long arithmetic progressions with common difference in D.

"Exactly δ" can equivalently be replaced by "at least δ." Indeed, if $\overline{d}(A) = a \geq \delta$, enumerate $A = \{a_1 < a_2 < \cdots\}$ and select a set of indices with counting function $\lfloor (\delta/a)m \rfloor$. The resulting subset $A' \subseteq A$ satisfies

$$|A' \cap [1, N]| = \left\lfloor \frac{\delta}{a} |A \cap [1, N]| \right\rfloor,$$

so $\overline{d}(A') = \delta$.

For $k \geq 2$, define

$$a_k(D) := \sup \left\{ \overline{d}(A) : A \subseteq \mathbb{N},\ A \text{ contains no } k\text{-term AP with difference in } D \right\}.$$

These numbers are nondecreasing in k. Put

$$a_\infty(D) := \lim_{k \to \infty} a_k(D).$$

The endpoint matters here, so it is useful to establish that the supremum defining $a_k(D)$ is attained.

Lemma 1: Attainment and finite formulations

For every $D \subseteq \mathbb{N}$ and $k \geq 2$, there is a set $A \subseteq \mathbb{N}$ having ordinary natural density $a_k(D)$ and containing no k-term AP with difference in D.

Moreover, if

$$r_k(D, N) := \max \left\{ |F| : F \subseteq [1, N],\ F \text{ contains no } k\text{-term AP with difference in } D \right\},$$

then

$$a_k(D) = \lim_{N \to \infty} \frac{r_k(D, N)}{N} = \inf_{N \geq 1} \frac{r_k(D, N)}{N}. \tag{1}$$

Proof. Restricting an admissible subset of $[1, N + M]$ to the first and last intervals gives

$$r_k(D, N + M) \leq r_k(D, N) + r_k(D, M).$$

Consequently the limit on the right of (1) exists.

Now consider the compact, shift-invariant space

$$X_{D,k} := \left\{ x \in \{0,1\}^{\mathbb{Z}} : \prod_{j=0}^{k-1} x_{n+jd} = 0 \text{ for every } n \in \mathbb{Z},\ d \in D \right\}.$$

Let $C = \{x : x_0 = 1\}$.

Every invariant probability measure μ on $X_{D,k}$ satisfies

$$N\mu(C) = \int \sum_{n=1}^{N} x_n \, d\mu(x) \le r_k(D,N).$$

Conversely, take extremal words of length N, extend them by zeros, and average their translates over intervals of length N. A weak-* limit is an invariant measure supported on $X_{D,k}$, with

$$\mu(C) = \lim_{N\to\infty} \frac{r_k(D,N)}{N}.$$

Choose an ergodic maximizing measure. For a generic point x, the set

$$A = \{n \ge 1 : x_n = 1\}$$

has natural density $\mu(C)$ and has the required avoidance property. Finally, any admissible $A \subseteq \mathbb{N}$ satisfies

$$|A \cap [1,N]| \le r_k(D,N),$$

which proves (1) and attainment. □

It follows that

$$\boxed{D \text{ is } \delta\text{-density large} \iff a_k(D) < \delta \quad \text{for every } k \ge 2.} \tag{2}$$

In particular,

$$D \text{ is } \delta\text{-density large for some } \delta < 1 \iff a_\infty(D) < 1, \tag{3}$$

whereas

$$D \text{ is } \delta\text{-density large for every } \delta > 0 \iff a_\infty(D) = 0. \tag{4}$$

Thus your conjecture is exactly the following zero–one statement:

$$\boxed{a_\infty(D) \in \{0,1\} \qquad \text{for every } D \subseteq \mathbb{N}.} \tag{5}$$

This is a particularly clean way to formulate the problem. It says:

> If there is a positive-density set avoiding k-term progressions with differences in D, for some k, then there are sets of density arbitrarily close to 1 avoiding sufficiently long progressions with differences in the same D.

The allowed progression length may increase, but the set of differences must remain unchanged.

For comparison, the color formulation fixes the number of colors and allows the progression length to vary; this is the formulation of r-largeness in Host–Kra–Maass, Section 7. arXiv

2. Products give a genuine partition-regularity result

The symbolic construction above also gives the equivalent dynamical description

$$a_k(D) = \max\left\{ \mu(A) : \mu\left(\bigcap_{j=0}^{k-1} T^{-jd} A \right) = 0 \text{ for every } d \in D \right\}, \tag{6}$$

where the maximum may be taken over invertible probability-preserving systems.

This makes the following inequality immediate.

Lemma 2: Product inequality

For $D, E \subseteq \mathbb{N}$ and $k, \ell \geq 2$,

$$a_{\max(k,\ell)}(D \cup E) \geq a_k(D) a_\ell(E).$$

Consequently,

$$\boxed{a_\infty(D) a_\infty(E) \leq a_\infty(D \cup E) \leq \min\{a_\infty(D), a_\infty(E)\}.} \tag{7}$$

Proof. Take systems and sets A, B realizing $a_k(D), a_\ell(E)$. In the product system, the set $A \times B$ has measure

$$\mu(A)\nu(B) = a_k(D) a_\ell(E).$$

A progression with difference in D is ruled out by its first coordinate, while one with difference in E is ruled out by its second coordinate. The upper bound follows from monotonicity in the set of forbidden differences. □

There is an exact fixed-density consequence.

Proposition 3: Density partition principle

Suppose

$$D = D_1 \cup \cdots \cup D_r, \qquad \delta_1, \ldots, \delta_r \in (0, 1).$$

If D is $(\delta_1 \cdots \delta_r)$-density large, then some D_i is δ_i-density large.

In particular, if D is δ-density large, then some cell of every r-partition of D is $\delta^{1/r}$-density large.

Proof. Otherwise choose k_i such that

$$a_{k_i}(D_i) \geq \delta_i.$$

With $K = \max_i k_i$, the product inequality gives

$$a_K(D) \geq \prod_{i=1}^{r} \delta_i,$$

contradicting (2). □

Therefore:

$$\boxed{\begin{array}{c}\text{The family of sets that are } \delta\text{-density large} \\ \text{for at least one } \delta \in (0,1) \text{ is partition regular.}\end{array}}$$

This is one respect in which your density analogue has a useful structural advantage: the corresponding Ramsey-property question for the family of 2-large sets is explicitly posed in Host–Kra–Maass, Question 7.6. The result above concerns the union over density thresholds, not partition regularity at a fixed threshold. arXiv

3. Bohr localization preserves the critical density exactly

Here is the strongest localization statement I obtained.

Theorem 4: Bohr localization

Let $B \subseteq \mathbb{N}$ be a Bohr neighborhood of zero. Then, for every $D \subseteq \mathbb{N}$,

$$\boxed{a_\infty(D \cap B) = a_\infty(D).} \tag{8}$$

The key point is that differences outside a Bohr neighborhood can be forbidden by a set of density arbitrarily close to 1.

Lemma 5: High-density witnesses outside a Bohr neighborhood

For every Bohr neighborhood B of zero,

$$a_\infty(\mathbb{N} \setminus B) = 1.$$

Proof, first for one frequency. Fix $\alpha \in \mathbb{T}$ and $\varepsilon > 0$, and put

$$E = \{d \geq 1 : \|d\alpha\| \geq \varepsilon\}.$$

We will construct an avoidance witness of measure arbitrarily close to 1.

Choose an open interval $I \subseteq \mathbb{T}$ of length $\eta > 0$, and let

$$C = \mathbb{T} \setminus I.$$

Thus $m(C) = 1 - \eta$.

For $L \geq 1$, define

$$K_L = \{\beta \in \mathbb{T} : \text{there is } x \in \mathbb{T} \text{ with } x + j\beta \in C \text{ for } 0 \leq j < L\}.$$

These are nested compact sets.

If $\beta \in \bigcap_L K_L$, compactness supplies x such that

$$x + j\beta \in C \qquad (j \geq 0).$$

Then β cannot be irrational. If β has reduced denominator q, its orbit consists of q equally spaced points. Such an orbit can avoid I only if

$$q \leq \eta^{-1}.$$

Hence, writing

$$Q = \operatorname{lcm}\bigl(1, 2, \ldots, \lfloor \eta^{-1} \rfloor\bigr),$$

we have

$$\bigcap_L K_L \subseteq \{\beta : Q\beta = 0\}.$$

Another compactness argument gives an L such that

$$K_L \subseteq \{\beta : \|Q\beta\| < \varepsilon\}. \tag{9}$$

Choose $\gamma \in \mathbb{T}$ with $Q\gamma = \alpha$, and consider the rotation

$$R(x) = x + \gamma.$$

If $d \in E$ and some x belonged to

$$\bigcap_{j=0}^{L-1} R^{-jd} C,$$

then $d\gamma \in K_L$. By (9),

$$\|d\alpha\| = \|Qd\gamma\| < \varepsilon,$$

a contradiction.

Thus C, of measure $1 - \eta$, avoids L-term progressions for every difference in E. Since η is arbitrary,

$$a_\infty(E) = 1.$$

For a basic Bohr neighborhood

$$B = \{d : \|d\alpha_i\| < \varepsilon_i,\ 1 \le i \le s\},$$

its complement is a finite union of sets of the preceding form. The product inequality gives

$$a_\infty(\mathbb{N} \setminus B) \ge 1^s = 1.$$

A general Bohr neighborhood contains a basic one, so the result follows by monotonicity. □

Proof of Theorem 4. Write

$$D = (D \cap B) \cup (D \setminus B).$$

Since $D \setminus B \subseteq \mathbb{N} \setminus B$, Lemma 5 implies

$$a_\infty(D \setminus B) = 1.$$

The product inequality therefore gives

$$a_\infty(D) \ge a_\infty(D \cap B) a_\infty(D \setminus B) = a_\infty(D \cap B).$$

The reverse inequality follows from $D \cap B \subseteq D$. □

In particular, every set that is density large at some density is Bohr recurrent. More strongly, its critical density does not deteriorate after Bohr localization.

There is a small endpoint distinction worth preserving:

> If D is δ-density large, then $D \cap B$ is δ'-density large for every $\delta' > \delta$.

The proof above does **not** automatically give $\delta' = \delta$, because $a_\infty(D)$ may equal δ even though every $a_k(D) < \delta$.

4. Passing to arithmetic progressions also preserves the critical density

There is an exact dilation identity at each fixed length:

$$\boxed{a_k(qD) = a_k(D) \qquad (q \geq 1).} \tag{10}$$

To see one inequality, if A avoids differences qD under T, then it avoids differences D under T^q.

For the reverse inequality, take a witness A for D under T, and form the cyclic tower

$$Y = X \times \{0, \dots, q-1\},$$

with

$$S(x,i) = \begin{cases} (x, i+1), & i < q-1, \\ (Tx, 0), & i = q-1. \end{cases}$$

Then

$$S^{qd}(x,i) = (T^d x, i).$$

The set $A \times \{0, \dots, q-1\}$ has the same measure as A and avoids differences qD.

Now define

$$D/q := \{n \geq 1 : qn \in D\}.$$

Since $q\mathbb{N}$ is a Bohr neighborhood of zero, (8) and (10) imply

$$\boxed{a_\infty(D/q) = a_\infty(D \cap q\mathbb{N}) = a_\infty(D).} \tag{11}$$

This is directly relevant to your proposed density-increment approach. If A contains no k-term progression with difference in D, then

$$A_{a,q} := \{n \geq 1 : a + qn \in A\}$$

contains no k-term progression with difference in D/q. Equation (11) says that this change of the difference set preserves the limiting critical density.

So **the arithmetic rescaling itself is not a loss in the proposed argument**.

5. Where the density increment still fails to close

There are two separate issues.

Unions of translates raise density but do not preserve the forbidden difference

Starting from a positive-density avoidance witness A, one would like to take

$$A' = \bigcup_{t \in F} (A - t)$$

and increase its density.

But a long progression

$$x, x+d, \dots, x+(L-1)d$$

in A' only gives a coloring of its indices according to which translate of A contains each point. Applying van der Waerden yields a progression in one translate with common difference

$$md$$

for some positive integer m, rather than necessarily d.

There is no reason for md to belong to D.

The simplest failure is

$$D = 2\mathbb{N} - 1, \qquad A = 2\mathbb{N}.$$

The set A avoids even two-term progressions with differences in D, but

$$A \cup (A + 1) = \mathbb{N}.$$

This is not a counterexample to your conjecture. It is a counterexample to the most direct proposed amplification operation.

Nor does (10) by itself fix this: dilation invariance is a statement about the extremal problem for mD, not permission to replace a particular $d \in D$ by a multiplier selected after the progression has been found.

A finite density increment must beat a recurrence scale that depends on the new dilation

Suppose

$$c := a_\infty(D) < 1$$

and choose $\gamma > c$.

By (11), for every q,

$$a_k(D/q) \le c < \gamma.$$

Thus Lemma 1 supplies a finite threshold $N(q, k, \gamma)$ such that sufficiently long subsets of density at least γ contain a k-term progression with difference in D/q.

Consequently, it would suffice to produce an arithmetic progression

$$P = a + q[1, L]$$

on which A has relative density at least γ, with

$$L \ge N(q, k, \gamma). \tag{12}$$

The problem is the dependence on q. Merely finding arbitrarily long progressions on which A is dense does not ensure (12), since the dilation changes along with the progression.

This dependence is not cosmetic. For example, for

$$D = \{n^2 : n \ge 1\}$$

and a prime q,

$$D/q = \{qm^2 : m \ge 1\}.$$

Its smallest allowed difference is q, so an interval of length less than

$$(k - 1)q + 1$$

cannot contain even one admissible k-term progression, regardless of its density.

Thus a successful finite density-increment proof needs simultaneous control of the density gain, the new dilation, and the available length. The localization theorem handles the limiting density parameter, but not this finite-scale control.

6. The precise missing amplification statement

A sufficient one-step lemma would be the following:

> **Density-amplification target.** There is an absolute constant $c_0 > 0$ such that, whenever A has density $a > 0$ and avoids k-term progressions with differences in D, there are K and a set A' avoiding K-term progressions with differences in the same D, with
>
> $$\overline{d}(A') \geq a + c_0 a(1-a). \tag{13}$$

Indeed, applying this to maximizing witnesses and letting $k \to \infty$ would give

$$a_\infty(D) \geq a_\infty(D) + c_0 a_\infty(D)\big(1 - a_\infty(D)\big),$$

which forces $a_\infty(D) \in \{0, 1\}$.

I do not have (13), or a substitute sufficient to rule out an intermediate value of $a_\infty(D)$.

A reduction only to nil-Bohr recurrence would not supply this missing step either: Ryan's construction intersects every nil-Bohr set while admitting a finite coloring with no monochromatic three-term progression along its differences. That example does **not** by itself refute your density conjecture; one would additionally have to establish density largeness at some fixed density. arXiv

My conclusion is that your density approach has real additional structure: the "some density" family is partition regular, and its critical density survives Bohr localization and arithmetic rescaling exactly. But the central assertion remains the zero–one law (5), and none of the amplification operations above proves it. The literature I checked supplies the nearby color formulations and nil-Bohr obstructions, but I did not find a result that supplies this density-amplification step. arXiv arXiv

---

6. User

2026-09-15 12:02:13 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T04:02:13.380000Z

Good progress. Please keep tryinng to prove my conjecture.

---

7. Assistant

2026-09-15 12:02:28 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T04:02:28.145138Z

I'll pursue a critical-density amplification that preserves differences in D exactly, specifically avoiding the multiplier obstruction from unions of translates.

8. Assistant

2026-09-15 12:08:36 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T04:08:36.178008Z

A product-based amplification has a new obstacle: progressions may alternate between coordinates. I'm testing whether an extension can block this, possibly guided by the nilsystem construction.

---

9. Assistant

2026-09-15 12:26:45 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T04:26:45.929424Z

A promising refinement is to run product coordinates at different speeds; I'm checking whether this can force any extracted subprogression to align with the correct coordinate.

---

10. Assistant

2026-09-15 12:45:15 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T04:45:15.553545Z

I obtained a **density-amplification lemma for every fixed progression length**, with the set D unchanged. It has two consequences that were missing from the previous answer:

$$\boxed{\text{At each fixed length, the density threshold improves by an amount independent of } D.}$$

And, writing $c(D) = a_\infty(D)$,

$$\boxed{0 < c(D) < 1 \quad \Longrightarrow \quad D \text{ is } c(D)\text{-density large.}}$$

The amplification is not yet uniform in the progression length, so **this does not complete the conjecture**. But it makes your proposed density-increment approach work in a precise form, and it removes the endpoint loss in the localization results from my previous answer.

1. A density-amplification theorem that preserves D

Retain the notation

$$a_k(D) = \sup\left\{\overline{d}(A) : A \subseteq \mathbb{N} \text{ contains no } k\text{-term AP with difference in } D\right\},$$

and

$$c(D) = \lim_{k\to\infty} a_k(D).$$

Recall the dynamical formulation established previously:

$$a_k(D) = \max \mu(A),$$

where the maximum is over invertible probability-preserving systems and measurable sets satisfying

$$\mu\left(\bigcap_{j=0}^{k-1} T^{-jd}A\right) = 0 \qquad (d \in D). \tag{1}$$

We will use finite van der Waerden numbers. Let $W(k,r)$ denote the least W such that every r-coloring of $\{0,\dots,W-1\}$ contains a monochromatic k-term AP. Their finiteness is the finite van der Waerden theorem; this formulation also appears as Theorem 2 in Robertson's discussion of the 2-large conjecture. arXiv

Theorem 1 — Length-dependent density amplification

For $k \geq 2$, set

$$W_k = W(k,k), \qquad M_k = \left\lfloor \frac{W_k - 1}{k-1} \right\rfloor, \qquad N_k = W_k + k - 1.$$

Then, for every $D \subseteq \mathbb{N}$,

$$\boxed{a_{N_k}(D) \geq a_k(D) + \big(1 - a_k(D)\big) a_k(D)^{M_k}.} \tag{2}$$

In particular, whenever $a_k(D) \in (0,1)$,

$$a_{N_k}(D) > a_k(D).$$

The constants here are not optimized. The important point is that **the gain depends only on k and the starting density, not on D**.

Step 1: An asymmetric coloring lemma

Every red-blue coloring of

$$\{0,\dots,N_k - 1\}$$

contains either a block of k consecutive red points, or a blue k-term AP with common difference in $\{1,\dots,M_k\}$.

Indeed, suppose there is no block of k consecutive red points. For each

$$n \in \{0,\dots,W_k - 1\},$$

choose

$$t(n) \in \{0,\dots,k-1\}$$

such that $n + t(n)$ is blue. This is possible because the interval

$$\{n,\dots,n+k-1\}$$

cannot be entirely red.

The map $n \mapsto t(n)$ is a k-coloring of an interval of length W_k. Hence there are b, q, t such that

$$t(b) = t(b+q) = \cdots = t(b+(k-1)q) = t.$$

Consequently,

$$b + t,\ b + t + q,\ \dots,\ b + t + (k-1)q$$

are all blue. Moreover,

$$1 \le q \le \left\lfloor \frac{W_k - 1}{k-1} \right\rfloor = M_k.$$

This proves the coloring lemma.

Step 2: Build roots to neutralize the extracted multiplier

Take a system (X, μ, T) and a set A satisfying (1), and write

$$a = \mu(A).$$

For each $q \in \{1, \dots, M_k\}$, form the cyclic tower

$$Y_q = X \times \{0, \dots, q-1\},$$

with product probability measure ν_q, and transformation

$$R_q(x,i) = \begin{cases} (x, i+1), & i < q-1, \\ (Tx, 0), & i = q-1. \end{cases}$$

Let

$$C_q = A \times \{0, \dots, q-1\}.$$

Then

$$\nu_q(C_q) = a$$

and

$$R_q^{qd}(x,i) = (T^d x, i).$$

It follows that

$$\nu_q\left(\bigcap_{j=0}^{k-1} R_q^{-jqd} C_q \right) = 0 \qquad (d \in D). \tag{3}$$

Now take

$$(Y, \nu, U) = \prod_{q=1}^{M_k} (Y_q, \nu_q, R_q), \qquad C = \prod_{q=1}^{M_k} C_q.$$

Thus

$$\nu(C) = a^{M_k}.$$

For every $d \in D$ and every $q \in \{1, \dots, M_k\}$, the q-th coordinate of this product gives

$$\nu\left(\bigcap_{j=0}^{k-1} U^{-jqd} C \right) = 0. \tag{4}$$

This is the crucial repair to the multiplier problem:

> A progression with difference qd in C would produce, in its q-th root coordinate, a progression with difference exactly d in A.

We have not assumed that $qd \in D$.

Step 3: Take a union to increase the measure

On

$$(Z, \rho, S) = (X \times Y, \mu \times \nu, T \times U),$$

define

$$B = (A \times Y) \cup (X \times C).$$

Its measure is

$$\rho(B) = a + (1-a)a^{M_k}. \tag{5}$$

I claim that

$$\rho\left(\bigcap_{j=0}^{N_k-1} S^{-jd}B\right) = 0 \qquad (d \in D). \tag{6}$$

To see this, suppose an orbit segment

$$z, S^d z, \dots, S^{(N_k-1)d} z$$

lies in B. Color index j red when its first coordinate lies in A, and blue otherwise. At every blue index, its second coordinate must lie in C.

Outside the relevant finite union of null sets, condition (1) rules out k consecutive red indices. The coloring lemma therefore supplies a blue k-term AP with some common difference $q \le M_k$. Its second coordinates lie in C along a U-orbit with step qd, contradicting (4).

Thus (6) holds. By the dynamical characterization of $a_{N_k}(D)$,

$$a_{N_k}(D) \ge a + (1-a)a^{M_k}.$$

Taking $a = a_k(D)$ proves (2). □

2. What this proves about decreasing the density threshold

Fix $\delta \in (0,1)$ and $k \ge 2$. Define

$$F_k(t) = t + (1-t)t^{M_k}.$$

This function is strictly increasing on $[0,1]$, and

$$F_k(0) = 0, \qquad F_k(\delta) > \delta.$$

Therefore there is a unique

$$\beta_{k,\delta} \in (0, \delta)$$

such that

$$\beta_{k,\delta} + \big(1 - \beta_{k,\delta}\big)\beta_{k,\delta}^{M_k} = \delta. \tag{7}$$

Corollary 2 — A uniform decrement at each fixed length

If D is δ-density large, then

$$\boxed{a_k(D) < \beta_{k,\delta} < \delta.} \tag{8}$$

Equivalently, **every set of upper density at least $\beta_{k,\delta}$ contains a k-term AP with difference in D.**

Indeed, δ-density largeness gives

$$a_{N_k}(D) < \delta.$$

Theorem 1 then gives

$$F_k(a_k(D)) < \delta,$$

and strict monotonicity yields (8).

A simpler, less sharp explicit decrement is

$$\boxed{a_k(D) < \delta - (1-\delta)\left(\frac{\delta}{2}\right)^{M_k}.} \tag{9}$$

For if $a_k(D)$ were at least the right-hand side, it would be at least $\delta/2$, and hence

$$a_k(D) + (1 - a_k(D))a_k(D)^{M_k} \geq \delta,$$

a contradiction.

This is stronger than merely observing that each $a_k(D) < \delta$: it gives a positive gap **uniform over every D satisfying the same δ-density-largeness hypothesis**.

In fact, the hypothesis needed for this particular conclusion is only that every density- δ set contains an N_k -term AP with difference in D.

3. The critical density is included

The same amplification resolves the endpoint issue from the previous answer.

Corollary 3 — Critical-density attainment

For every $D \subseteq \mathbb{N}$ with

$$0 < c(D) < 1,$$

one has

$$a_k(D) < c(D) \qquad \text{for every } k \geq 2.$$

Consequently, D is $c(D)$-density large.

Proof. Suppose $a_k(D) = c(D)$ for some k. Theorem 1 would imply

$$c(D) \geq a_{N_k}(D) \geq c(D) + (1 - c(D))c(D)^{M_k} > c(D),$$

a contradiction. □

Thus, for every $\delta \in (0,1)$, we now have the exact characterization

$$\boxed{D \text{ is } \delta\text{-density large} \quad \Longleftrightarrow \quad c(D) \le \delta.} \tag{10}$$

The forward implication follows by taking the supremum of $a_k(D) < \delta$. For the reverse implication, the only previously unresolved case was $c(D) = \delta$, which Corollary 3 handles.

In particular, a counterexample to your conjecture would have a genuine minimum admissible density:

$$\{\delta \in (0,1) : D \text{ is } \delta\text{-density large}\} = [c(D), 1)$$

for some $c(D) \in (0,1)$.

It cannot exhibit the alternative endpoint behavior $(c(D), 1)$.

Exact localization now follows

Previously we proved

$$c(D \cap B) = c(D)$$

for a Bohr neighborhood B of zero. Combining this with (10) gives

$$\boxed{D \text{ is } \delta\text{-density large} \quad \Longleftrightarrow \quad D \cap B \text{ is } \delta\text{-density large.}} \tag{11}$$

There is **no need to replace** δ **by a larger** δ'.

Similarly, for

$$D/q = \{n \in \mathbb{N} : qn \in D\},$$

the previously established identity $c(D/q) = c(D)$ now gives

$$\boxed{D \text{ is } \delta\text{-density large} \quad \Longleftrightarrow \quad D/q \text{ is } \delta\text{-density large.}} \tag{12}$$

So a density-increment argument may pass to arithmetic progressions and rescale the differences without any loss in the original density-largeness hypothesis.

4. A considerably better amplification for pairs

For $k = 2$, the root construction can be made much more efficient.

Proposition 4

For every $D \subseteq \mathbb{N}$,

$$\boxed{a_5(D) \ge 2a_2(D) - a_2(D)^2.} \tag{13}$$

Proof. Let $A \subseteq X$, $\mu(A) = a$, satisfy

$$\mu(A \cap T^{-d}A) = 0 \qquad (d \in D).$$

Take just the square-root tower

$$Y = X \times \{0,1\}, \qquad R(x,0) = (x,1), \quad R(x,1) = (Tx, 0),$$

and put

$$C = A \times \{0, 1\}.$$

Then

$$\nu(C) = a, \qquad \nu(C \cap R^{-2d}C) = 0 \qquad (d \in D).$$

On $X \times Y$, with transformation $T \times R$, let

$$B = (A \times Y) \cup (X \times C).$$

Thus

$$(\mu \times \nu)(B) = 2a - a^2.$$

A five-term orbit segment in B would give a red-blue coloring of $\{0, 1, 2, 3, 4\}$ in which no red pair has distance 1, and no blue pair has distance 2.

Such a coloring is impossible. If an internal index $j \in \{1, 2, 3\}$ is red, its two neighbors must be blue, contradicting the blue distance- 2 restriction. Therefore $1, 2, 3$ are all blue, which again gives a blue pair at distance 2.

The corresponding dynamical intersection has measure zero, proving (13). □

Consequently,

$$D \text{ is } \delta\text{-density large} \quad \Longrightarrow \quad \boxed{a_2(D) < 1 - \sqrt{1 - \delta}.} \tag{14}$$

For example, if D is $1/2$-density large, then every set of upper density at least

$$1 - \frac{1}{\sqrt{2}} \approx 0.292893$$

contains two elements whose positive difference belongs to D.

This is a substantial decrement for pairs, although it does not yet give recurrence for arbitrarily small positive densities.

5. A further reduction using the nilrecurrence work

There is also a useful structural consequence when we incorporate the high-density nil-Bohr obstruction from the earlier write-up.

That write-up establishes that, for every finite-step nil-Bohr neighborhood B of zero and every $\eta < 1$, there is a set of upper density η whose sufficiently long APs all have differences in B. Thus

$$c(\mathbb{N} \setminus B) = 1.$$

This is precisely the “High-density nil-Bohr obstruction” theorem in the file.
[fixed_density_nilrecurrence.tex, L334-L345]

Using the product inequality from the preceding answer, we therefore obtain

$$c(D \cap B) = c(D)$$

for **finite-step nil-Bohr neighborhoods**, not merely Bohr neighborhoods. By (10), this also preserves δ-density largeness at the exact same δ.

Now the Bergelson–Host–Kra recurrence structure, in the form stated by Huang–Shao–Ye, says that for an ergodic system, a positive-measure set A, and a fixed k, there are a finite-step nil-Bohr neighborhood B and a set E of zero upper Banach density such that

$$B \setminus E \subseteq \left\{ n \geq 1 : \mu\left(\bigcap_{j=0}^{k-1} T^{-jn} A \right) > 0 \right\}.$$

The positivity of the nilsequence at its base point and the zero-upper-Banach-density exceptional set are addressed immediately before and in their Theorem 2.2.4. arXiv

It follows that:

> **If D is a counterexample to your conjecture, then some finite-step nil-Bohr neighborhood B satisfies**
>
> $$d^*(D \cap B) = 0, \qquad c(D \cap B) = c(D). \tag{15}$$

Indeed, choose k with $a_k(D) > 0$ and an ergodic avoidance witness A. Every $d \in D$ has zero k-fold intersection, so the displayed recurrence inclusion forces

$$D \cap B \subseteq E.$$

Nil-Bohr localization preserves $c(D)$, giving (15).

Thus any intermediate critical density survives inside a uniformly sparse exceptional set, even after a single finite-step nil-Bohr restriction. This pinpoints why nilrecurrence alone does not settle the density question: the remaining obstruction can be concentrated entirely on exceptional times.

6. What remains to finish the conjecture

The new general inequality is

$$a_{N_k}(D) - a_k(D) \geq (1 - a_k(D)) a_k(D)^{M_k}. \tag{16}$$

This proves a positive increment at every finite stage where $a_k(D) > 0$. But M_k depends on k, and the bound becomes very small as k grows.

If we try to iterate by setting

$$k_{j+1} = N_{k_j},$$

we get increasing lower bounds on the admissible avoidance densities, but the successive guaranteed increments might be summable. **Strict increase at every stage does not force the limit to equal 1.**

The two-root construction illustrates why a direct iteration is not automatic. For pairs, the two forbidden patterns—red distance 1, blue distance 2—cannot coexist on a long interval. For three-term progressions, the periodic coloring

$$\mathrm{RRBBRRBB}\cdots$$

has neither three consecutive red points nor a blue three-term progression with common difference 2. Thus the same two-root argument cannot simply be reapplied to a longer-progression avoidance witness.

The general construction overcomes this by supplying a root coordinate for **every possible extracted multiplier**. That succeeds, but costs the factor a^{M_k}.

So the precise remaining challenge is now:

$$\boxed{\text{Can the multiplier-neutralizing construction be made with a density gain that stays positive independently of the original forbidden length } k?}$$

A bound

$$a_{K(k,a)}(D) \geq a + \varepsilon(a), \qquad \varepsilon(a) > 0,$$

whenever $a_k(D) \geq a$, with $\varepsilon(a)$ independent of k, would force the desired zero–one law.

The progress here is that density amplification itself now works, with D exactly preserved; the unresolved issue is its uniformity in the progression length. The argument also gives exact critical-density attainment and exact-density localization, which remove two possible sources of loss from a subsequent proof.

---

11. User

2026-09-15 15:06:56 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T07:06:56.822000Z

Great progress! Please keep working on this.

---

12. Assistant

2026-09-15 15:07:13 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T07:07:13.913893Z

I'll focus on making the amplification gain uniform in progression length, while checking whether the endpoint result and existing construction provide enough reusable structure.

---

13. Assistant

2026-09-15 15:10:17 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T07:10:17.108166Z

The bottleneck is the product-measure cost from separate root towers, so I'm testing whether one joint extension can neutralize all needed multipliers at once.

---

14. Assistant

2026-09-15 15:19:20 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T07:19:20.453104Z

I found that adjoining finitely many integer multiples of D preserves its critical density exactly. I'm testing this alongside length-two auxiliary patterns to avoid worsening constants.

15. Assistant

2026-09-15 15:54:34 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T07:54:34.214502Z

I obtained a sharper amplification inequality, together with a quantitative obstruction to making the previous construction uniform. The new inequality is

$$\boxed{a_{2k}(D) \geq a_k(D) + (1 - a_k(D))\, a_2\left(\bigcup_{q=1}^{k} qD\right).} \tag{1}$$

In particular,

$$\boxed{a_{2k}(D) \geq a_k(D) + (1 - a_k(D))a_2(D)^k.} \tag{2}$$

Thus, whenever there is a positive-density **pair-avoidance** witness, we can increase the density while only **doubling** the forbidden progression length. This is substantially better than the van der Waerden length increase in my previous answer.

I still do not have the uniform amplification needed to prove your conjecture. However, I can now identify two limitations of the root construction rigorously, rather than merely observing that its estimates deteriorate. I also obtained a structural restriction on any near-extremal avoidance witness.

1. A more flexible amplification principle

Recall

$$a_k(D) = \max\left\{\mu(A) : \mu\left(\bigcap_{j=0}^{k-1} T^{-jd}A\right) = 0 \text{ for every } d \in D\right\},$$

where we may work with invertible probability-preserving systems, and

$$c(D) = \lim_{k\to\infty} a_k(D).$$

The conjecture is equivalent to

$$c(D) \in \{0, 1\}.$$

The useful refinement is to allow the original avoidance witness and the auxiliary witness to have **different forbidden lengths**.

Proposition 1: Finite-multiplier amplification

Let $k, \ell, N \geq 2$, and let $F \subseteq \mathbb{N}$ be finite. Suppose every red-blue coloring of

$$\{0, \dots, N-1\}$$

contains either k consecutive red indices, or a blue ℓ-term arithmetic progression with common difference in F.

Write

$$FD = \bigcup_{q\in F} qD.$$

Then

$$\boxed{a_N(D) \geq a_k(D) + (1 - a_k(D))a_\ell(FD).} \tag{3}$$

Consequently,

$$\boxed{a_N(D) \geq a_k(D) + (1 - a_k(D))a_\ell(D)^{|F|}.} \tag{4}$$

Proof. Take a k-term avoidance witness $A \subseteq X$ for D, and an ℓ-term avoidance witness $C \subseteq Y$ for FD. Write their transformations as T, U, and their measures as

$$\mu(A) = a, \qquad \nu(C) = b.$$

On the product system, put

$$B = (A \times Y) \cup (X \times C).$$

Then

$$(\mu \times \nu)(B) = a + (1 - a)b. \tag{5}$$

Fix $d \in D$. Suppose an orbit segment

$$(x, y), (T^d x, U^d y), \ldots, (T^{(N-1)d} x, U^{(N-1)d} y)$$

lies entirely in B. Color index j red when $T^{jd}x \in A$, and blue otherwise. Every blue index then satisfies $U^{jd}y \in C$.

A block of k consecutive red indices gives a forbidden k-term progression in A with difference d. A blue ℓ-term progression with index difference $q \in F$ gives a forbidden ℓ-term progression in C with difference $qd \in FD$.

More formally, the event that the whole orbit segment lies in B is contained in a finite union of translates of the corresponding zero-measure intersection events. Hence

$$(\mu \times \nu)\left(\bigcap_{j=0}^{N-1} (T \times U)^{-jd} B\right) = 0.$$

Taking maximizing witnesses proves (3).

Finally, the product inequality and dilation identity from the preceding answers give

$$a_\ell(FD) \geq \prod_{q \in F} a_\ell(qD) = a_\ell(D)^{|F|}.$$

This proves (4). □

The distinction between (3) and (4) is important. Formula (3) permits **any** auxiliary witness for FD, including one built using a correlated construction. Formula (4) is what independent root coordinates guarantee.

A mixed-length van der Waerden consequence

Let $W(\ell, k)$ be the van der Waerden number for an ℓ-term progression in a k-coloring. Its finiteness is the finite van der Waerden theorem. arXiv Put

$$W = W(\ell, k), \qquad M = \left\lfloor \frac{W-1}{\ell-1} \right\rfloor, \qquad N = W + k - 1.$$

Then

$$\boxed{a_N(D) \geq a_k(D) + (1 - a_k(D)) a_\ell(D)^M.} \tag{6}$$

Indeed, if there are no k consecutive red indices, then for each $n < W$ choose

$$t(n) \in \{0, \dots, k-1\}$$

such that $n + t(n)$ is blue. A monochromatic ℓ-term progression for the coloring $n \mapsto t(n)$ gives a blue ℓ-term progression whose common difference is at most M.

Unlike the previous amplification theorem, **we may keep ℓ fixed while k increases**. Thus we can keep reusing an initial positive-density ℓ-term avoidance witness.

The case $\ell = 2$ gives a particularly clean result.

2. Doubling the length is enough when the auxiliary witness avoids pairs

Define

$$D^{[k]} = \bigcup_{q=1}^{k} qD, \qquad p_k(D) = a_2(D^{[k]}), \qquad p = a_2(D).$$

Theorem 2: Pair-assisted amplification

For every $k \geq 2$,

$$\boxed{a_{2k}(D) \geq a_k(D) + (1 - a_k(D)) p_k(D) \geq a_k(D) + (1 - a_k(D)) p^k.} \tag{7}$$

Proof of the coloring step. Consider a red-blue coloring of

$$\{0, \dots, 2k-1\}$$

with no k consecutive red indices.

There must be at least two blue indices: otherwise the at most one blue index separates the interval into red blocks whose combined length is at least $2k - 1$, so one has length at least k.

Now take two consecutive blue indices. Their distance is at most k, since a larger distance would leave at least k consecutive red indices between them.

Therefore the coloring contains either k consecutive red indices, or a blue pair whose distance belongs to $\{1, \dots, k\}$. Proposition 1 applies.

For the second inequality, take a pair-avoidance witness of measure p. For each $q \leq k$, its q-root tower supplies a witness of the same measure avoiding pairs at differences in qD. Their product has measure p^k and avoids pairs at every difference in $D^{[k]}$. Thus

$$p_k(D) \geq p^k.$$

□

The length $2k$ is optimal for this elementary coloring argument: on $2k-1$ indices,

$$\underbrace{R\cdots R}_{k-1}B\underbrace{R\cdots R}_{k-1}$$

has no k consecutive red indices and has only one blue index.

Consequence for a fixed density threshold

Suppose D is δ-density large. Then $a_{2k}(D)<\delta$, so

$$a_k(D)+(1-a_k(D))p^k<\delta.$$

Rearranging gives

$$\boxed{a_k(D)<\frac{\delta-p^k}{1-p^k}=\delta-\frac{(1-\delta)p^k}{1-p^k}.} \tag{8}$$

When $p>0$, this is an explicit decrement below δ at every length k, obtained by testing the original hypothesis only at length $2k$.

Likewise, if $c=c(D)<1$, then

$$c-a_k(D)\geq(1-a_k(D))p^k\geq(1-c)p^k. \tag{9}$$

Thus a hypothetical intermediate critical density with $a_2(D)>0$ must satisfy the quantitative endpoint separation

$$\boxed{c(D)-a_k(D)\geq(1-c(D))a_2(D)^k.} \tag{10}$$

This strengthens the qualitative conclusion $a_k(D)<c(D)$ proved previously.

There is an important restriction: if $a_2(D)=0$, this particular estimate gives no new gain. Such a D is already a set of measurable recurrence for pairs, but it could still be relevant to the conjecture through a higher-order avoidance witness. The mixed-length formula (6) remains available in that case.

3. Why a shared auxiliary extension does not immediately finish the argument

I initially hoped that the factor p^k was merely the cost of making the root coordinates independent. Formula (3) lets us replace that product construction by the best possible auxiliary witness.

But even the optimum has an unavoidable decay.

Proposition 3: An upper bound on the auxiliary density

If $D\neq\varnothing$, then

$$\boxed{p_k(D)=a_2(D^{[k]})\leq\frac{1}{k+1}.} \tag{11}$$

Proof. Fix $d\in D$, and let C be a pair-avoidance witness for $D^{[k]}$. The sets

$$C,\ U^{-d}C,\ U^{-2d}C,\dots,U^{-kd}C$$

are pairwise disjoint modulo null sets.

Indeed, for $0 \le i < j \le k$, their intersection has the same measure as

$$C \cap U^{-(j-i)d} C,$$

which is zero because $(j-i)d \in D^{[k]}$.

Therefore

$$(k+1)\nu(C) \le 1.$$

Taking the maximum proves (11). □

So the pair-assisted construction gives

$$p^k \le p_k(D) \le \frac{1}{k+1}, \tag{12}$$

and **correlating the root coordinates cannot make this auxiliary measure bounded below independently of k.**

This does not rule out using the optimal $p_k(D)$ more cleverly. It does rule out one proposed repair: replace the independent product by a shared extension of uniformly positive auxiliary measure while retaining pair avoidance for all $d, 2d, \dots, kd$.

The failure of the direct iteration can also be seen explicitly. Applying (7) at lengths

$$k_j = 2^j k_0$$

gives

$$1 - a_{k_{j+1}}(D) \le (1 - a_{k_j}(D))(1 - p^{k_j}).$$

Consequently,

$$c(D) \ge 1 - (1 - a_{k_0}(D)) \prod_{j=0}^{\infty} (1 - p^{2^j k_0}). \tag{13}$$

For $0 < p < 1$, this infinite product is positive, since

$$\sum_{j=0}^{\infty} p^{2^j k_0} < \infty.$$

Thus the resulting bound is genuinely stronger, but it does not force $c(D) = 1$.

4. The original same-length construction necessarily uses many multipliers

There is a different limitation for the previous construction, where both the original and auxiliary witnesses avoid k-term progressions.

It is not just that the available van der Waerden bound is poor. **Any finite-multiplier coloring argument of that specific form requires exponentially many multipliers.**

Theorem 4: A lower bound on the multiplier set

Let $F \subseteq \mathbb{N}$ be finite. Suppose there is an N such that every red-blue coloring of $\{0, \dots, N-1\}$ contains either

$$k \text{ consecutive red indices}$$

or

$$\text{a blue } k\text{-term AP with common difference in } F.$$

Then

$$\boxed{|F| + 1 > \frac{2^k}{e\,k^2}.} \tag{14}$$

Proof. Write $r = |F|$. We prove the contrapositive using the symmetric Lovász local lemma: events of probability at most p_0, each dependent on at most Δ others, can all be avoided when

$$e\,p_0(\Delta + 1) \le 1.$$

This follows from the general local lemma by taking a constant value of its parameter x; a primary source stating the general form is Moser–Tardos, Theorem 1.1. arXiv

Color a finite interval independently red or blue, each with probability $1/2$. Declare the following events bad: a particular k-term consecutive block is entirely red, or a particular k-term AP with difference in F is entirely blue.

Each bad event has probability

$$2^{-k}.$$

A vertex belongs to at most k red-block events, and at most k blue-progression events for each $q \in F$. Hence it belongs to at most

$$k(r+1)$$

bad events in total.

Each bad event involves k vertices. Therefore its closed dependency neighborhood has size at most

$$k^2(r+1).$$

The local lemma applies whenever

$$e\,2^{-k}k^2(r+1) \le 1. \tag{15}$$

Under (15), every finite interval admits a coloring with neither forbidden pattern. In particular, no N has the asserted universal coloring property. This proves (14). □

For the construction using independent copies of a density- a witness, the auxiliary measure is

$$a^{|F|}.$$

At fixed $a \in (0,1)$, (14) shows that this factor cannot remain bounded below as k grows. In fact, the forced number of factors is already exponential in k.

The scope of this obstruction is precise:

> Improving the van der Waerden estimates, while retaining the same “red consecutive block versus blue progression” argument and independent root intersection, cannot produce a length-uniform density increment.

It does **not** exclude a correlated construction at the increasing auxiliary length k, a different rule for forming the enlarged set, or an argument that uses additional structure of the original avoidance witness.

5. A structural constraint on near-extremal witnesses

There is another consequence of the exact arithmetic localization identity

$$c(D/q) = c(D)$$

from the preceding answers. It places a useful restriction on where a density increment could come from.

Proposition 5: Conditional-density bound on periodic factors

Let A be a k-term avoidance witness for D in an invertible probability-preserving system (X, μ, T). Put

$$a = \mu(A), \qquad c = c(D).$$

Then, for every $q \geq 1$,

$$\boxed{\mathbb{E}(1_A \mid \mathcal{I}(T^q)) \leq a_k(D/q) \leq c \quad \text{almost everywhere.}} \tag{16}$$

Here $\mathcal{I}(T^q)$ denotes the T^q-invariant sigma-algebra.

Proof. Let $H \in \mathcal{I}(T^q)$ have positive measure. Restrict to H, normalize the measure, and use the transformation T^q.

Whenever $n \in D/q$, we have $qn \in D$, and hence

$$\mu\left(H \cap \bigcap_{j=0}^{k-1} (T^q)^{-jn} A\right) = 0.$$

Thus $A \cap H$, in the normalized system on H, is a k-term avoidance witness for D/q. Consequently,

$$\frac{\mu(A \cap H)}{\mu(H)} \leq a_k(D/q) \leq c(D/q) = c.$$

Since this holds for every such invariant H, it is exactly the conditional-expectation bound. □

Let

$$\mathcal{K}_{\mathrm{rat}} = \bigvee_{m \geq 1} \mathcal{I}(T^{m!}), \qquad g = \mathbb{E}(1_A \mid \mathcal{K}_{\mathrm{rat}}).$$

The sigma-algebras in this join are increasing. Martingale convergence and (16) give

$$0 \leq g \leq c.$$

Since $\int g \, d\mu = a$,

$$\begin{aligned}\|g-a\|_2^2 &= \int g^2\,d\mu - a^2\\ &\le c\int g\,d\mu - a^2\\ &= a(c-a).\end{aligned}$$

Therefore

$$\boxed{\|\mathbb{E}(1_A \mid \mathcal{K}_{\mathrm{rat}}) - a\|_2 \le \sqrt{a(c-a)}.} \tag{17}$$

In particular, if

$$a \ge c-\varepsilon,$$

then

$$\|\mathbb{E}(1_A \mid \mathcal{K}_{\mathrm{rat}}) - a\|_2 \le \sqrt{c\varepsilon}. \tag{18}$$

Thus **near-extremal avoidance witnesses are nearly balanced on their entire rational factor**, uniformly over the periods involved.

This matters for your proposed arithmetic density-increment argument. A genuine asymptotic density increase on a residue component can never exceed $c(D)$, and a near-extremal witness has very little rational-factor variance available in the first place. An argument that crosses the critical density by passing to arithmetic progressions must therefore exploit a **finite-scale** density increase, with the necessary control of the progression length and its dilation—not merely a persistent density bias on a periodic component.

This conclusion concerns the rational factor. It does not establish the analogous bound for irrational eigenfunctions or higher-order nilfactors.

6. Where this leaves the proof attempt

The stronger amplification separates two possibilities for a hypothetical counterexample.

When $a_2(D) > 0$, there is an elementary amplification with only a doubling of length:

$$a_{2k}(D) \ge a_k(D) + (1-a_k(D))a_2(D)^k.$$

When pair recurrence already holds but some higher-order recurrence fails, the mixed-length formula allows us to reuse a fixed higher-order avoidance witness:

$$a_{W(\ell,k)+k-1}(D) \ge a_k(D) + (1-a_k(D))a_\ell(D)^{\lfloor (W(\ell,k)-1)/(\ell-1)\rfloor}.$$

Neither estimate yet excludes $0 < c(D) < 1$. The quantitative obstructions explain why the most immediate repairs do not work: in the pair-assisted version, even the optimal auxiliary measure tends to zero; in the previous same-length version, exponentially many independent root coordinates are unavoidable for that coloring argument.

A sufficient next step remains an increment

$$a_K(D) \ge a_k(D) + \varepsilon$$

for avoidance densities in a fixed compact subinterval of $(0,1)$, where $\varepsilon > 0$ is independent of k, although K may depend on k. Such a statement would contradict $a_k(D) \to c(D) \in (0,1)$.

What is now clearer is that this increment must use more than the bounded-gap condition on the complement that drives the current coloring argument. The new doubling inequality gives a better finite-stage tool; the multiplier lower bound identifies a genuine limitation of the earlier tool; and the rational-factor estimate restricts the structure of any near-extremal obstruction. None of these, individually or together, yet supplies the uniform increment needed to complete your conjecture.

---

16. User

2026-09-15 16:44:14 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T08:44:14.781000Z

consider the equivalent definition of delta-density regular in terms of measure-preserving systems that you mentioned earlier. We can rephrase the result from the other chat that if a set is delta-density regular for some delta in (0,1), then for the class of nilsystems it is delta density regular for every delta in (0,1). Can we prove the stronger theorem that if the set is delta density regular (for all systems) for some dleta in (0,1), then it is delta-density regular for the class of distal systems for all delta in (0,1)?

---

17. Assistant

2026-09-15 16:44:35 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T08:44:35.241273Z

I'll test whether the nilsystem argument survives compact extensions, while carefully distinguishing measurable distal systems from topological distal systems where necessary.

---

18. Assistant

2026-09-15 17:02:59 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T09:02:59.313051Z

The inverse-limit step does work via conditional densities, so the real remaining obstacle is preserving recurrence through a compact extension with an arbitrary cocycle.

---

19. Assistant

2026-09-15 17:12:02 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T09:12:02.593415Z

A significant warning is that Alweiss's obstruction already has a two-step distal compact-group skew-product realization; proving this theorem would therefore settle its unresolved fixed-density status as well.

20. Assistant

2026-09-15 17:20:20 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T09:20:20.743966Z

The inverse-limit extension can be proved completely, but I have not yet proved the compact-extension step needed for arbitrary distal systems. Your proposed theorem is nevertheless a useful intermediate target. In particular, even its **two-step distal case would imply that Alweiss's example is not δ-density regular for any $\delta < 1$.**

I will explain that implication explicitly below. It comes from realizing Alweiss's obstruction inside a compact-group skew product, rather than merely inside a symbolic system.

First, here is the part of the proposed theorem that does go through.

1. Precise formulation

Write

$$\mathrm{MR}_D(X, \mu, T)$$

for the property

$$\forall A \in \mathcal{B},\ \mu(A) > 0,\ \forall k \geq 2,\ \exists d \in D: \quad \mu\left(\bigcap_{j=0}^{k-1} T^{-jd} A\right) > 0.$$

Your hypothesis, for some fixed $\delta_0 \in (0, 1)$, is

$$\begin{gathered}\text{For every probability-preserving system } (X, \mu, T),\\ \text{every } A \text{ with } \mu(A) \geq \delta_0, \text{ and every } k \geq 2,\\ \exists d \in D: \quad \mu\left(\bigcap_{j=0}^{k-1} T^{-jd} A\right) > 0.\end{gathered} \tag{H_{δ_0}}$$

The desired conclusion is $\mathrm{MR}_D(X, \mu, T)$ for every measurably distal system.

Here I use the measure-theoretic notion of distality. One characterization is that measurably distal systems form the smallest class containing the trivial system and closed under factors, compact group extensions, and inverse limits. This formulation also accounts for compact homogeneous-space extensions, through passage to factors. arXiv

Taking your nilsystem theorem as established, we can prove:

> **Theorem 1.** If D satisfies H_{δ_0}, then D is a set of multiple measurable recurrence for every measure-preserving inverse limit of nilsystems. There is no requirement of a uniform bound on the nilpotency steps.

The important point in proving this is to use **conditional densities close to one**, rather than subtracting an approximation error from a possibly very small recurrence intersection.

2. A lifting lemma that proves the inverse-limit step

Lemma 2: Near-full conditional density lifts recurrence

Let $\mathcal{F} \subseteq \mathcal{B}$ be a T-invariant factor, and suppose that D is a set of k-term measurable recurrence on this factor.

Let $A \in \mathcal{B}$, and put

$$f = \mathbb{E}(1_A \mid \mathcal{F}).$$

Suppose that, for some $\varepsilon < 1/k$,

$$B := \{f > 1 - \varepsilon\}$$

has positive measure. Then

$$\mu\left(\bigcap_{j=0}^{k-1} T^{-jd} A\right) > 0$$

for some $d \in D$.

Proof. Since B belongs to the factor, choose $d \in D$ such that

$$C := \bigcap_{j=0}^{k-1} T^{-jd} B$$

has positive measure.

The elementary union bound gives

$$1_{\bigcap_{j=0}^{k-1} T^{-jd} A} \geq 1 - \sum_{j=0}^{k-1} 1_{T^{-jd} A^c}.$$

Taking conditional expectations and using invariance of the factor,

$$\mathbb{E}\left(1_{\bigcap_{j=0}^{k-1} T^{-jd} A} \mid \mathcal{F}\right) \geq 1 - \sum_{j=0}^{k-1} \left(1 - f \circ T^{jd}\right).$$

On C, every $f \circ T^{jd}$ is greater than $1 - \varepsilon$. Consequently,

$$\begin{aligned} \mu\left(\bigcap_{j=0}^{k-1} T^{-jd} A\right) &\geq \int_C \mathbb{E}\left(1_{\bigcap_{j=0}^{k-1} T^{-jd} A} \mid \mathcal{F}\right) d\mu \\ &\geq (1 - k\varepsilon)\mu(C) > 0. \end{aligned}$$

□

Notice that this lemma makes **no compactness assumption** on the extension.

Proposition 3: Recurrence along a fixed D passes to inverse limits

Suppose

$$\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots, \qquad \bigvee_m \mathcal{F}_m = \mathcal{B} \quad \text{modulo null sets},$$

and D is a set of multiple measurable recurrence on every factor $\mathcal{F}_m$. Then it is a set of multiple measurable recurrence on X.

Proof. Fix A of positive measure and $k \geq 2$. Choose $\varepsilon < 1/k$. By martingale convergence,

$$\mathbb{E}(1_A \mid \mathcal{F}_m) \longrightarrow 1_A \quad \text{almost everywhere.}$$

Therefore, for some m,

$$\mu\{\mathbb{E}(1_A \mid \mathcal{F}_m) > 1 - \varepsilon\} > 0.$$

Apply Lemma 2. □

Theorem 1 now follows immediately from your nilsystem theorem.

There is also a useful restriction on any possible counterexample.

Corollary 4: A conditional-density obstruction

Suppose D is k-term recurrent on a factor $\mathcal{F}$, but

$$\mu\left(\bigcap_{j=0}^{k-1} T^{-jd} A\right) = 0 \qquad (d \in D).$$

Then

$$\boxed{\mathbb{E}(1_A \mid \mathcal{F}) \leq 1 - \frac{1}{k} \quad \text{almost everywhere.}} \tag{1}$$

Otherwise, for some $\varepsilon < 1/k$, the set in Lemma 2 would have positive measure.

Thus, under your hypothesis, a distal avoidance witness must satisfy (1) over **every pro-nil factor**. It cannot be nearly full on a positive-measure family of fibers over such a factor.

3. The remaining step really is a compact extension

For fixed D, the class

$$\mathscr{G}_D = \{(X, \mu, T) : \mathrm{MR}_D(X, \mu, T)\}$$

is closed under factors, and Proposition 3 proves closure under inverse limits.

Therefore the following would suffice to prove your distal theorem:

> **Compact-extension assertion needed.** Suppose D satisfies H_{δ_0}. If Y is measurably distal and belongs to $\mathscr{G}_D$, and $X \to Y$ is a compact group extension, then $X \in \mathscr{G}_D$.

I do not yet have a proof of this assertion.

It is important not to substitute the standard compact-extension theorem from the proof of Szemerédi's theorem. That theorem preserves positivity of the lower Cesàro average of the multiple correlations over **all integers**. For example, this is exactly the meaning of "MR" in Eisner's Definition 2.1 and Proposition 7.4. It is not a theorem about retaining an arbitrary prescribed set of times D. arXiv

The issue is visible in a compact group skew product

$$T(y,z) = (Sy, \rho(y)z).$$

Writing

$$\rho_n(y) = \rho(S^{n-1}y)\cdots\rho(y),$$

we have

$$T^{jn}(y,z) = \big(S^{jn}y, \rho_{jn}(y)z\big).$$

Recurrence of

$$y, S^n y, \dots, S^{(k-1)n}y$$

in the base does not control

$$\rho_n(y), \rho_{2n}(y), \dots, \rho_{(k-1)n}(y).$$

Even for a continuous cocycle over a compact base, continuity of ρ does not give a modulus of continuity for all the accumulated cocycles ρ_n that is uniform in n.

The usual compactness-and-coloring approach can also extract a recurrence time of the form qn. That is precisely where our earlier multiplier problem reappears: $n \in D$ need not imply $qn \in D$. The identity $c(D/q) = c(D)$ does not, by itself, allow a multiplier selected after the orbit segment has been examined.

4. An exact high-density formulation of what must be proved

There is a useful way to isolate the missing theorem without committing to a particular tower argument.

For a distal system, A of positive measure, and $k \geq 2$, define

$$R(A,k) = \left\{ n \geq 1 : \mu\left(\bigcap_{j=0}^{k-1} T^{-jn} A \right) > 0 \right\},$$

and put

$$E(A,k) = \mathbb{N} \setminus R(A,k).$$

Recall our extremal parameter

$$c(E) = \sup_{\ell \geq 2} a_\ell(E).$$

Then your proposed distal theorem is equivalent to

$$\boxed{c(E(A,k)) = 1 \quad \text{for every distal system, every positive-measure } A, \text{ and every } k \geq 2.} \tag{2}$$

Proof of the equivalence. Suppose (2) holds and D satisfies H_{δ_0}. If D failed the particular distal recurrence test, then

$$D \subseteq E(A,k).$$

Monotonicity would give

$$c(D) \geq c(E(A,k)) = 1,$$

contradicting $c(D) \le \delta_0 < 1$.

Conversely, suppose some distal test has

$$c(E(A,k)) < 1.$$

Choose

$$c(E(A,k)) < \delta_0 < 1.$$

Then $a_\ell(E(A,k)) < \delta_0$ for every ℓ, so $E(A,k)$ satisfies H_{δ_0}. But it fails the original distal test by definition. □

In concrete terms, (2) asks for the following construction:

> Given a failed distal recurrence test and any $\eta < 1$, construct an auxiliary probability-preserving system and a set B with $\nu(B) > \eta$ such that, for some L,
>
> $$\nu\left(\bigcap_{j=0}^{L-1} U^{-jn} B\right) > 0 \quad \Longrightarrow \quad n \in R(A,k). \tag{3}$$

The auxiliary system in (3) need not be distal. This freedom is exactly where the **all-systems** hypothesis on D could be used.

Your nilsystem argument supplies high-density tests of this kind for nilpotent return conditions. The desired strengthening is an analogous high-density test for arbitrary distal recurrence conditions.

5. Alweiss's obstruction already has a two-step distal realization

Here is the concrete connection to your other question about Alweiss's example.

Use the version of the construction in Ben Green's exposition. There are positive numbers (α_i) with

$$\sum_i \alpha_i < \infty,$$

chosen sufficiently rapidly decreasing, and a set S_{A} consisting of the integers n for which some index i satisfies

$$\left\| n\alpha_i - \frac{1}{2} \right\| \le \varepsilon_*, \qquad \|n\alpha_j\| \le \varepsilon_* \quad (j \ne i), \qquad \sum_{j \ne i} \|n\alpha_j\| < 4, \tag{4}$$

where

$$\varepsilon_* = \frac{1}{200}.$$

With the specified choice of parameters, this set intersects every nil-Bohr neighborhood of zero. These are Definition 2.1 and Theorem 1.1 in that exposition. arxiv.org

The following skew-product realization follows directly from these formulas.

Proposition 5

There is a two-step distal probability-preserving system (X, μ, T) and a measurable set A with

$$\mu(A) = \frac{1}{4096}$$

such that

$$\boxed{A \cap T^{-n}A \cap T^{-2n}A = \varnothing \qquad (n \in S_{\mathrm{A}}).} \tag{5}$$

Construction. Write $e(t) = e^{2\pi i t}$, and let

$$K = \mathbb{T}^{\mathbb{N}}, \qquad F = \mathbb{C}/(8\mathbb{Z}[i]).$$

Equip $X = K \times F$ with product Haar measure.

Define the continuous function

$$\phi(x) = \sum_{i=1}^{\infty} \big(e(\alpha_i) - 1\big) e(x_i). \tag{6}$$

The series converges uniformly, since

$$|e(\alpha_i) - 1| \le 2\pi\alpha_i$$

and $\sum_i \alpha_i < \infty$.

Set

$$T(x, z) = \big(x + \alpha,\, z + \phi(x) \bmod 8\mathbb{Z}[i]\big). \tag{7}$$

This preserves product Haar measure.

It is an isometric compact-group extension of the rotation $x \mapsto x + \alpha$. More directly, two points with different base coordinates remain separated in the base under every iterate; two points with the same base coordinate keep the same fiber difference. Thus the system is topologically distal and has a distal tower of height at most two. Its Haar measure makes it measurably distal as well.

For each fixed n,

$$T^n(x, z) = \big(x + n\alpha,\, z + \Phi_n(x)\big),$$

where telescoping gives

$$\Phi_n(x) = \sum_{i} \big(e(n\alpha_i) - 1\big) e(x_i). \tag{8}$$

Again, the series is uniformly convergent for each fixed n.

The second difference of the fiber coordinate is therefore

$$\begin{aligned} Z_n(x) &:= \Phi_{2n}(x) - 2\Phi_n(x) \\ &= \sum_{i} \big(e(n\alpha_i) - 1\big)^2 e(x_i). \end{aligned} \tag{9}$$

Fix $n \in S_{\mathrm{A}}$, and choose i as in (4). Its contribution to (9) has magnitude

$$|e(n\alpha_i) - 1|^2 = 4\sin^2(\pi n\alpha_i) \ge 4\cos^2(\pi\varepsilon_*) \ge 2.$$

The total magnitude of the remaining contributions is at most

$$\begin{aligned}\sum_{j\neq i}|e(n\alpha_j)-1|^2 &\le 4\pi^2\sum_{j\neq i}\|n\alpha_j\|^2\\ &\le 4\pi^2\varepsilon_*\sum_{j\neq i}\|n\alpha_j\|\\ &< 16\pi^2\varepsilon_* < 1.\end{aligned}$$

Since the distinguished contribution is also at most 4, we obtain, uniformly in x,

$$1 \le |Z_n(x)| \le 5. \tag{10}$$

Every nonzero point of $8\mathbb{Z}[i]$ has magnitude at least 8. Hence

$$\operatorname{dist}\big(Z_n(x), 8\mathbb{Z}[i]\big) \ge 1. \tag{11}$$

Now let $C \subset F$ be the image of the square

$$[0,1/8)+i[0,1/8), \qquad A = K \times C.$$

Its normalized Haar measure is

$$\mu(A) = \frac{(1/8)^2}{8^2} = \frac{1}{4096}.$$

If $(x,z), T^n(x,z), T^{2n}(x,z)$ all belonged to A, choose their fiber representatives z_0, z_1, z_2 in that square. Then

$$Z_n(x) \equiv z_0 - 2z_1 + z_2 \pmod{8\mathbb{Z}[i]}.$$

But

$$|z_0 - 2z_1 + z_2| < 4\frac{\sqrt{2}}{8} < 1,$$

contradicting (11). This proves (5). □

This is a **measurable** obstruction, not merely a failure of recurrence for an individual point.

What this does—and does not—show

It does **not** refute your proposed theorem: we have not established that S_{A} satisfies H_{δ_0} for any $\delta_0 < 1$.

Instead, it gives the implication

$$\boxed{\begin{array}{c}\text{Your theorem for two-step distal systems}\\ \Longrightarrow\\ S_{\mathrm{A}} \text{ is not } \delta\text{-density regular for any } \delta < 1.\end{array}} \tag{12}$$

Indeed, any fixed-density universal regularity of S_{A} would, by that theorem, force recurrence for the set A above, contradicting (5).

So this proposed distal strengthening is directly connected to the Alweiss question you have been investigating. It would already settle that question without requiring the full zero–one conjecture for arbitrary systems.

6. Where the proof attempt stands

The inverse-limit argument is complete and preserves the set of times D exactly. Consequently, your nilsystem result extends to all pro-nil systems, including inverse limits with unbounded nilpotency step.

For arbitrary distal systems, the unresolved issue is the high-density test construction (3), or a compact-extension lemma that implies it. The standard compact-extension recurrence theorem does not supply that lemma, because it concerns averages over unrestricted times rather than the same prescribed D. arXiv

The explicit skew product above also shows why the nil-Bohr consequence of your hypothesis cannot be the only input: a nil-Bohr recurrent set can already fail three-term measurable recurrence in a two-step distal system. Any successful proof must use the stronger hypothesis that D is fixed-density regular **for all probability-preserving systems**.

Thus I have not proved the full distal assertion, but the reduction is now precise: the limit stages work, the compact successor stage remains, and even its two-step case would resolve the fixed-density status of Alweiss's example.

---

21. User

2026-09-15 17:42:58 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T09:42:58.396000Z

Can we maybe use the characterization of distal systems using those cesaro average norms that resemble the host-kra norms, the norms we discussed in the other chat? It would be really nice to have this distal systems intermediary result.

---

22. Assistant

2026-09-15 17:43:16 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T09:43:16.868709Z

I'll pin down the Cesàro/seminorm characterization of distal systems and test whether it can prove the missing compact-extension step while preserving recurrence times in D.

---

23. Assistant

2026-09-15 17:49:32 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T09:49:32.764479Z

I found Assani–Presser's precise seminorm characterization of successive maximal isometric factors. I'm now testing a relative self-joining argument that retains phase information the norms alone may lose.

24. Assistant

2026-09-15 17:58:37 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T09:58:37.693863Z

The self-joining preserves every forbidden time in D, but lowers density; moreover, maximal conditional seminorm can coexist with zero recurrence correlation, confirming that cocycle phase information is essential.

25. Assistant

2026-09-15 18:14:36 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-15T10:14:36.944999Z

These seminorms give a natural way to formulate the missing compact-extension step. In particular, they suggest a precise *relative degree-lowering statement* which, if proved from your fixed-density hypothesis, would establish the distal intermediary theorem.

I have not proved that statement. What I did obtain is a quantitative lifting lemma using the finite-dimensional data underlying the seminorms. It avoids the approximation problem from our previous discussion, where an uncontrolled approximation error could exceed the measure of a recurrence intersection.

There is an important distinction: **the scalar seminorm detects relative compactness, but selecting a time in D requires retaining the phases of the conditional correlations.** Below I make both points precise.

1. The relevant distal seminorms

Let (X, μ, T) be an ergodic probability-preserving system, and let $\mathcal{Y}$ be an invariant factor. In the complex-valued convention, put

$$\mathsf{N}_{\mathcal{Y}}(f)^4 := \lim_{H\to\infty} \frac{1}{H} \sum_{h=1}^{H} \left\| \mathbb{E}\left(f\, \overline{T^h f} \mid \mathcal{Y} \right) \right\|_2^2. \tag{1}$$

Here $T^h f = f \circ T^h$.

If $\mathcal{K}(\mathcal{Y})$ denotes the maximal compact extension of $\mathcal{Y}$ inside X, the characterization is

$$\boxed{\mathsf{N}_{\mathcal{Y}}(f) = 0 \quad \Longleftrightarrow \quad \mathbb{E}(f \mid \mathcal{K}(\mathcal{Y})) = 0.} \tag{2}$$

Thus, starting with the trivial factor $\mathcal{A}_0$ and defining

$$\mathcal{A}_{s+1} = \mathcal{K}(\mathcal{A}_s),$$

these seminorms detect the successive stages of the maximal distal tower. A precise reference is Assani–Presser, Definition 4 and Lemma 1; they write the defining formula in its real-valued form. arXiv

The resemblance to the Host–Kra recursion is real, but the recursion is different. The Host–Kra seminorms iterate lower-order uniformity seminorms of

$$f\, \overline{T^h f},$$

whereas (1) takes the **full L^2-norm of its conditional expectation onto the preceding distal factor**. The resulting maximal-isometric factors need not be the Host–Kra nilfactors. arXiv Annals of Mathematics

For our problem, the especially useful consequence of (2) is:

> If $X \to Y$ is a compact extension, then
>
> $$\mathsf{N}_{\mathcal{Y}}(f) = 0 \quad \Longrightarrow \quad f = 0. \tag{3}$$

This is exactly the kind of conclusion one could use in an induction up the distal tower.

2. A norm-based statement that would complete the successor step

Write (H_{δ_0}) for your hypothesis that D is δ_0-density regular for **all** systems.

Suppose that $X \to Y$ is a compact extension of a distal system, and that we have already proved full multiple recurrence along D on Y. Assume, toward a contradiction, that $A \subseteq X$ has positive measure and

$$\mu\left(\bigcap_{j=0}^{k-1} T^{-jd} A\right) = 0 \qquad (d \in D). \tag{4}$$

Set

$$f_A := 1_A - \mathbb{E}(1_A \mid \mathcal{Y}).$$

The following implication would suffice:

$$\boxed{(\mathrm{H}_{\delta_0}) \text{ and } (4) \quad \Longrightarrow \quad \mathsf{N}_{\mathcal{Y}}(f_A) = 0.} \tag{5}$$

Here the induction hypothesis of full recurrence on the distal base Y is also available.

Indeed, because the extension is compact, (3) would imply $f_A = 0$. Hence A would be measurable with respect to $\mathcal{Y}$, modulo null sets. Recurrence on Y would then contradict (4).

So your proposed norm strategy has a very concrete target: show that an avoidance witness has no genuinely new compact component above the preceding distal factor.

I have not established (5). It is not a consequence of the characteristic-factor theorem alone: that theorem concerns Cesàro averages, whereas (4) concerns every individual time in a prescribed set D.

Two calculations help clarify what a proof of (5) must do.

3. What the seminorm construction preserves—and what it loses

The relative self-joining preserves the forbidden times

Disintegrate

$$\mu = \int_Y \mu_y \, d\nu(y),$$

and let

$$\lambda = \mu \times_Y \mu = \int_Y \mu_y \otimes \mu_y \, d\nu(y).$$

For $F = f \otimes \overline{f}$, the mean ergodic theorem gives

$$\mathsf{N}_{\mathcal{Y}}(f)^4 = \left\| \mathbb{E}_\lambda\big(F \mid \mathcal{I}(T \times T)\big) \right\|_2^2. \tag{6}$$

Indeed, the inner product

$$\langle F, (T \times T)^h F \rangle_{L^2(\lambda)}$$

is exactly the summand in (1). This is the relative-self-joining identity behind the maximal-isometric characterization. arXiv

Now let

$$C = A \times_Y A.$$

If

$$A_{d,k} := \bigcap_{j=0}^{k-1} T^{-jd} A,$$

then

$$\lambda\left(\bigcap_{j=0}^{k-1} (T \times T)^{-jd} C \right) = \int_Y \mu_y(A_{d,k})^2 \, d\nu(y). \tag{7}$$

Consequently, every forbidden time in (4) remains forbidden in the relative self-joining. **There is no multiplier loss here.**

But its density moves in the wrong direction:

$$\lambda(C) = \int_Y \mu_y(A)^2 \, d\nu(y) \le \mu(A). \tag{8}$$

Thus the basic squaring operation underlying the seminorm is not, on its own, the high-density amplification we need.

Also, simply proving that an indicator has positive seminorm cannot exploit distality. If $a = \mu(A)$, then in every system

$$\begin{aligned} \mathsf{N}_{\mathcal{Y}}(1_A)^4 &\ge \lim_{H\to\infty} \frac{1}{H} \sum_{h=1}^{H} \mu(A \cap T^{-h}A)^2 \\ &\ge \left(\lim_{H\to\infty} \frac{1}{H} \sum_{h=1}^{H} \mu(A \cap T^{-h}A) \right)^2 \\ &\ge a^4. \end{aligned} \tag{9}$$

This is why the centered function f_A, rather than 1_A itself, is the natural object in (5).

Even maximal conditional-correlation norms need not imply phase alignment

Consider a circle extension

$$T(y, z) = (Sy, z + \rho(y)).$$

Write

$$\rho_n(y) = \sum_{i=0}^{n-1} \rho(S^i y).$$

For $f(y,z) = e(z)$, where $e(t) = e^{2\pi i t}$,

$$\mathbb{E}\big(f\,\overline{T^n f} \mid \mathcal{Y}\big)\,(y) = e(-\rho_n(y)).$$

Therefore

$$\boxed{\left\|\mathbb{E}\big(f\,\overline{T^n f} \mid \mathcal{Y}\big)\right\|_2^2 = 1 \qquad \text{for every } n.} \tag{10}$$

Thus the conditional norm is maximal regardless of the cocycle's accumulated phase.

Here is an explicit illustration involving the actual recurrence intersections. Take the skew shift

$$T(x,z) = (x+\alpha,\, z+x), \qquad \alpha \notin \mathbb{Q}.$$

Then

$$T^n(x,z) = \left(x + n\alpha,\ z + nx + \binom{n}{2}\alpha\right).$$

For the three fiber coordinates at times $0, n, 2n$,

$$z_{2n} - 2z_n + z_0 = n^2\alpha \pmod 1. \tag{11}$$

Put

$$A = \mathbb{T} \times [0, 1/16), \qquad E_\alpha = \{n \geq 1 : \|n^2\alpha\|_{\mathbb{T}} \geq 1/4\}.$$

If the three points lay in A, representatives of their fiber coordinates in $[0,1/16)$ would satisfy

$$|z_{2n} - 2z_n + z_0| < 1/8,$$

contradicting (11). Hence

$$A \cap T^{-n}A \cap T^{-2n}A = \varnothing \qquad (n \in E_\alpha), \tag{12}$$

while the conditional norm in (10) remains maximal at every one of these times.

This phenomenon also occurs for the centered indicator. For an interval I of length $\eta < 1/2$, set $g = 1_I - \eta$. Since

$$x \longmapsto nx + \binom{n}{2}\alpha$$

is Haar-uniform for every $n \geq 1$,

$$\|\mathbb{E}(g(z)\,g(T^n(x,z)) \mid x)\|_2^2 = \frac{2\eta^3}{3} - \eta^4 > 0 \qquad (n \geq 1). \tag{13}$$

To check the value, write $c(t) = m(I \cap (I - t))$. Then

$$\int c(t)\,dt = \eta^2, \qquad \int c(t)^2\,dt = \frac{2\eta^3}{3},$$

and integrate $(c(t) - \eta^2)^2$.

This is not a counterexample to your theorem: no fixed-density regularity is asserted for E_α. It shows that a proof must use (H_{δ_0}) in an essential way, beyond a lower bound on these norms—even beyond lower bounds at every time in D.

It also does not rule out using polarized seminorm identities or the full conditional correlations. Those retain substantially more information than the single scalar quantity in (10).

4. Retaining that information gives a quantitative lifting lemma

The finite-dimensional structure underlying a compact extension is useful here. Such extensions can be analyzed using invariant finite-rank $L^\infty(Y)$-modules with conditionally orthonormal bases; these modules provide the needed approximations to bounded functions. This is the relative-discrete-spectrum characterization, stated in Definition 2.4 and Proposition 2.8 of *Distal systems in topological dynamics and ergodic theory*. arXiv

Let

$$\psi = (\psi_1, \dots, \psi_r)^{\mathsf{T}}$$

be a conditionally orthonormal basis:

$$\mathbb{E}(\psi_i \overline{\psi_j} \mid Y) = \delta_{ij}.$$

Invariance supplies unitary matrices $M_n(y)$ satisfying

$$\psi(T^n x) = M_n(y)\psi(x), \qquad y = \pi(x). \tag{14}$$

Write an approximating function in the module as

$$g(x) = v(y)^* \psi(x).$$

Then conditional orthonormality gives the exact formula

$$\boxed{\int |T^n g - g|^2 \, d\mu_y = \|M_n(y)^* v(S^n y) - v(y)\|_{\mathbb{C}^r}^2 .} \tag{15}$$

Unlike the absolute-square conditional correlation in (10), this expression retains the discrepancy in phase.

Proposition: Phase-sensitive lifting

Let $A \subseteq X$, $k \geq 2$, and let g have the preceding representation. Suppose that $E \subseteq Y$ has positive measure and, for almost every $y \in E$,

$$\mu_y(A) \geq b > 0, \qquad \int |1_A - g|^2 \, d\mu_y \leq \varepsilon^2. \tag{16}$$

Suppose there exist $d \in D$ and a positive-measure set

$$B \subseteq \bigcap_{j=0}^{k-1} S^{-jd} E$$

such that, for almost every $y \in B$,

$$\sum_{j=1}^{k-1} \left\| M_{jd}(y)^* v(S^{jd} y) - v(y) \right\|^2 \leq \theta. \tag{17}$$

Then

$$\boxed{\mu\left(\bigcap_{j=0}^{k-1} T^{-jd}A\right) \geq \big[b - 6(k-1)\varepsilon^2 - 3\theta\big]\nu(B).} \tag{18}$$

In particular, the desired recurrence follows whenever

$$6(k-1)\varepsilon^2 + 3\theta < b. \tag{19}$$

Proof

Put $f = 1_A$. For every n,

$$|f - T^n f|^2 \leq 3|f-g|^2 + 3|g - T^n g|^2 + 3|T^n(g-f)|^2. \tag{20}$$

The conditional measures satisfy

$$(T^n)_*\mu_y = \mu_{S^n y}$$

on a common conull set. Therefore, for $y \in B$ and $1 \leq j < k$, the first and third terms in the conditional integral of (20) are each at most $3\varepsilon^2$. By (15),

$$\int |f - T^{jd}f|^2 \, d\mu_y \leq 6\varepsilon^2 + 3\left\|M_{jd}(y)^* v(S^{jd}y) - v(y)\right\|^2.$$

Summing and using (17),

$$\sum_{j=1}^{k-1} \int |f - T^{jd}f|^2 \, d\mu_y \leq 6(k-1)\varepsilon^2 + 3\theta. \tag{21}$$

Because f is an indicator,

$$1_{\bigcap_{j=0}^{k-1} T^{-jd}A} \geq 1_A - \sum_{j=1}^{k-1} |f - T^{jd}f|^2. \tag{22}$$

Indeed, if $x \in A$ but is missing from the intersection, at least one term in the sum equals 1.

Integrating (22) conditionally over $y \in B$, and applying (16) and (21), proves (18). □

Why this addresses the earlier approximation problem

A global estimate such as

$$\mu\left(\bigcap T^{-jd}A\right) \geq \text{positive quantity} - C\|1_A - g\|_2$$

is dangerous: the positive quantity may be arbitrarily small, depending on d.

In (18), the main term and the error are both proportional to $\nu(B)$. **There is no need to know a lower bound for the measure of the base recurrence intersection.**

Furthermore, compactness does supply the initial approximation in (16). Choose $b > 0$ for which

$$E_0 = \{y : \mu_y(A) \geq b\}$$

has positive measure. Approximate 1_A sufficiently well in L^2 by a finite-module function g; then discard from E_0 the fibers on which the conditional approximation error exceeds ε^2. With the approximation chosen sufficiently accurately, a positive-measure set E remains.

Thus the unsolved part is specifically **the selection of $d \in D$ and B satisfying (17)**, not the approximation or the final positivity argument.

5. The remaining phase-selection problem

Already for a circle extension, a sufficient selection statement would be:

> Assume (H_{δ_0}), and assume full multiple recurrence along D has been established on the distal base (Y, ν, S). For every circle cocycle ρ, every positive-measure $E \subseteq Y$, every k, and every $\epsilon > 0$, there is $d \in D$ such that
>
> $$\nu\left(\bigcap_{j=0}^{k-1} S^{-jd}E \cap \bigcap_{j=1}^{k-1}\{y : \|\rho_{jd}(y)\|_{\mathbb{T}} < \epsilon\}\right) > 0. \tag{23}$$

For general finite-rank modules, the analogue retains the unitary matrices in (17). These are sufficient phase-sensitive targets; I am **not** claiming that the Cesàro-norm characterization proves them.

A norm-based proof could instead establish the annihilation implication (5). That would bypass explicit phase selection and would be an especially clean argument. But the usual relative-self-joining calculation does not yet derive (5) from the all-systems fixed-density hypothesis.

Finally, a proof for all distal systems must permit successive compact extensions beyond finite height. Distal systems of arbitrary countable height exist; it is not enough to prove something for each finite $\mathcal{A}_s$ and silently identify their join with every distal system. The inverse-limit lifting we already proved handles limit stages, so a compact-extension argument valid over an arbitrary distal base would address this issue. arXiv

My present conclusion is that your suggestion gives a useful analytic formulation, but not yet the distal theorem. The clearest sufficient norm statement is (5). The phase-sensitive alternative is (17), and Proposition (18) proves that this alternative really does lift recurrence without changing D or losing control through a small base intersection. What remains is to extract one of these conclusions from fixed-density regularity for all systems.

---

26. User

2026-09-16 14:27:38 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T06:27:38.438000Z

good progress. Please keep working towards this intermediary theorem.

---

27. Assistant

2026-09-16 14:28:01 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T06:28:01.509029Z

I'll focus on the compact-extension step, testing whether the all-systems density hypothesis can force the cocycle phase alignment that scalar seminorms do not detect.

28. Assistant

2026-09-16 14:31:06 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T06:31:06.907018Z

The relative self-joining preserves forbidden times but lowers density, so I'm investigating invariant joinings nearer the diagonal as a possible amplification mechanism.

29. Assistant

2026-09-16 14:40:12 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T06:40:12.697614Z

I'm also reducing the compact-extension problem to unitary matrix cocycles, which handles homogeneous fibers K/H without requiring the fiber itself to be a group.

30. Assistant

2026-09-16 15:01:47 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T07:01:47.870816Z

I have pushed the compact-extension reduction further. **Any failure of recurrence at a compact stage can be replaced by a failure for a fixed spherical cap in a finite-dimensional sphere extension of the same base.** The set D and the forbidden progression length remain unchanged.

I also obtained a quantitative version of the seminorm obstruction, and an explicit finite weighted consequence of your all-systems hypothesis. Together, these give a more concrete target than the earlier qualitative statement that an avoidance witness must have a nonzero relative seminorm.

I have not yet proved the distal intermediary theorem. The results below isolate a finite-dimensional phase-selection problem that is sufficient to finish the compact-extension step, and show exactly how an approximate high-density construction could finish it.

1. A failure in any compact extension produces a spherical-cap failure

Write your hypothesis as

$$(\mathrm{H}_{\delta_0}): \quad \forall (Z,\lambda,R),\ \forall B \text{ with } \lambda(B)\ge\delta_0,\ \forall L\ge 2,\ \exists d\in D: \quad \lambda\left(\bigcap_{j=0}^{L-1} R^{-jd}B\right)>0.$$

For the structural reduction that follows, we do not yet need this hypothesis.

Theorem 1: Spherical-cap reduction

Let

$$\pi:(X,\mu,T)\longrightarrow(Y,\nu,S)$$

be a compact extension of standard ergodic probability-preserving systems. Suppose $A\subseteq X$, $\mu(A)>0$, and $k\ge 2$ satisfy

$$\mu\left(\bigcap_{j=0}^{k-1} T^{-jd}A\right)=0 \qquad (d\in D). \tag{1}$$

Then there exist a finite r, a measurable cocycle

$$U: Y \longrightarrow \mathrm{U}(r),$$

a positive-measure set $E\subseteq Y$, a unit vector $u\in\mathbb{C}^r$, and $\eta>0$, such that

$$\boxed{\nu\left(\bigcap_{j=0}^{k-1} S^{-jd}E \cap \left\{y:\max_{1\le j<k}\|U_{jd}(y)u-u\|<\eta\right\}\right)=0 \quad (d\in D).} \tag{2}$$

Here

$$U_n(y)=U(S^{n-1}y)\cdots U(y).$$

Consequently, on the sphere extension

$$R(y,z)=(Sy,U(y)z), \qquad (Y\times\mathbb{S}^{2r-1},\nu\times\sigma),$$

where σ is normalized surface measure, the set

$$B=E\times\{z:\|z-u\|<\eta/4\} \tag{3}$$

has positive measure and satisfies

$$\boxed{(\nu\times\sigma)\left(\bigcap_{j=0}^{k-1} R^{-jd}B\right)=0 \qquad (d\in D).} \tag{4}$$

Thus, for the compact-extension step, **arbitrary measurable sets in arbitrary compact homogeneous fibers can be replaced by ordinary caps in finite-dimensional spheres**. The sphere extension is an auxiliary extension of Y; it need not be a factor of the original X.

The structural input is the finite-rank invariant-module characterization of compact extensions. It applies to isometric extensions with homogeneous fibers K/H, not only to compact-group fibers. Edeko–Kreidler give the relative-discrete-spectrum formulation and its conditionally orthonormal module version in Definition 2.4 and Proposition 2.8. arXiv

Proof

Put $f=1_A$, and disintegrate

$$\mu=\int_Y \mu_y\,d\nu(y).$$

Choose $b>0$ such that

$$E_0=\{y:\mu_y(A)\ge b\}$$

has positive measure. Set

$$\eta=\sqrt{\frac{b}{100k}}. \tag{5}$$

By compactness of the extension, approximate f in L^2 by a function in a finite-rank invariant module. Write this approximation as

$$g(x) = v(y)^* \psi(x), \qquad \psi = (\psi_1, \dots, \psi_r)^{\mathsf{T}},$$

where

$$\mathbb{E}(\psi_i \overline{\psi_j} \mid Y) = \delta_{ij}. \tag{6}$$

Choose the approximation sufficiently accurately that a positive-measure subset $E_1 \subseteq E_0$ satisfies

$$\int |f - g|^2 \, d\mu_y \le \eta^2 \qquad (y \in E_1). \tag{7}$$

Conditional orthonormality gives

$$\|v(y)\| = \|g\|_{L^2(\mu_y)}.$$

Hence, on E_1,

$$\sqrt{b} - \eta \le \|v(y)\| \le 1 + \eta. \tag{8}$$

In particular, these coefficient vectors stay away from zero and lie in a bounded subset of $\mathbb{C}^r$.

After passing to a positive-measure subset $E \subseteq E_1$, there is a nonzero vector v_0, with $\|v_0\| < 2$, such that

$$\|v(y) - v_0\| < \eta \qquad (y \in E). \tag{9}$$

Set

$$u = \frac{v_0}{\|v_0\|}.$$

Invariance of the module gives a unitary cocycle satisfying

$$\psi(T^n x) = U_n(y) \psi(x). \tag{10}$$

Consequently,

$$\int |T^n g - g|^2 \, d\mu_y = \left\| U_n(y)^* v(S^n y) - v(y) \right\|^2. \tag{11}$$

Fix $d \in D$. Suppose that y belongs to the set inside (2). Then all the points

$$y, S^d y, \dots, S^{(k-1)d} y$$

belong to E, and

$$\|U_{jd}(y) u - u\| < \eta \qquad (1 \le j < k).$$

Since U_{jd} is unitary,

$$\|U_{jd}^* u - u\| = \|U_{jd} u - u\|.$$

Using (9), we get

$$\begin{aligned}\left\|U_{jd}(y)^* v(S^{jd}y) - v(y)\right\| &\le \|v(S^{jd}y) - v_0\| \\ &\quad + \|U_{jd}(y)^* v_0 - v_0\| + \|v_0 - v(y)\| \\ &< 2\eta + \|v_0\|\eta \\ &< 4\eta.\end{aligned} \tag{12}$$

The conditional approximation bounds at y and $S^{jd}y$, together with (11), now give

$$\|f - T^{jd} f\|_{L^2(\mu_y)} < \eta + 4\eta + \eta = 6\eta. \tag{13}$$

For indicators, we have the pointwise inequality

$$1_{\bigcap_{j=0}^{k-1} T^{-jd}A} \ge 1_A - \sum_{j=1}^{k-1} |f - T^{jd} f|^2. \tag{14}$$

Therefore, on the set inside (2),

$$\begin{aligned}\mu_y\left(\bigcap_{j=0}^{k-1} T^{-jd}A\right) &\ge b - 36(k-1)\eta^2 \\ &> b - \frac{36b}{100} \\ &> \frac{b}{2}.\end{aligned} \tag{15}$$

If that base set had positive measure, integrating (15) would contradict (1). This proves (2).

Finally, if (y, z) belongs to the intersection in (4), then every $S^{jd}y$ belongs to E, while

$$\|z - u\| < \eta/4, \qquad \|U_{jd}(y)z - u\| < \eta/4.$$

Thus

$$\|U_{jd}(y)u - u\| \le \|U_{jd}(y)(u - z)\| + \|U_{jd}(y)z - u\| < \eta/2.$$

The projection of the intersection in (4) is therefore contained in the null set from (2). This proves (4). □

What this removes from the problem

The earlier phase-sensitive lifting lemma involved varying coefficient vectors

$$U_{jd}(y)^* v(S^{jd}y) - v(y).$$

The theorem above freezes the coefficient vector. We only need to make

$$U_d(y)u,\ U_{2d}(y)u, \dots, U_{(k-1)d}(y)u$$

return close to **one fixed vector** u on a positive-measure base recurrence set.

In particular, it is not necessary to prove

$$\|U_{jd}(y) - I\|_{\mathrm{op}} \ll 1.$$

That stronger demand would unnecessarily require control of the cocycle on every vector.

2. The seminorm obstruction has a quantitative finite-rank form

Recall the relative seminorm

$$\mathsf{N}_Y(h)^4 = \lim_{H\to\infty} \frac{1}{H}\sum_{n=1}^{H} \left\| \mathbb{E}(h\,\overline{T^n h} \mid Y)\right\|_2^2. \tag{16}$$

This is the seminorm detecting the maximal compact extension of Y; the relative-self-joining description appears in Assani–Presser's Definition 4 and Lemma 1. arXiv

Here is a useful explicit estimate.

Proposition 2: Rank-sensitive seminorm bound

Let $\mathcal{M}$ be a T-invariant relative module of rank r, with a conditionally orthonormal basis, and let $P_{\mathcal{M}}$ denote the orthogonal projection onto its L^2-closure. Then

$$\boxed{\mathsf{N}_Y(h)^4 \geq \frac{\|P_{\mathcal{M}}h\|_2^4}{r}.} \tag{17}$$

Proof

On the relatively independent self-joining

$$\lambda = \mu \times_Y \mu,$$

define

$$F_h(x,x') = h(x)\overline{h(x')}$$

and

$$K_{\mathcal{M}}(x,x') = \sum_{i=1}^{r} \psi_i(x)\overline{\psi_i(x')}. \tag{18}$$

The kernel $K_{\mathcal{M}}$ is invariant under $T\times T$: both vectors of basis functions are multiplied by the same unitary matrix over the common base point.

Conditional orthonormality also gives

$$\|K_{\mathcal{M}}\|_{L^2(\lambda)}^2 = r \tag{19}$$

and

$$\begin{aligned} \langle F_h, K_{\mathcal{M}}\rangle &= \sum_{i=1}^{r}\int_Y \left|\mathbb{E}(h\overline{\psi_i}\mid Y)\right|^2 \, d\nu \\ &= \|P_{\mathcal{M}}h\|_2^2. \end{aligned} \tag{20}$$

Let P_{inv} be projection onto the invariant functions of the relative self-joining. The mean ergodic identity is

$$\mathsf{N}_Y(h)^4 = \|P_{\mathrm{inv}}F_h\|_2^2. \tag{21}$$

Since $K_{\mathcal{M}}$ is invariant, Cauchy–Schwarz gives

$$\|P_{\mathcal{M}}h\|_2^4 = |\langle P_{\mathrm{inv}}F_h, K_{\mathcal{M}}\rangle|^2 \leq r\,\mathsf{N}_Y(h)^4.$$

This is (17). □

Application to a hypothetical failed compact-extension step

Suppose recurrence along D has already been established on Y, but A satisfies (1). Put

$$a = \mu(A), \qquad p = \mathbb{E}(1_A \mid Y), \qquad h = 1_A - p.$$

Our conditional-density lifting argument shows that

$$p \le 1 - \frac{1}{k} \quad \text{almost everywhere.} \tag{22}$$

Indeed, a positive-measure set of fibers on which $p > 1 - 1/k$, with a small fixed margin, would allow base recurrence and the conditional union bound to lift the intersection to A.

It follows that

$$\begin{aligned} \|h\|_2^2 &= \int_Y p(1-p)\, d\nu \\ &\ge \frac{1}{k} \int_Y p \, d\nu \\ &= \frac{a}{k}. \end{aligned} \tag{23}$$

Because the extension is compact, choose a finite-rank invariant module $\mathcal{M}$ so that

$$\|P_{\mathcal{M}} h\|_2^2 \ge \frac{a}{2k}.$$

Then Proposition 2 yields

$$\boxed{\mathsf{N}_Y(1_A - \mathbb{E}(1_A \mid Y))^4 \ge \frac{a^2}{4k^2 r}.} \tag{24}$$

Thus the obstruction is not merely "some nonzero relative compact component." A hypothetical failure supplies a **particular finite rank** r and a definite amount of relative seminorm.

This sharpens the norm strategy: one would only need to contradict the bound in (24) for the finite module extracted from the witness. A bound uniform over every possible rank is not required.

What remains unproved is an upper bound coming from (H_{δ_0}) that contradicts (24).

3. The all-systems hypothesis gives an explicit finite averaging certificate

There is a useful strengthening of the way we have been using (H_{δ_0}). For every fixed progression length, it supplies one finite weighted average that works uniformly over **all systems and all admissible functions**.

Proposition 3: Universal finite weighted recurrence

Assume (H_{δ_0}). For every $L \ge 2$, there are a finite nonempty $F \subseteq D$, positive weights w_d with

$$\sum_{d \in F} w_d = 1,$$

and $\epsilon > 0$, such that

$$\boxed{\sum_{d \in F} w_d \int_Z \prod_{j=0}^{L-1} g(R^{jd} z)\, d\lambda(z) \geq \epsilon} \tag{25}$$

for every probability-preserving system and every measurable $g : Z \to [0,1]$ satisfying

$$\int g\, d\lambda \geq \delta_0.$$

There is an elementary formula for the weights.

Proof

Let

$$r_L(D,N) = \max\{|Q| : Q \subseteq [1,N] \text{ contains no } L\text{-term AP with difference in } D\}.$$

By our earlier finite extremal formulation, choose N such that

$$r := r_L(D,N) < \delta_0 N. \tag{26}$$

Set

$$F = \{d \in D : (L-1)d < N\}, \qquad c_d = N - (L-1)d. \tag{27}$$

For any $Q \subseteq [1,N]$, let $e(Q)$ count the L-term progressions in Q with difference in D. Deleting one vertex from each of its progressions leaves an admissible set after at most $e(Q)$ deletions. Therefore

$$|Q| \leq r + e(Q). \tag{28}$$

More generally, for any $t_1, \dots, t_N \in [0,1]$, choose a random subset Q by including the indices independently with probabilities t_i. Taking expectations in (28) gives

$$\sum_{i=1}^{N} t_i \leq r + \sum_{d \in F} \sum_{i=1}^{N-(L-1)d} \prod_{j=0}^{L-1} t_{i+jd}. \tag{29}$$

Apply this pointwise with

$$t_i = g(R^i z)$$

and integrate. Invariance yields

$$N \int g\, d\lambda \leq r + \sum_{d \in F} c_d \int \prod_{j=0}^{L-1} g \circ R^{jd}\, d\lambda. \tag{30}$$

Consequently, with

$$W = \sum_{d \in F} c_d, \qquad w_d = \frac{c_d}{W},$$

we obtain (25) with

$$\boxed{\epsilon = \frac{\delta_0 N - r}{W} > 0.} \tag{31}$$

The set F is nonempty because (26) implies $r < N$. $\square$

This gives more than a qualitative choice of a good d: **the same finite collection of times, the same weights, and the same positive lower bound work in every auxiliary system.**

It also means we may work with $[0,1]$-valued functions rather than indicators. That is potentially helpful for a construction using continuous functions of matrix coefficients or smooth functions on the sphere.

4. How this changes the remaining construction

Combining the reductions, the compact-extension problem can be concentrated on the following configuration:

$$R(y,z) = (Sy, U(y)z), \qquad B = E \times C, \tag{32}$$

where C is a fixed spherical cap, and

$$(\nu \times \sigma)\left(\bigcap_{j=0}^{k-1} R^{-jd} B\right) = 0 \qquad (d \in D). \tag{33}$$

The base is distal and, in a tower induction, already has the desired recurrence property.

To contradict (H_{δ_0}), it would suffice to construct the following.

> **Finite-dimensional amplification target.**
> From the fixed configuration (32)–(33), find a length L such that, for every finite $F \subseteq D$ and every $\tau > 0$, there are an auxiliary probability-preserving system (Z, λ, V) and a function $g : Z \to [0,1]$ with
>
> $$\int g \, d\lambda \geq \delta_0$$

> and
>
> $$\int \prod_{j=0}^{L-1} g \circ V^{jd} \, d\lambda < \tau \qquad (d \in F). \tag{34}$$

If this construction were available, first fix L, then take F, w_d, ϵ from Proposition 3, and apply (34) with $\tau = \epsilon/2$. The weighted average would be both at least ϵ and less than $\epsilon/2$.

Thus an approximate construction would suffice. We do not need an auxiliary set having exactly zero intersections for all $d \in D$ simultaneously.

The important permissible dependence is

$$L = L(\delta_0, k, U, E, C).$$

It may depend on the specific cocycle, the base set, the cap, and the finite dimension. It need not be uniform over all compact extensions.

But it must be chosen **before** the finite set F.

That order cannot be relaxed. For any finite $F \subseteq \mathbb{N}$ and any density $\gamma < 1$, choose a prime

$$q > \max F, \qquad 1 - \frac{1}{q} > \gamma.$$

In the rotation on $\mathbb{Z}/q\mathbb{Z}$, the complement of one point has measure greater than γ, yet its q-fold intersection at every step $d \in F$ is empty. Therefore, allowing the forbidden length to depend freely on F gives such constructions even when $D = \mathbb{N}$.

This is a useful distinction for the next attempt: approximating the cocycle after the finite certificate is known is legitimate; increasing the progression length at that point is not.

5. What would now finish the distal induction

The structural theorem shows that it is enough to prove the following phase-selection assertion:

$$\boxed{\begin{array}{c}\text{Under } (\mathrm{H}_{\delta_0}),\ \text{for every unitary cocycle over}\\ \text{a distal base already satisfying the conclusion,}\\ \text{every positive-measure } E,\ \text{every unit vector } u,\ \text{and every } \eta > 0,\\ \exists d \in D \text{ for which the base set in } (2) \text{ has positive measure.}\end{array}} \tag{35}$$

Only one vector must return; the entire unitary matrix need not approach the identity. Alternatively, the finite-dimensional amplification target (34) would rule out every failure of (35).

Once the compact successor step is established in that form, the inverse-limit argument from our earlier answer handles the limit stages of the distal tower. Those stages cannot simply be omitted: distal towers of arbitrary countable height occur, as in Glasner–Weiss's construction. arXiv

The current gap is still the conversion of the phase separation in (2) into a high-density test of a fixed length. The scalar Cesàro seminorm does not yet supply that conversion.

What has been established here is a more restrictive form for any obstruction: it can be taken to be finite-dimensional, with a fixed spherical cap, and its relative compact component has the explicit lower bound (24). On the hypothesis side, we now have the explicit finite weighted inequality (25). These are concrete tools for the remaining step, but I do not yet have the argument joining them to prove the intermediary theorem.

---

31. User

2026-09-16 15:08:26 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T07:08:26.786000Z

This seems like good progress! Please keep working.

---

32. Assistant

2026-09-16 15:08:48 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T07:08:48.668631Z

I'll focus on converting the finite-dimensional spherical-cap obstruction into a fixed-length high-density test, using the weighted recurrence certificate to absorb approximation errors.

33. Assistant

2026-09-16 15:11:54 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T07:11:54.591122Z

The key issue is uniformity over finite subsets of D; I'm testing whether compact fibers or relative joinings reduce this to orbit-return control at the original times.

---

34. Assistant

2026-09-16 15:38:59 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T07:38:59.065723Z

I found a constraint: fixed-length arbitrarily accurate tests on every finite subset of D already compactify to exact global avoidance, so we need a genuinely weaker criterion.

---

35. Assistant

2026-09-16 15:54:39 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T07:54:39.429232Z

I obtained a reduction that **uses the fixed-density hypothesis**, rather than only the structure of the compact extension:

> **Any counterexample at a compact-extension stage yields a subset of D, still fixed-density regular at some density below 1, on which one fixed finite-dimensional matrix coefficient has a uniformly negative conditional average.**

This replaces the spherical-cap obstruction by a single algebraic expression with a quantitative negative bias.

I also found an important limitation of the seminorm approach: **two ergodic two-step distal systems can have exactly the same entire collection of distal-tower seminorms, as functionals on the same underlying function space, but different multiple-recurrence behavior along a prescribed set of times.** Thus the proof must retain the time-dependent cocycle information, not just the seminorm values.

I have not yet excluded the negative matrix-coefficient obstruction, so this is not a proof of the intermediary theorem. Here are the precise statements and proofs.

1. Reduction to one negatively biased matrix coefficient

Write (H_δ) for your fixed-density hypothesis:

$$\text{For every probability-preserving system } (Z,\lambda,R),$$
$$\lambda(B)\ge\delta,\quad L\ge 2 \quad\Longrightarrow\quad \exists d\in D:\lambda\left(\bigcap_{j=0}^{L-1}R^{-jd}B\right)>0.$$

At an induction stage, suppose the desired conclusion has already been proved for a distal base (Y,ν,S). More precisely, assume:

$$\begin{gathered}\text{Every set satisfying } (\mathrm{H}_{\varepsilon}) \text{ for some } \varepsilon < 1 \\ \text{is a set of multiple measurable recurrence for } Y.\end{gathered} \tag{P_Y}$$

Start from the spherical-cap reduction established in the preceding answer. A failed compact-extension step supplies a measurable cocycle

$$U : Y \longrightarrow G := \mathrm{U}(r),$$

a positive-measure set $E \subseteq Y$, a unit vector $u \in \mathbb{C}^r$, $k \geq 2$, and $\eta > 0$, such that, writing

$$B_d := \bigcap_{j=0}^{k-1} S^{-jd} E,$$

we have

$$\nu\left(B_d \cap \left\{ y : \max_{1 \leq j < k} \|U_{jd}(y)u - u\| < \eta \right\}\right) = 0 \qquad (d \in D). \tag{1}$$

The finite-dimensional cocycle description is consistent with the invariant finite-rank module characterization of compact extensions. arXiv

Here is the further reduction.

Theorem 1: A single-coefficient obstruction

Suppose D satisfies (H_{δ_0}), the base satisfies (P_Y), and (1) holds.

Then there exist

$$D' \subseteq D, \qquad \delta' \in (0,1), \qquad c > 0,$$

a finite-dimensional nontrivial irreducible unitary representation

$$\pi : G^{k-1} \longrightarrow \mathrm{U}(V),$$

and a unit vector $v \in V$, such that D' satisfies $(\mathrm{H}_{\delta'})$, and for every $d \in D'$,

$$\nu(B_d) > 0$$

and

$$\boxed{\frac{1}{\nu(B_d)} \operatorname{Re} \int_{B_d} \left\langle \pi\big(U_d(y), U_{2d}(y), \ldots, U_{(k-1)d}(y)\big)v, v \right\rangle \, d\nu(y) \leq -c.} \tag{2}$$

The constant c, and a bound for the dimension of V, can be chosen in terms of r, k, η. In particular, the negative bias in (2) does not deteriorate with d or with the measure of B_d.

The last point needs a qualification: this is a **uniform relative bias**. It does not assert a positive lower bound for $\nu(B_d)$.

Proof: Construct a positive-definite phase detector

For an integer $M \geq 1$, define

$$p_M(g) = \left(\frac{1 + \operatorname{Re}\langle gu, u\rangle}{2}\right)^M = \left(1 - \frac{\|gu - u\|^2}{4}\right)^M. \tag{3}$$

Then

$$0 \le p_M \le 1, \qquad p_M(1_G) = 1.$$

Moreover, p_M is a matrix coefficient of a finite-dimensional unitary representation. Indeed,

$$\frac{1 + \mathrm{Re}\langle gu, u\rangle}{2}$$

is the coefficient of $1 \oplus \mathrm{std} \oplus \overline{\mathrm{std}}$ at the unit vector

$$\left(\frac{1}{\sqrt{2}}, \frac{u}{2}, \frac{\bar{u}}{2}\right).$$

Taking the M-th tensor power gives (3).

Let

$$b_M := \int_G p_M(g)\, dm_G(g).$$

Haar-distributed gu is uniformly distributed on $\mathbb{S}^{2r-1}$. A spherical cap of radius $M^{-1/2}$ has measure bounded below by a constant times $M^{-(2r-1)/2}$, and on that cap,

$$p_M(g) \ge \left(1 - \frac{1}{4M}\right)^M \ge \frac{3}{4}.$$

Consequently,

$$b_M \ge c_r M^{-(2r-1)/2} \tag{4}$$

for a constant $c_r > 0$.

Now put

$$P_M(g_1, \dots, g_{k-1}) = \prod_{j=1}^{k-1} p_M(g_j). \tag{5}$$

This is again a finite-dimensional unitary matrix coefficient, now on G^{k-1}, with Haar mean

$$b := \int_{G^{k-1}} P_M \, dm_G^{\otimes(k-1)} = b_M^{k-1}. \tag{6}$$

If some $\|g_j u - u\| \ge \eta$, then

$$P_M(g_1, \dots, g_{k-1}) \le \left(1 - \frac{\eta^2}{4}\right)^M. \tag{7}$$

The right side decreases exponentially, whereas the lower bound for b from (4) decreases only polynomially. We can therefore choose M so that

$$b < \frac{1}{2}, \qquad \left(1 - \frac{\eta^2}{4}\right)^M \le \frac{b}{2}. \tag{8}$$

By (1), for every $d \in D$ with $\nu(B_d) > 0$,

$$\frac{1}{\nu(B_d)} \int_{B_d} P_M\big(U_d(y), \dots, U_{(k-1)d}(y)\big)\, d\nu(y) \le \frac{b}{2}. \tag{9}$$

Decompose the detector into irreducible coefficients

Decompose the finite-dimensional representation defining P_M into its trivial and nontrivial irreducible summands. This gives

$$P_M(g) = b + \sum_{\ell=1}^{s} w_\ell \phi_\ell(g), \tag{10}$$

where

$$w_\ell > 0, \qquad \sum_{\ell=1}^{s} w_\ell = 1 - b,$$

and

$$\phi_\ell(g) = \langle \pi_\ell(g) v_\ell, v_\ell \rangle$$

for a nontrivial irreducible representation π_ℓ and a unit vector v_ℓ.

The constant term is exactly b: Haar averaging a unitary representation is orthogonal projection onto its invariant vectors.

Combining (9) and (10),

$$\sum_{\ell=1}^{s} w_\ell \operatorname{Re}\left[\frac{1}{\nu(B_d)} \int_{B_d} \phi_\ell\big(U_d, \dots, U_{(k-1)d}\big)\, d\nu\right] \le -\frac{b}{2}.$$

Therefore at least one ℓ satisfies

$$\operatorname{Re}\left[\frac{1}{\nu(B_d)} \int_{B_d} \phi_\ell\big(U_d, \dots, U_{(k-1)d}\big)\, d\nu\right] \le -\frac{b}{2(1-b)}. \tag{11}$$

Use fixed-density regularity to select one coefficient

Partition D into

$$D_0 = \{d \in D : \nu(B_d) = 0\}$$

and $D_1, \dots, D_s$, assigning each remaining d to an index for which (11) holds.

We use the finite partition principle from our earlier work:

$$D \text{ satisfies } (\mathrm{H}_{\delta_0}) \quad \Longrightarrow \quad \text{some cell of an } s+1 \text{ partition satisfies } (\mathrm{H}_{\delta_0^{1/(s+1)}}). \tag{12}$$

For completeness, if every cell failed the latter property, take one avoidance witness for each cell and form their product. Its measure is at least δ_0, and at the maximum of the finitely many forbidden lengths it avoids every difference in D, a contradiction.

The good cell in (12) cannot be D_0, because D_0 fails the recurrence test given by E in the base, contradicting (P_Y).

Thus some D_ℓ, $\ell \ge 1$, has the required fixed-density property. Set

$$D' = D_\ell, \qquad \delta' = \delta_0^{1/(s+1)}, \qquad c = \frac{b}{2(1-b)}.$$

Equation (11) is exactly (2). □

This is a narrower obstruction than a general failure of phase selection: one representation, one vector, one negative constant, and a set of times that still satisfies the original kind of hypothesis.

2. An exact identity for the relative seminorm

There is a reason we should preserve the matrix coefficient in (2), rather than immediately replace it by an absolute-square seminorm.

Consider an ergodic compact-group extension

$$T(y,z) = (Sy, \rho(y)z)$$

on

$$(X,\mu) = (Y \times K, \nu \times m_K).$$

For bounded f, define

$$\mathsf{N}_Y(f)^4 = \lim_{H\to\infty} \frac{1}{H} \sum_{n=1}^{H} \left\| \mathbb{E}(f\,\overline{T^n f} \mid Y) \right\|_2^2. \tag{13}$$

This is the relative seminorm detecting the maximal isometric extension, as discussed in Assani–Presser. arXiv

Proposition 2: The full relative seminorm is cocycle-independent

Define the fiber correlation

$$\Gamma_f(t) = \int_{Y\times K} f(y,z)\overline{f(y,zt)}\, d\nu(y)\, dm_K(z).$$

Then

$$\boxed{\mathsf{N}_Y(f)^4 = \int_K |\Gamma_f(t)|^2\, dm_K(t).} \tag{14}$$

In particular, among ergodic compact-group extensions on this same product probability space, the functional $f \mapsto \mathsf{N}_Y(f)$ is independent of the cocycle ρ.

Proof

On the relative self-joining, use coordinates

$$(y,z,z') = (y,z,zt).$$

Because the cocycle acts by left multiplication,

$$(\rho(y)z)^{-1}(\rho(y)z') = z^{-1}z' = t.$$

Thus the relative self-joining is isomorphic to

$$(X\times K, \mu\times m_K, T\times \mathrm{id}).$$

Since T is ergodic, the invariant sigma-algebra of this joining consists exactly of the functions of t.

For

$$F(y,z,z') = f(y,z)\overline{f(y,z')},$$

conditional expectation onto that invariant sigma-algebra is $\Gamma_f(t)$. The relative-self-joining mean ergodic identity gives

$$\mathsf{N}_Y(f)^4 = \|\mathbb{E}(F \mid \mathcal{I}(T\times T))\|_2^2,$$

which is (14). □

For an abelian fiber, the formula becomes particularly transparent. Write

$$f(y,z) = \sum_{\chi\in\widehat{K}} a_\chi(y)\chi(z).$$

Then

$$\Gamma_f(t) = \sum_{\chi\in\widehat{K}} \|a_\chi\|_{L^2(\nu)}^2\, \overline{\chi(t)},$$

and hence

$$\boxed{\mathsf{N}_Y(f)^4 = \sum_{\chi\in\widehat{K}} \|a_\chi\|_{L^2(\nu)}^4.} \tag{15}$$

So this seminorm measures the distribution of L^2-mass among fiber frequencies. The accumulated cocycle phases have disappeared completely.

3. Even the entire distal-tower seminorm collection can agree

The preceding identity leads to a sharper example than the single-character example from our earlier discussion.

Fix an irrational α, and on $\mathbb{T}^2$ consider

$$T_1(x,z) = (x+\alpha, z+x), \qquad T_2(x,z) = (x+\alpha, z+2x). \tag{16}$$

These are affine nilsystems of the type discussed in the recurrence literature. arXiv

Both are ergodic, both have the same Kronecker factor

$$Y = \sigma(x),$$

and both have distal height exactly two.

To verify the spectral assertions, let

$$e_{p,q}(x,z) = e(px+qz).$$

For $m = 1, 2$,

$$T_m e_{p,q} = e(p\alpha)e_{p+mq,q}. \tag{17}$$

When $q \neq 0$, an eigenfunction would have Fourier coefficients of constant magnitude along an infinite progression in the p-index, which is impossible in ℓ^2 unless those coefficients vanish. Thus all eigenfunctions come from $q = 0$, giving precisely the x-factor. Irrationality of α also gives ergodicity.

For

$$\hat{f}(p,q)=\int_{\mathbb{T}^2} f(x,z)e(-px-qz)\,dx\,dz,$$

the distal-tower seminorms are therefore, for both transformations,

$$\boxed{\mathsf{N}_1(f)^4=\sum_{p\in\mathbb{Z}}|\hat{f}(p,0)|^4,} \tag{18}$$

and, by (15),

$$\boxed{\mathsf{N}_2(f)^4=\sum_{q\in\mathbb{Z}}\left(\sum_{p\in\mathbb{Z}}|\hat{f}(p,q)|^2\right)^2.} \tag{19}$$

At every subsequent stage the preceding factor is the whole system, so the mean ergodic theorem gives

$$\boxed{\mathsf{N}_s(f)=\|f\|_2 \qquad (s\ge 3).} \tag{20}$$

Thus **every one of these seminorms agrees for every bounded** f under the identity identification of the two probability spaces.

Nevertheless, their three-term recurrence behavior differs substantially.

An explicit difference in recurrence

Fix $0<a<1/8$, let

$$I=[0,a), \qquad A=\mathbb{T}\times I,$$

and define

$$Q_\alpha=\left\{n\ge 1:\left\|n^2\alpha-\frac{1}{2}\right\|_{\mathbb{T}}<\frac{a}{32}\right\}. \tag{21}$$

This is an infinite set. Indeed, the usual differencing argument gives equidistribution of $n^2\alpha$: for every nonzero integer ℓ and positive h, the difference of the phases $\ell\alpha(n+h)^2$ and $\ell\alpha n^2$ is a linear phase with irrational frequency.

We have

$$T_m^n(x,z)=\left(x+n\alpha,\ z+mnx+m\binom{n}{2}\alpha\right).$$

Consequently,

$$\mu(A\cap T_m^{-n}A\cap T_m^{-2n}A)=J_a(mn^2\alpha), \tag{22}$$

where

$$J_a(t)=\int_{I\times I}1_I(2v-z+t)\,dz\,dv. \tag{23}$$

Here the arguments are taken modulo 1. Formula (22) follows by using the first two fiber coordinates as the Haar-uniform variables z,v.

If $J_a(t)>0$, then

$$t = w - 2v + z \pmod 1$$

for some $z, v, w \in I$, so necessarily

$$\|t\|_{\mathbb{T}} < 2a.$$

For $n \in Q_\alpha$, this rules out $t = n^2\alpha$. Therefore

$$\boxed{\mu(A \cap T_1^{-n} A \cap T_1^{-2n} A) = 0 \qquad (n \in Q_\alpha).} \tag{24}$$

On the other hand,

$$\|2n^2\alpha\|_{\mathbb{T}} < a/16.$$

If

$$z, v \in [3a/8, 5a/8],$$

then

$$2v - z \in [a/8, 7a/8],$$

so adding any representative of $2n^2\alpha$ in $(-a/16, a/16)$ leaves the result inside I. Hence

$$\boxed{\mu(A \cap T_2^{-n} A \cap T_2^{-2n} A) \geq \frac{a^2}{16} \qquad (n \in Q_\alpha).} \tag{25}$$

The distinction is captured by the phase

$$\rho_{2n}(x) - 2\rho_n(x) = mn^2\alpha, \tag{26}$$

not by the tower seminorms.

This is **not a counterexample to your intermediary theorem**: Q_α does not have the fixed-density hypothesis, by the nilsystem conclusion we have been using. It demonstrates that the seminorm values alone cannot select the correct times, even when all the factors and all the seminorms agree.

It does not rule out a proof using seminorms of time-dependent expressions together with (H_{δ_0}).

4. What the remaining step now looks like

Theorem 1 reduces a hypothetical failed successor step to

$$\mathrm{Re}\, \mathbb{E}_{y \in B_d} \left\langle \pi\big(U_d(y), \ldots, U_{(k-1)d}(y)\big) v, v \right\rangle \leq -c \qquad (d \in D'), \tag{27}$$

where D' is still fixed-density regular.

This is the obstruction I would now try to eliminate. It has several advantages over the original indicator formulation: the representation is finite-dimensional, the expression is fixed, and the negative gap is uniform after conditioning on the base recurrence event.

There is one distinction to preserve. Although

$$(g_1, \ldots, g_{k-1}) \longmapsto \langle \pi(g_1, \ldots, g_{k-1}) v, v \rangle$$

is positive definite on the compact product group, the sequence obtained by substituting

$$\big(U_d(y), U_{2d}(y), \ldots, U_{(k-1)d}(y)\big)$$

and integrating over B_d is **not generally a positive-definite sequence in** d. The different powers do not form one cocycle over one base transformation. Thus a direct application of a one-parameter spectral or van der Corput criterion would be unjustified.

The explicit example above shows the issue in its simplest form: the relevant phase can be $e(n^2\alpha)$, even though the fiber is only one-dimensional.

The unresolved implication is therefore a phase-sensitive one:

$$\boxed{\text{fixed-density regularity for all systems} \quad \Longrightarrow \quad \text{no obstruction of the form } (27) \text{ over a distal base.}}$$

Proving this would finish the compact-extension step, and our inverse-limit argument would then finish the distal induction.

So this pass has not completed the theorem. It has, however, narrowed the remaining counterexample to **one nontrivial finite-dimensional coefficient on a still-regular subset of** D, and established that an argument based only on the scalar distal seminorms loses exactly the information that distinguishes the relevant recurrence phenomena.

---

36. User

2026-09-16 16:00:28 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T08:00:28.910000Z

Great progress! Please keep working.

---

37. Assistant

2026-09-16 16:00:52 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T08:00:52.450890Z

I'll try to use fixed-density regularity to eliminate the negative matrix-coefficient obstruction while preserving the cocycle phases and the same allowed differences.

---

38. Assistant

2026-09-16 16:03:57 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T08:03:57.566838Z

The remaining issue is spatial as well as phase-sensitive: I'm testing relative joinings or auxiliary extensions that retain where bad phases occur while producing a usable density increase.

---

39. Assistant

2026-09-16 16:13:04 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T08:13:04.141139Z

A new limitation appears: negative cocycle bias alone is insufficient, even for $D = \mathbb{N}$; the argument must retain cap avoidance. I'm exploring exact compact-extension coordinate changes.

40. Assistant

2026-09-16 16:33:34 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T08:33:34.060178Z

I obtained a **same-D lifting theorem for arbitrary joinings with compact rotations**, not just independent products. In particular, it proves the compact-extension step for cocycles cohomologous to constants.

But I also found an important correction to the last proposed route:

$$\boxed{\text{A uniformly negative matrix coefficient is not, by itself, an obstruction to recurrence.}}$$

In fact, the precise type of coefficient extracted in my last answer can equal $-1/8$ at **every positive time**, even in an ergodic system measurably isomorphic to a nilsystem. Thus the proposed step of ruling out that coefficient from fixed-density regularity is false. The original spherical-cap obstruction remains valid; the passage to one signed coefficient discarded essential information.

I have not completed the general distal theorem. Below are the explicit counterexample to that auxiliary assertion, the information we need to retain instead, and the lifting theorem that does work.

1. A negative coefficient can coexist with full multiple recurrence

Let

$$\omega = e^{2\pi i/3}, \qquad K = \{1, \omega, \omega^2\},$$

equipped with uniform measure. On $Y = \mathbb{T}^2$, with Haar measure ν, take the skew shift

$$S(x, z) = (x + \alpha, z + x), \qquad \alpha \notin \mathbb{Q}.$$

This is an affine two-step nilsystem; affine unipotent systems are included in the nilsystem framework discussed by Host–Kra–Maass. arXiv

Define

$$h(x, z) = \begin{cases} 1, & 0 \le z < 1/2, \\ \omega, & 1/2 \le z < 1, \end{cases}$$

and the K-valued cocycle

$$\rho(y) = \omega\, h(Sy) h(y)^{-1}.$$

Its iterates telescope:

$$\rho_n(y) = \omega^n h(S^n y) h(y)^{-1}. \tag{1}$$

Consider the compact extension

$$T(y, w) = (Sy, \rho(y) w)$$

on $Y \times K$, with product measure. The coordinate change

$$\Psi(y, w) = (y, h(y)^{-1} w)$$

conjugates T to

$$(y, w) \longmapsto (Sy, \omega w). \tag{2}$$

Thus this is simply a product with a three-point rotation, written in nonconstant fiber coordinates.

It is ergodic. Indeed, the eigenvalues of S are $e(p\alpha)$, $p \in \mathbb{Z}$, so it has no nontrivial eigenvalue in common with the three-point rotation. One can check the eigenvalue assertion directly from

$$Se_{p,q} = e(p\alpha)e_{p+q,q}, \qquad e_{p,q}(x, z) = e(px + qz) :$$

an eigenfunction with $q \neq 0$ would require Fourier coefficients of constant nonzero magnitude along an infinite progression, contradicting square summability.

The coefficient is exactly $-1/8$

Take the nontrivial one-dimensional representation

$$\pi(g_1, g_2) = g_1^{-2} g_2$$

of $\mathbb{T}^2$. Define

$$\Phi(n) = \int_Y \pi(\rho_n(y), \rho_{2n}(y))\, d\nu(y).$$

By (1),

$$\Phi(n) = \int_Y h(y)\, h(S^{2n}y)\, h(S^n y)^{-2}\, d\nu(y).$$

Because $h^3 = 1$, this simplifies to

$$\Phi(n) = \int_Y h(y) h(S^n y) h(S^{2n} y)\, d\nu(y). \tag{3}$$

For every $n \geq 1$, the three random variables

$$h(y), \quad h(S^n y), \quad h(S^{2n} y)$$

are independent, each taking $1, \omega$ with equal probabilities. Here is the verification.

Write

$$z_j = z + jnx + \binom{jn}{2}\alpha \pmod 1, \qquad j = 0, 1, 2,$$

and let $s_j = 1$ when $z_j \in [0, 1/2)$, and $s_j = -1$ otherwise.

Each s_j is balanced. Every pair z_i, z_j, $i \neq j$, is Haar-uniform on $\mathbb{T}^2$, because the corresponding integer linear map in (x, z) has nonzero determinant $(i - j)n$. Thus

$$\mathbb{E}(s_i s_j) = 0.$$

Also, replacing z by $z + 1/2$ changes all three signs, so

$$\mathbb{E}(s_0 s_1 s_2) = 0.$$

Expanding

$$\mathbb{P}(s_0=\varepsilon_0, s_1=\varepsilon_1, s_2=\varepsilon_2)=\frac{1}{8}\,\mathbb{E}\prod_{j=0}^{2}(1+\varepsilon_j s_j)$$

now gives $1/8$ for every sign pattern. This proves independence.

Consequently, (3) gives

$$\boxed{\Phi(n)=\left(\frac{1+\omega}{2}\right)^3=-\frac{1}{8}\qquad(n\geq 1).} \tag{4}$$

This has precisely the form of the last answer's proposed obstruction: $k=3$, $E=Y$, a nontrivial irreducible representation, and a uniform negative constant. There is no small base-return measure involved.

Yet we may take $D=\mathbb{N}$, which satisfies your fixed-density hypothesis at every density by Furstenberg's multiple recurrence theorem. We may equally take $D=3\mathbb{N}$, by applying that theorem to the third power of any transformation. arXiv

There are genuine, quantitatively positive returns at the same times

Let

$$A=Y\times\{1\},\qquad \mu(A)=\frac{1}{3}.$$

For $n\in 3\mathbb{N}$, membership in

$$A\cap T^{-n}A\cap T^{-2n}A$$

is equivalent to

$$w=1,\qquad h(y)=h(S^n y)=h(S^{2n}y).$$

The latter event has probability $2/8=1/4$. Therefore

$$\boxed{\mu(A\cap T^{-n}A\cap T^{-2n}A)=\frac{1}{12}\qquad(n\in 3\mathbb{N}).} \tag{5}$$

Thus a uniformly negative coefficient occurs at times when this very concrete fiber set has a uniformly positive triple intersection.

The consequence for our proof attempt is precise: the deduction of a negative coefficient from a genuine cap-avoidance witness can remain valid, but the negative coefficient alone is too weak to serve as the contradiction.

2. What was lost in the extraction

There are two ways to see the missing information.

A change of fiber coordinates reverses the apparent obstruction

In the coordinates from (2), the cocycle is constant:

$$\rho'(y)=\omega.$$

Hence

$$\pi(\rho'_n, \rho'_{2n}) = \omega^{-2n}\omega^{2n} = 1$$

at every time.

So the same compact extension has the coefficient

$$-\frac{1}{8}$$

in one choice of fiber coordinates and

$$1$$

in another.

Recurrence is unchanged under this conjugacy. What changes along with the cocycle is the location of the fiber test set. A fixed cap becomes a base-dependent cap. Any proof that changes the cocycle coordinates must carry that test set along.

The full positive detector does not have the same problem

Recall the detector

$$p_M(g) = \left(\frac{1 + \operatorname{Re} g}{2}\right)^M, \qquad P_M(g_1, g_2) = p_M(g_1)p_M(g_2).$$

For the third roots of unity,

$$p_M(1) = 1, \qquad p_M(\omega) = p_M(\omega^2) = 4^{-M}.$$

Using the independent values of h_0, h_1, h_2, for $n \in 3\mathbb{N}$ we obtain

$$\boxed{\int_Y P_M(\rho_n, \rho_{2n})\, d\nu = \frac{1}{4} + \frac{1}{2}\,4^{-M} + \frac{1}{4}\,16^{-M} \geq \frac{1}{4}.} \tag{6}$$

The signed coefficient in (4) is negative, but the full positive detector correctly retains the substantial mass of simultaneous returns.

This identifies the point where the preceding reduction became too weak: **one negative summand does not preserve the smallness of the whole nonnegative detector.**

The distal seminorm characterization is still useful for identifying the compact component. Its relative-self-joining interpretation, however, does not supply the additional localization information needed here. arXiv

3. A lifting theorem that does preserve the required information

Here is the positive result from this pass.

Theorem: Lifting through arbitrary joinings with compact rotations

Let (Y, ν, S) be an invertible standard probability-preserving system, and let $D \subseteq \mathbb{N}$.

Assume that, for **every Bohr neighborhood B of zero**, the set

$$D \cap B$$

is a set of multiple measurable recurrence for Y.

Let K be a compact metrizable group with Haar measure m_K, let $a \in K$, and let λ be **any joining** of

$$(Y, \nu, S) \quad \text{and} \quad (K, m_K, L_a).$$

Then D is a set of multiple measurable recurrence for

$$(Y \times K, \lambda, S \times L_a).$$

In particular, the joining need not be the independent product.

The point of the proof is that the fiber approximation error is controlled **relative to the available fiber mass**. It therefore remains valid even when the eventual base recurrence intersection is very small.

Proof

Disintegrate the joining over Y:

$$\lambda = \int_Y \kappa_y \, d\nu(y),$$

where κ_y is a probability measure on K. Invariance gives, on a common conull set,

$$\kappa_{S^n y} = (L_{a^n})_* \kappa_y \qquad (n \in \mathbb{Z}). \tag{7}$$

Fix a measurable $A \subseteq Y \times K$ with $\lambda(A) > 0$, and a length $k \geq 2$. Write

$$A_y = \{z : (y, z) \in A\}, \qquad \varepsilon = \frac{1}{8k}.$$

Step 1: Find a nearly full fiber cell

Choose increasing finite measurable partitions of K that generate its Borel sigma-algebra and whose atoms have Haar-null boundaries. Such partitions can be obtained from a countable basis of metric balls with Haar-null boundaries.

Conditional expectations onto

$$\mathcal{B}_Y \otimes \sigma(\mathcal{P}_m)$$

converge to 1_A. Therefore there are an atom C of one partition, a positive-measure set $E_0 \subseteq Y$, and $b > 0$, such that, writing

$$q(y) = \kappa_y(C),$$

we have

$$q(y) \geq b, \qquad \kappa_y(C \setminus A_y) \leq \varepsilon q(y) \qquad (y \in E_0). \tag{8}$$

To justify the positive lower bound b, first obtain a positive-measure set on which $q(y) > 0$ and the second inequality holds, and then restrict to one of the sets $\{q \geq 1/m\}$.

Step 2: Control the boundary of the cell

Because C has Haar-null boundary, choose a compact set and an open set with

$$F \subseteq \operatorname{int} C, \qquad \overline{C} \subseteq O,$$

such that

$$m_K(O \setminus F) < \frac{\varepsilon b\, \nu(E_0)}{2}.$$

Since the K-marginal of λ is Haar,

$$\int_Y \kappa_y(O \setminus F)\, d\nu(y) = m_K(O \setminus F).$$

After discarding a set of measure less than $\nu(E_0)/2$, we obtain $E \subseteq E_0$, $\nu(E) > 0$, such that

$$\kappa_y(O \setminus F) < \varepsilon b \le \varepsilon q(y) \qquad (y \in E). \tag{9}$$

Step 3: Choose the time through a Bohr localization

Let

$$H = \overline{\{a^n : n \in \mathbb{Z}\}} \subseteq K.$$

This is a compact abelian group. Choose an identity neighborhood $V \subseteq H$ so small that, for every $g \in V$ and $0 \le j < k$,

$$g^j F \subseteq C, \qquad g^{-j} C \subseteq O. \tag{10}$$

Compactness of $F, \overline{C}$ and continuity of multiplication give such a neighborhood.

The set

$$B_V = \{n \ge 1 : a^n \in V\}$$

is a Bohr neighborhood of zero. By hypothesis, choose $d \in D \cap B_V$ such that

$$E_d = \bigcap_{j=0}^{k-1} S^{-jd} E$$

has positive measure.

Step 4: Lift the intersection fiber by fiber

Fix $y \in E_d$. From (9),

$$\kappa_y(F) \ge (1 - \varepsilon) q(y). \tag{11}$$

Also, by (7), (9), and (10),

$$\begin{aligned} q(S^{jd} y) &= \kappa_y(a^{-jd} C) \\ &\le \kappa_y(O) \\ &\le (1 + \varepsilon) q(y). \end{aligned} \tag{12}$$

For a point $z \in F$, condition (10) guarantees $a^{jd} z \in C$. Therefore

$$\begin{aligned}\kappa_y\left(F \setminus \bigcap_{j=0}^{k-1} a^{-jd}A_{S^{jd}y}\right) &\le \sum_{j=0}^{k-1} \kappa_y\big(a^{-jd}(C \setminus A_{S^{jd}y})\big)\\ &= \sum_{j=0}^{k-1} \kappa_{S^{jd}y}(C \setminus A_{S^{jd}y})\\ &\le \varepsilon \sum_{j=0}^{k-1} q(S^{jd}y)\\ &\le k\varepsilon(1+\varepsilon)q(y).\end{aligned}$$

Combining this with (11),

$$\begin{aligned}\kappa_y\left(\bigcap_{j=0}^{k-1} a^{-jd}A_{S^{jd}y}\right) &\ge \big[1-\varepsilon-k\varepsilon(1+\varepsilon)\big]q(y)\\ &> \frac{1}{2}q(y)\\ &\ge \frac{b}{2}.\end{aligned} \tag{13}$$

Integrating over E_d yields

$$\boxed{\lambda\left(\bigcap_{j=0}^{k-1}(S\times L_a)^{-jd}A\right) \ge \frac{b}{2}\,\nu(E_d) > 0.}$$

This proves the theorem. □

4. Consequences for the distal induction

Let

$$\mathcal{C} = \{D \subseteq \mathbb{N} : D \text{ satisfies } (\mathrm{H}_\delta) \text{ for some } \delta \in (0,1)\}.$$

Recall the Bohr-localization result proved earlier in our conversation:

$$D \in \mathcal{C}, \quad B \text{ Bohr}_0 \quad\Longrightarrow\quad D \cap B \in \mathcal{C}. \tag{14}$$

Suppose the induction hypothesis holds on Y: every $D \in \mathcal{C}$ is a set of multiple measurable recurrence for Y. Then (14) supplies the hypothesis of the joining theorem.

Hence:

$$\boxed{\begin{array}{c}\text{The desired property passes from } Y\\ \text{to every joining of } Y \text{ with a compact rotation,}\\ \text{and to every factor of such a joining.}\end{array}} \tag{15}$$

In particular: cocycles cohomologous to constants

Suppose a compact-group cocycle has the form

$$\rho(y) = b(Sy)\,a\,b(y)^{-1},$$

where $b : Y \to K$ is measurable and $a \in K$ is fixed. The change of coordinates

$$(y,z) \longmapsto (y, b(y)^{-1}z)$$

conjugates

$$T(y,z) = (Sy, \rho(y)z)$$

to the product transformation

$$S \times L_a.$$

Thus the desired recurrence conclusion holds for this extension.

This also explains the example in Section 1. Its cocycle looks nonconstant and gives a negative coefficient, but it is cohomologous to a constant. The actual recurrence follows by localizing in the base so that the transfer function b varies little, and imposing an appropriate Bohr condition on a^d.

The distinction from an arbitrary compact extension is essential. Relative compactness permits genuinely variable unitary cocycles; it does not say that they can be made constant by a measurable change of coordinates. The finite-rank module description of compact extensions captures this broader class. arXiv

5. The obstruction that must be retained

The counterexample suggests that the remaining argument should use the **whole nonnegative fiber-intersection function**, not one of its signed Fourier coefficients.

For a sphere extension

$$T(y,z) = (Sy, U(y)z),$$

allow the cap center to depend measurably on the base:

$$u : Y \to \mathbb{S}^{2r-1}.$$

For a positive-measure base set E, define

$$E_d = \bigcap_{j=0}^{k-1} S^{-jd}E$$

and

$$Q_\eta(y,d) = \sigma\left\{ z : \max_{0\le j<k} \|U_{jd}(y)z - u(S^{jd}y)\| < \eta \right\}, \tag{16}$$

with $U_0 = I$.

For

$$A = \{(y,z) : y \in E,\ \|z - u(y)\| < \eta\},$$

the recurrence correlation is exactly

$$\boxed{\mu\left(\bigcap_{j=0}^{k-1} T^{-jd}A\right) = \int_{E_d} Q_\eta(y,d)\, d\nu(y).} \tag{17}$$

This formulation has the invariance that the extracted coefficient lacked. Under a coordinate change $z = b(y)z'$, replace

$$U(y) \quad \text{by} \quad b(Sy)^{-1}U(y)b(y), \qquad u(y) \quad \text{by} \quad b(y)^{-1}u(y).$$

Then $Q_\eta(y,d)$ is unchanged, by unitary invariance of sphere measure.

There is also a useful positive-kernel version. Put

$$\mathcal{K}_M(y,d) = \int_{\mathbb{S}^{2r-1}} \prod_{j=0}^{k-1} \left(\frac{1 + \mathrm{Re}\langle U_{jd}(y)z, u(S^{jd}y)\rangle}{2} \right)^M d\sigma(z). \tag{18}$$

This too is unchanged under the same coordinate transformation.

If $Q_\eta(y,d) = 0$, the intersection of the relevant open caps is empty. Hence for every z, at least one factor in (18) is at most $1 - \eta^2/4$, giving

$$\boxed{\mathcal{K}_M(y,d) \le \left(1 - \frac{\eta^2}{4}\right)^M \qquad (M \ge 1).} \tag{19}$$

Thus a genuine avoidance witness supplies **uniform exponential decay of an entire family of positive kernels**, not merely a negative average of one coefficient. Equation (6) shows explicitly how the latter can hold while the former fails completely.

I have not yet converted (19) into the fixed-length high-density test needed to contradict (H_δ). That remains the missing argument for genuinely variable compact cocycles. The standard compact-extension proof of multiple recurrence concerns averages over unrestricted times and does not automatically provide this same- D conclusion. arXiv

So the general intermediary theorem remains unproved here. The corrected route is to preserve the full, nonnegative cap-intersection data—including its behavior under changes of fiber coordinates. The joining theorem establishes a valid successor step for additional compact rotations and for cocycles cohomologous to constants; the explicit $-1/8$ example shows why the previous single-coefficient route cannot establish the remaining cases.

---

41. User

2026-09-16 17:15:24 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T09:15:24.243000Z

Let's focus in on a special case of the intermediary theorem. If D is delta density large for some delta in (0,1), can we show that D is delta density large for every delta in (0,1) for the class of 2-step distal systems?

---

42. Assistant

2026-09-16 17:15:50 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T09:15:50.756647Z

I'll focus on compact extensions of Kronecker systems, where the two-step structure may provide the phase control missing from the general distal argument.

43. Assistant

2026-09-16 17:18:57 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T09:18:57.969576Z

The remaining case is genuinely variable measurable cocycles over rotations; affine or cohomologous-to-constant cocycles are already covered, so I'll test whether rotation structure simplifies their iterates.

---

44. Assistant

2026-09-16 17:28:08 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T09:28:08.970268Z

I now have two partial results: generic cocycles are rigid along D. More concretely, if $\delta_0 \leq 1/m$, every arbitrary measurable m-point extension of a Kronecker system satisfies the conclusion.

---

45. Assistant

2026-09-16 17:42:04 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T09:42:04.382429Z

I can now prove the conclusion for a class of two-step distal systems with arbitrary measurable cocycles: finite-fibre extensions of Kronecker systems, provided the original density threshold is at most the reciprocal of the fibre size. The endpoint is included.

More precisely:

> **Finite-fibre theorem.** Suppose $D \subseteq \mathbb{N}$ is δ_0-density large for all measure-preserving systems, where $0 < \delta_0 < 1$. Let
>
> $$T(y, i) = (y + \theta, \rho(y)i)$$
>
> on
>
> $$K \times \{1, \dots, m\},$$
>
> with product of Haar measure and uniform counting measure, where $y \mapsto y + \theta$ is an ergodic compact rotation and
>
> $$\rho : K \longrightarrow \mathrm{Sym}(m)$$
>
> is **any measurable cocycle**. If
>
> $$\boxed{\delta_0 \leq \frac{1}{m},}$$
>
> then D is a set of multiple measurable recurrence for T. Equivalently, it is δ-density large for this system for every $\delta > 0$.

Thus, in particular,

$$\boxed{\delta_0 \leq \tfrac{1}{2} \quad \Longrightarrow \quad \text{the desired conclusion for every two-point extension of a Kronecker system.}}$$

There is no continuity assumption on ρ, and no assumption that it is cohomologous to a constant or to a nilsystem cocycle. These are two-step measurably distal systems because the second extension has finite, isometric fibres. The general compact-extension framework also allows infinite compact homogeneous fibres, so this does not cover the whole class. arXiv

For arbitrary two-step distal systems, and arbitrary $\delta_0 < 1$, I still do not have a complete proof. The new ingredient is a conditional-density principle over the **entire Kronecker factor**, stronger than the rational-factor estimate we obtained earlier.

1. The stronger conditional-density principle

Use the measure-preserving formulation of your hypothesis:

$$\begin{gathered}\text{For every probability-preserving } (Z,\lambda,R),\ \lambda(B)\ge\delta_0,\ L\ge 2,\\ \exists d\in D:\quad \lambda\left(\bigcap_{j=0}^{L-1} R^{-jd}B\right)>0.\end{gathered}\tag{H_{δ_0}}$$

Here is the principle that leads to the finite-fibre theorem.

> **Conditional-density theorem.** Assume (H_{δ_0}). Let $X\to Y$ be any extension of a Kronecker system—not necessarily a compact extension—and let $A\subseteq X$ have positive measure. If
>
> $$\boxed{\|\mathbb{E}(1_A\mid Y)\|_\infty\ge\delta_0,}\tag{1}$$
>
> then, for every $k\ge 2$, there is $d\in D$ such that
>
> $$\mu\left(\bigcap_{j=0}^{k-1} T^{-jd}A\right)>0.\tag{2}$$

The distinction from the original hypothesis is substantial: $\mu(A)$ itself can be arbitrarily small. It suffices that A has sufficiently large **conditional density on some Kronecker fibres**, in the essential-supremum sense.

Even the equality case in (1) works; the essential supremum need not be attained.

The proof has two parts. First, one can spread high conditional density across most of a compact base, using different fixed translates of A on separated compact pieces. Second, our root construction supplies the strict increment needed to handle the endpoint.

2. Extremal notation and the localization input

As before, define

$$a_k(D)=\sup\left\{\mu(A):\mu\left(\bigcap_{j=0}^{k-1}T^{-jd}A\right)=0\text{ for every }d\in D\right\},$$

where the supremum is over probability-preserving systems, and put

$$c(D)=\sup_{k\ge 2}a_k(D).$$

Your hypothesis implies

$$c(D) \le \delta_0 < 1. \tag{3}$$

We will use the Bohr-localization identity established earlier:

$$\boxed{c(D \cap B) = c(D) \quad \text{whenever } B \text{ is a Bohr neighborhood of zero.}} \tag{4}$$

Recall the mechanism behind (4). Products of avoidance witnesses give

$$c(E \cup F) \ge c(E)c(F).$$

The complement of a Bohr neighborhood has avoidance witnesses of density arbitrarily close to 1, so

$$c(\mathbb{N} \setminus B) = 1.$$

Applying the product inequality to

$$D = (D \cap B) \cup (D \setminus B)$$

gives (4). Thus the argument below only uses the already-proved critical-density localization, not the stronger endpoint formulation of fixed-density localization.

3. Spreading conditional density across a compact base

Lemma 1: Kronecker-fibre bound

Suppose $X \to Y$ is an extension of a Kronecker system and

$$\mu\left(\bigcap_{j=0}^{k-1} T^{-jd} A\right) = 0 \qquad (d \in D). \tag{5}$$

Then

$$\boxed{\|\mathbb{E}(1_A \mid Y)\|_\infty \le c(D).} \tag{6}$$

This improves the earlier bound over $\mathcal{I}(T^q)$: it applies to the full compact factor, including irrational eigenfunctions.

Proof

First suppose the base is an ergodic rotation

$$S : y \longmapsto y + \theta$$

on a compact metrizable abelian group K, with Haar measure ν. Fix a compatible translation-invariant metric on K.

Write

$$p(y) = \mu_y(A), \qquad \alpha = \operatorname*{ess\,sup}_{y \in K} p(y).$$

Choose

$$0 < b < \alpha$$

and let

$$E = \{y : p(y) > b\}.$$

Then $\nu(E) > 0$.

Fix $\varepsilon > 0$. Ergodicity gives finitely many nonnegative integers

$$n_1, \dots, n_s$$

such that

$$\nu\left(\bigcup_{i=1}^{s} S^{-n_i} E\right) > 1 - \varepsilon.$$

Partition this union into disjoint measurable sets $P_i \subseteq S^{-n_i} E$. By inner regularity, choose disjoint compact sets

$$K_i \subseteq P_i$$

such that

$$\nu\left(\bigcup_{i=1}^{s} K_i\right) > 1 - 2\varepsilon. \tag{7}$$

The nonempty compact sets K_i are pairwise positively separated. Choose $\eta > 0$ smaller than all their pairwise distances, and define the Bohr neighborhood

$$B_\eta = \{d \geq 1 : \operatorname{dist}(d\theta, 0) < \eta\}.$$

Now define a new set in the **same extension**:

$$C = \bigcup_{i=1}^{s} \left(\pi^{-1}(K_i) \cap T^{-n_i} A\right). \tag{8}$$

For $y \in K_i$, its conditional measure is

$$\mu_y(C) = \mu_y(T^{-n_i} A) = p(S^{n_i} y) > b.$$

Consequently,

$$\mu(C) > b(1 - 2\varepsilon). \tag{9}$$

The important observation is that C remains a k-term avoidance witness along $D \cap B_\eta$.

Indeed, suppose

$$x, T^d x, \dots, T^{(k-1)d} x \in C$$

with $d \in B_\eta$. Their base points lie in the union of the K_i, and consecutive base points are less than η apart. They therefore all lie in the **same** compact piece K_i. By (8),

$$T^{n_i} x,\ T^{n_i + d} x, \dots, T^{n_i + (k-1)d} x \in A.$$

For $d \in D$, this is a null event by (5). More formally, the intersection for C is contained in the finite union

$$\bigcup_{i=1}^{s} T^{-n_i} \left(\bigcap_{j=0}^{k-1} T^{-jd} A\right).$$

Hence

$$b(1-2\varepsilon) < \mu(C) \le a_k(D \cap B_\eta) \le c(D \cap B_\eta) = c(D).$$

Let $\varepsilon \downarrow 0$, then $b \uparrow \alpha$, to obtain (6).

For a nonergodic compact rotation, disintegrate over its ergodic components. Since D is countable, (5) holds on almost every component simultaneously for every $d \in D$. The preceding argument then proves the same bound. □

Why the base being Kronecker matters

The construction uses one neighborhood of zero for which every allowed base displacement is uniformly small:

$$\operatorname{dist}(S^d y, y) < \eta \qquad \text{for all } y.$$

That prevents the orbit segment from switching between the compact pieces K_i, so the translating integer n_i stays fixed throughout the progression.

There is no extracted multiplier of d. This is precisely where this argument avoids the difficulty in the general compact-extension proof.

4. Making the conditional bound strict

Lemma 1 alone gives only $\alpha \le c(D)$. To obtain the conditional-density theorem at the endpoint, we amplify α, not merely the global measure $\mu(A)$.

Let W_k be the van der Waerden number for a monochromatic k-term progression in a k-coloring, and put

$$M_k = \left\lfloor \frac{W_k - 1}{k-1} \right\rfloor, \qquad L_k = W_k + k - 1. \tag{10}$$

The finiteness of these numbers is the finite van der Waerden theorem. arXiv

Lemma 2: Conditional-density amplification

Under the hypotheses of Lemma 1, let

$$\alpha = \|\mathbb{E}(1_A \mid Y)\|_\infty .$$

Then

$$\boxed{\alpha + (1-\alpha)\alpha^{M_k} \le c(D).} \tag{11}$$

Proof

For each $1 \le q \le M_k$, form the q-root tower

$$X_q = X \times \{0, \dots, q-1\}$$

with

$$R_q(x,i) = \begin{cases} (x, i+1), & i < q-1, \\ (Tx, 0), & i = q-1, \end{cases}$$

and set

$$A_q = A \times \{0, \dots, q-1\}.$$

Then

$$R_q^{qd}(x,i) = (T^d x, i),$$

so

$$\mu_q\left(\bigcap_{j=0}^{k-1} R_q^{-jqd} A_q\right) = 0 \qquad (d \in D). \tag{12}$$

The corresponding tower over the base Y is again a Kronecker system. For example, if $g \circ S = \lambda g$ and $\zeta^q = \lambda$, then

$$(y,i) \longmapsto \zeta^i g(y)$$

is an eigenfunction of the root transformation. Taking all q-th roots and a complete family of base eigenfunctions gives discrete spectrum.

Take the independent product

$$(Z, \lambda, V) = (X, \mu, T) \times \prod_{q=1}^{M_k} (X_q, \mu_q, R_q),$$

and define

$$H = \left(A \times \prod_q X_q\right) \cup \left(X \times \prod_q A_q\right). \tag{13}$$

Its natural base is the product of the original Kronecker system and the root Kronecker systems, hence is again a compact rotation, though possibly nonergodic.

Conditional on the base coordinates,

$$\lambda_{\mathbf{y}}(H) = p(y_0) + (1 - p(y_0)) \prod_{q=1}^{M_k} p(y_q).$$

Because the base coordinates have product measure,

$$\boxed{\left\| \mathbb{E}(1_H \mid \text{product base}) \right\|_\infty = \alpha + (1-\alpha)\alpha^{M_k}.} \tag{14}$$

It remains to verify that H is an L_k-term avoidance witness along the **original** D.

Every red-blue coloring of

$$\{0, \dots, L_k - 1\}$$

contains either k consecutive red indices or a blue k-term progression whose common difference is at most M_k. To see this, if there are no k consecutive red indices, for every $n < W_k$ choose

$$t(n) \in \{0, \dots, k-1\}$$

with $n + t(n)$ blue. A monochromatic k-term progression for the k-coloring $n \mapsto t(n)$ supplies the required blue progression.

Now suppose an L_k-term V^d-orbit segment lies in H. Color an index red when its first coordinate belongs to A, and blue otherwise. At every blue index, all auxiliary coordinates belong to their respective A_q.

A consecutive red k-block contradicts the original avoidance property. A blue k-term progression of index difference q contradicts (12) in its q-th root coordinate. Thus the full intersection has measure zero for every $d \in D$.

Apply Lemma 1 to H over its product Kronecker base. Equation (14) gives (11). □

Consequence: a strict fibre-density gap

Since $\mu(A) > 0$, we have $\alpha > 0$. If $c(D) < 1$, (11) first rules out $\alpha = 1$, and then gives

$$\boxed{\alpha < c(D) \le \delta_0.} \tag{15}$$

This proves the conditional-density theorem: an avoidance witness cannot have essential conditional supremum at least δ_0.

There is also a quantitative version. Let $\beta_{k,\delta_0} \in (0, \delta_0)$ be the unique solution of

$$\beta + (1 - \beta)\beta^{M_k} = \delta_0.$$

Then any k-term avoidance witness satisfies

$$\boxed{\|\mathbb{E}(1_A \mid Y)\|_\infty \le \beta_{k,\delta_0} < \delta_0.} \tag{16}$$

The gap depends only on k, δ_0, not on the system or cocycle.

5. Proof of the finite-fibre theorem

Now suppose

$$X = K \times \{1, \dots, m\}, \qquad T(y, i) = (y + \theta, \rho(y)i),$$

with uniform fibre measures.

For every measurable $A \subseteq X$,

$$p(y) = \mu_y(A) = \frac{\#\{i : (y, i) \in A\}}{m}.$$

Thus

$$p(y) \in \left\{0, \frac{1}{m}, \frac{2}{m}, \dots, 1\right\}. \tag{17}$$

If $\mu(A) > 0$, then $p > 0$ on a positive-measure set, and therefore

$$\|p\|_\infty \ge \frac{1}{m}. \tag{18}$$

Assume $\delta_0 \le 1/m$. Equations (18) and (1) give multiple recurrence for A along D.

Equivalently, a hypothetical k-term avoidance witness would satisfy both

$$\frac{1}{m} \le \|p\|_\infty$$

and, by (15),

$$\|p\|_\infty < \delta_0 \le \frac{1}{m},$$

a contradiction. □

This is a full conclusion at every positive target density and every progression length, not merely a bound for sets of measure at least $1/m$. The set A may be supported over an arbitrarily small positive-measure portion of the base.

A stronger threshold for pair recurrence

For pairs alone, the threshold can be improved:

$$\boxed{\delta_0 \le \frac{2}{m} - \frac{1}{m^2} \quad \Longrightarrow \quad D \text{ is a set of measurable pair recurrence for every such } m\text{-point extension.}} \tag{19}$$

Here is the short argument. If A avoids pairs along D, use only its square-root tower and the union construction

$$H = (A \times X_2) \cup (X \times A_2).$$

Its essential conditional supremum is

$$2\alpha - \alpha^2.$$

It avoids five-term progressions along D: a red-blue coloring of five consecutive indices cannot simultaneously have no adjacent red pair and no blue pair at distance 2. Apply the strict bound (15) to H:

$$2\alpha - \alpha^2 < c(D) \le \delta_0.$$

But $\alpha \ge 1/m$, contradicting (19).

For a two-point extension, this gives

$$\delta_0 \le \frac{3}{4} \quad \Longrightarrow \quad \text{pair recurrence at every positive measure.}$$

The all-length conclusion proved above still requires $\delta_0 \le 1/2$.

6. What remains for all two-step distal systems

The finite-fibre argument succeeds because a nonempty fibre section has a uniform positive lower bound on its measure. In an infinite compact homogeneous fibre, this is no longer true: a positive-measure set can satisfy

$$0 < \mu_y(A) \ll \delta_0$$

on every fibre it meets. Such a witness is not excluded by (16).

It is important here not to identify finite-rank relative modules with finite fibres. A circle extension already has rank-one modules generated by individual fibre characters, while its fibre measure is nonatomic. The finite-rank module characterization of compact extensions therefore does not provide the lower bound

$1/m$ used in (18). arXiv

Nevertheless, the conditional-density estimate sharpens the possible obstruction in the full two-step case. If A avoids k-term progressions along D, then over its Kronecker base,

$$\mu_y(A) \le \beta_{k,\delta_0} < \delta_0 \quad \text{almost everywhere.}$$

Moreover, writing $a = \mu(A)$, $c = c(D)$, and $p = \mathbb{E}(1_A \mid Y)$, Lemma 1 gives

$$\begin{aligned} \|p - a\|_2^2 &= \int p^2 \, d\nu - a^2 \\ &\le c \int p \, d\nu - a^2 \\ &= a(c - a). \end{aligned} \tag{20}$$

Thus near-extremal avoidance witnesses are nearly balanced on the **whole Kronecker factor**, not merely on its rational part.

This is compatible with the distal-seminorm viewpoint: the first distal factor is the Kronecker factor, and the next is its maximal isometric extension. But (20) now comes directly from fixed-density largeness and an exact geometric localization argument, rather than from discarding the phases in conditional correlations. arXiv

So the unrestricted two-step intermediary theorem remains unproved here. What is now established is the conditional-density theorem and, as a consequence, the full intermediary conclusion for arbitrary measurable m-point extensions whenever $\delta_0 \le 1/m$, including the endpoint. The remaining task is to obtain a comparable amplification inside nonatomic compact fibres, where no uniform minimum positive fibre measure is available.

---

46. User

2026-09-17 10:44:01 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-17T02:44:01.195000Z

good progress. Please keep working.

---

47. Assistant

2026-09-17 10:44:25 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-17T02:44:25.622507Z

I'll test whether the finite-fiber argument extends to higher thresholds, then seek a circle-fiber analogue that preserves every recurrence time in D.

---

48. Assistant

2026-09-17 10:47:32 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-17T02:47:32.054896Z

The key bottleneck is the minimum positive fiber mass; I'm testing whether finite-group symmetries can amplify it without changing the allowed difference.

49. Assistant

2026-09-17 11:06:32 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-17T03:06:32.708172Z

The density restriction can now be removed completely for pair recurrence in every two-point extension of a Kronecker system. Previously, I had this only for $\delta_0 \le 3/4$.

The new result is:

> **Theorem.** Suppose $D \subseteq \mathbb{N}$ is δ_0-density large for all measure-preserving systems, for some $\delta_0 < 1$. Let
>
> $$T(y,i) = (Sy, i + \rho(y) \pmod 2)$$
>
> be any measurable two-point extension of an ergodic compact rotation, with the product of Haar measure and uniform counting measure. Then
>
> $$\boxed{\mu(A) > 0 \implies \exists d \in D : \ \mu(A \cap T^{-d}A) > 0.}$$
>
> There is no regularity or cohomological assumption on ρ.

These systems are a subclass of the two-step measurably distal systems under consideration: the first stage is a compact rotation, and the second is a finite isometric extension. The general compact-extension framework also includes infinite compact homogeneous fibres. arXiv

This does not yet prove multiple recurrence for two-point extensions, or the full intermediary theorem. But it gives an exact, arbitrarily large density amplification for a genuine compact-extension obstruction. The amplification is

$$\boxed{\frac{1}{2} \ \longmapsto \ 1 - 2^{-r}}$$

at the level of conditional fibre density, while keeping the allowed differences in D.

The mechanism is different from our earlier van der Waerden amplification: powers-of-two roots force an orbit segment to visit **every binary pattern**, not merely some monochromatic subprogression.

1. The two ingredients already available

Recall

$$a_\ell(D) = \sup\left\{\mu(B) : \mu\left(\bigcap_{j=0}^{\ell-1} R^{-jd}B\right) = 0 \text{ for every } d \in D\right\}, \qquad c(D) = \sup_{\ell \ge 2} a_\ell(D).$$

Your hypothesis implies

$$c(D) \le \delta_0 < 1. \tag{1}$$

We proved that Bohr localization preserves this critical density:

$$c(D \cap B) = c(D) \qquad (B \text{ a Bohr neighborhood of } 0). \tag{2}$$

The second ingredient is the conditional-density bound from the preceding answer:

> **Kronecker-fibre bound.** If $X \to Y$ is an extension of a Kronecker system and $C \subseteq X$ avoids ℓ-term progressions along D, then
>
> $$\boxed{\|\mathbb{E}(1_C \mid Y)\|_\infty \le c(D).} \tag{3}$$

Only this non-strict version is needed below.

Here is a reminder of why (3) works. Suppose the conditional density of C is greater than b on a positive-measure base set E. Finitely many translates of E cover almost all of an ergodic compact base. Choose separated compact pieces K_i inside those translates, and on K_i use the corresponding fixed translate $R^{-n_i}C$. This produces a set of global measure arbitrarily close to b.

Restrict the allowable differences to a Bohr neighborhood small enough that consecutive base points cannot move between distinct K_i. Any progression in the new set then lies in one fixed translate of C, so it is still forbidden. Equation (2) gives $b \le c(D)$. Letting b approach the essential supremum proves (3). For a nonergodic compact rotation, apply the same argument on its ergodic components.

The key feature is that the translating integer n_i stays fixed along the progression: **the difference remains d, not an extracted multiple of d.**

2. A binary-array lemma

The new combinatorial observation is elementary but useful.

Lemma 1: Dyadic antiperiodicity fills the whole cube

Let $r \ge 1$, $L = 2^r$, and suppose

$$b_i : \{0, \dots, L-1\} \longrightarrow \{0,1\}, \qquad 0 \le i < r,$$

satisfy

$$b_i(t + 2^i) = 1 - b_i(t) \qquad \text{whenever } t + 2^i < L. \tag{4}$$

Then the L vectors

$$\big(b_0(t), \dots, b_{r-1}(t)\big), \qquad 0 \le t < L,$$

run through every element of $\{0,1\}^r$, exactly once.

In particular, one of them is $(0, \dots, 0)$.

Proof. Induct on r. The assertion for $r = 1$ is immediate.

For $i < r-1$, applying (4) twice gives period 2^{i+1}. Since 2^{i+1} divides $L/2$,

$$b_i(t + L/2) = b_i(t) \qquad (0 \le t < L/2).$$

For the last coordinate, (4) gives

$$b_{r-1}(t + L/2) = 1 - b_{r-1}(t).$$

By induction, the first $r-1$ coordinates run through $\{0,1\}^{r-1}$ exactly once during the first half of the interval. Each such vector appears again in the second half, with the last bit reversed. Thus every r-bit vector appears exactly once. □

This is stronger than a Ramsey conclusion. No subprogression or multiplier has to be selected: the entire interval already contains every pattern.

3. A local half-fibre obstruction forces $c(D) = 1$

The main argument actually applies beyond two-point extensions.

Theorem 2: Half-fibre amplification

Let

$$\pi : (X, \mu, T) \longrightarrow (Y, \nu, S)$$

be an extension of a Kronecker system. Suppose there are $E \subseteq Y$, $\nu(E) > 0$, and $A \subseteq \pi^{-1}(E)$ such that

$$\mu_y(A) = \frac{1}{2} \qquad \text{for almost every } y \in E. \tag{5}$$

If

$$\mu(A \cap T^{-d}A) = 0 \qquad (d \in D), \tag{6}$$

then

$$\boxed{c(D) = 1.} \tag{7}$$

There is no compactness assumption on $X \to Y$ in this statement. The compactness is needed only for the base and the fibre-spreading bound.

Step 1: Pair avoidance becomes exact complementation

Fix $d \in D$. On

$$E \cap S^{-d}E,$$

both A and $T^{-d}A$ have conditional measure $1/2$:

$$\mu_y(T^{-d}A) = \mu_{S^d y}(A) = \frac{1}{2}.$$

Equation (6) says that these two half-fibre sets are disjoint almost everywhere. Therefore they partition the fibre:

$$\boxed{1_A(T^d x) = 1 - 1_A(x)} \tag{8}$$

for almost every x whose base point belongs to $E \cap S^{-d}E$.

This exact identity is the extra information that was absent from our earlier general avoidance construction.

Step 2: Take roots of orders $1, 2, 4, \ldots, 2^{r-1}$

Fix $r \geq 1$, and put

$$q_i = 2^i, \qquad 0 \leq i < r.$$

For each i, form the q_i-root tower

$$X_i = X \times \{0, \dots, q_i - 1\},$$

with transformation

$$R_i(x, a) = \begin{cases} (x, a+1), & a < q_i - 1, \\ (Tx, 0), & a = q_i - 1. \end{cases}$$

Then

$$R_i^{q_i d}(x, a) = (T^d x, a). \tag{9}$$

Its base is the analogous root tower over Y, which is again a Kronecker system. For example, when $Y = K$ and $Sy = y + \theta$, the root base is a compact abelian group on $K \times \{0, \dots, q_i - 1\}$, with the carry in the clock coordinate adding θ. Its rotation by one clock unit is exactly the root transformation.

Write

$$A_i = A \times \{0, \dots, q_i - 1\}, \qquad E_i = E \times \{0, \dots, q_i - 1\}.$$

Take the independent product system

$$(Z, \lambda, V) = \prod_{i=0}^{r-1} (X_i, \mu_i, R_i),$$

with its product Kronecker base $Y^{(r)}$.

Step 3: Retain every binary fibre pattern except the all-zero pattern

Define $C_r \subseteq Z$ by requiring that every base coordinate belong to E_i, and that at least one coordinate belong to A_i:

$$C_r = \{(z_0, \dots, z_{r-1}) : \pi_i(z_i) \in E_i \ \forall i,\ z_i \in A_i \text{ for at least one } i\}. \tag{10}$$

On the positive-measure base set

$$E^{(r)} = \prod_{i=0}^{r-1} E_i,$$

the r membership bits are conditionally independent and balanced. Hence

$$\mathbb{E}(1_{C_r} \mid Y^{(r)}) = \begin{cases} 1 - 2^{-r}, & \text{on } E^{(r)}, \\ 0, & \text{outside } E^{(r)}. \end{cases}$$

In particular,

$$\boxed{\left\| \mathbb{E}(1_{C_r} \mid Y^{(r)}) \right\|_\infty = 1 - 2^{-r}.} \tag{11}$$

Step 4: C_r avoids 2^r-term progressions along the original D

Set $L = 2^r$, and fix $d \in D$. Suppose an orbit segment

$$z, V^d z, \dots, V^{(L-1)d} z$$

lies in C_r.

For $0 \le i < r$, define the membership bit

$$b_i(t) = 1_{A_i}(R_i^{td} z_i), \qquad 0 \le t < L.$$

Every relevant base coordinate lies in E_i. Whenever $t + q_i < L$, equation (9) identifies the change from time t to time $t + q_i$ with applying T^d to the original X-coordinate. Thus (8) gives

$$b_i(t + q_i) = 1 - b_i(t). \tag{12}$$

Lemma 1 now supplies a time t at which

$$b_0(t) = \cdots = b_{r-1}(t) = 0.$$

But this contradicts the definition of C_r.

The identities used here hold outside a finite union of translated null sets for each fixed d. Consequently,

$$\boxed{\lambda\left(\bigcap_{t=0}^{2^r-1} V^{-td} C_r\right) = 0 \qquad (d \in D).} \tag{13}$$

Step 5: Apply the Kronecker-fibre bound

Equations (11), (13), and (3) give

$$1 - 2^{-r} \le c(D).$$

Since r is arbitrary,

$$c(D) = 1.$$

This proves Theorem 2. □

The improvement over the earlier root construction is substantial:

$$\text{old gain: } a + (1-a)a^M; \qquad \text{new half-fibre density: } 1 - 2^{-r}.$$

The new expression tends to 1. It is not an iteration whose guaranteed increments might have a finite total below 1.

4. Deduction for every positive-measure set in a two-point extension

Now let

$$X = Y \times \{0,1\}, \qquad T(y,i) = (Sy, i + \rho(y) \pmod 2),$$

with uniform conditional fibre measures.

Take any measurable $B \subseteq X$ with $\mu(B) > 0$. Let

$$E = \{y : B_y \neq \varnothing\}.$$

This has positive measure. Choose one point of B_y measurably for each $y \in E$, and let $A \subseteq B$ be the resulting graph. Then

$$\mu_y(A) = \frac{1}{2} \qquad (y \in E).$$

If B failed pair recurrence along D, its subset A would also satisfy

$$\mu(A \cap T^{-d}A) = 0 \qquad (d \in D).$$

Theorem 2 would give $c(D) = 1$, contradicting (1).

Thus every positive-measure B has a pair return along D, with **no restriction on δ_0 other than $\delta_0 < 1$.**

Notice that $\mu(B)$ may be arbitrarily small: the graph A can be supported over an arbitrarily small positive-measure part of the base.

A stronger contrapositive

The proof gives more than the stated recurrence conclusion:

> If D fails pair recurrence in a two-point extension of a Kronecker system, then for every $\gamma < 1$ there is a finite-fibre extension of a Kronecker system containing a set of measure greater than γ that avoids sufficiently long progressions along D.

To see the global-density assertion, first use C_r and fibre spreading to obtain a set of global measure close to $1 - 2^{-r}$, avoiding differences in $D \cap B$ for a Bohr neighborhood B. Then take its product with the high-density rotation witness that forbids differences outside B, as in our Bohr-localization proof.

All these auxiliary systems remain finite-fibre extensions of compact rotations. The final forbidden length may exceed 2^r, because of the additional Bohr-complement test.

Thus, in this special case, **a low-density failure really has been amplified into failures at every density below one, within the two-step class itself**.

5. Pair-return sets now admit exact critical-density localization

There is also a useful new localization consequence.

Let X be a two-point extension of a Kronecker system, and let $A \subseteq X$ have positive measure. Define

$$R_X(A) = \{n \geq 1 : \mu(A \cap T^{-n}A) > 0\}.$$

Corollary 3

For every $D \subseteq \mathbb{N}$,

$$\boxed{c\big(D \cap R_X(A)\big) = c(D).} \tag{14}$$

Proof. Put

$$F = \mathbb{N} \setminus R_X(A).$$

The set F fails the pair-recurrence test supplied by A. Theorem 2 and the graph reduction therefore give

$$c(F) = 1.$$

Recall the product inequality

$$c(P \cup Q) \geq c(P)c(Q).$$

Since $D \setminus R_X(A) \subseteq F$, monotonicity gives

$$c(D \setminus R_X(A)) = 1.$$

Hence

$$\begin{aligned} c(D) &= c\big((D \cap R_X(A)) \cup (D \setminus R_X(A))\big) \\ &\geq c(D \cap R_X(A)). \end{aligned}$$

The reverse inequality follows from $D \cap R_X(A) \subseteq D$. $\square$

So the new conclusion is not merely that D meets each pair-return set. We can **restrict D to such a return set without increasing its critical density at all**.

In particular, finitely many pair-return conditions from possibly different two-point extensions can be imposed simultaneously while preserving $c(D)$. This enlarges the collection of recurrence conditions available for localization beyond the Bohr conditions used in the preceding proof.

6. Why the argument does not yet prove three-term recurrence

The decisive implication was

$$\text{two disjoint half-fibre sets} \quad \Longrightarrow \quad \text{they are exact complements.}$$

For a three-term avoidance witness, the corresponding assertion is weaker. In a two-point extension, after restricting to a graph over E, a forbidden triple says that the relevant three membership bits cannot all be 1; fibre symmetry also rules out their all being 0. But it does **not** determine the next bit from the previous one.

The dyadic cube lemma genuinely fails under that weaker condition.

For example, take three coordinates with root orders

$$q_0 = 1, \qquad q_1 = 2, \qquad q_2 = 4,$$

and define, for all integers t,

$$b_i(t) = 1_{\{t \equiv i \pmod 3\}}, \qquad i = 0, 1, 2.$$

For each i, the triple

$$b_i(t),\ b_i(t + q_i),\ b_i(t + 2q_i)$$

contains both 0 and 1, because q_i is nonzero modulo 3.

Nevertheless,

$$b_0(t) + b_1(t) + b_2(t) = 1 \qquad \text{for every } t.$$

Thus the vector sequence never visits $(0, 0, 0)$. It stays forever in the set that retains every pattern except the all-zero pattern.

This is only a counterexample to that attempted extension of the **combinatorial lemma**, not a counterexample to your intermediary theorem. It shows why the pair proof cannot simply be repeated with “no monochromatic triple” replacing exact complementation.

The threshold-free pair case for two-point extensions is now proved, together with the stronger localization identity (14). What remains is to obtain comparable amplification from three-term or higher-order avoidance, and to deal with compact fibres that are not two-point fibres. The exact complementation identity provides the required amplification for pairs; its replacement for longer patterns is the next missing step.

Saved alternate messages

50. User

2026-09-16 18:05:38 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T10:05:38.253000Z

please keep working on the 2-step distal case.

51. Assistant

2026-09-16 18:06:00 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T10:06:00.795970Z

I'll focus on amplifying small conditional densities in nonatomic compact fibers, while also checking whether the finite-fiber argument yields thresholds beyond $1/m$.

52. Assistant

2026-09-16 18:09:06 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T10:09:06.204853Z

I'm now testing tuple constructions and fiber symmetries, whose invariant relative positions may provide extra amplification beyond what the cocycle-independent Kronecker conditional-density bound gives.

53. Assistant

2026-09-16 18:18:12 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T10:18:12.630504Z

I'm separating automatic high-density recurrence from control of genuinely small fiber sets, while keeping the all-systems hypothesis explicit in the argument.

54. Assistant

2026-09-16 18:35:15 (Asia/Shanghai, UTC+08:00)
Source timestamp: 2026-09-16T10:35:15.255701Z

I can now prove the desired conclusion for a **comeagre family of two-step distal cocycles, with no restriction on the starting density** $\delta_0 < 1$. This family includes genuinely two-step systems that are **not inverse limits of nilsystems**.

This does **not** prove the conclusion for every two-step distal system. But it gives a full all-density result in a class with nonatomic compact fibres, rather than another finite-fibre threshold improvement.

The stronger statement is that, generically, we can restrict D to times when the entire system is close to the identity **without losing its original universal fixed-density regularity**.

1. The new theorem

Fix an aperiodic ergodic compact rotation (Y, ν, S), a compact metrizable group G, and a closed subgroup $H \leq G$. For a measurable cocycle

$$\rho : Y \longrightarrow G,$$

consider

$$T_\rho(y, zH) = (Sy, \rho(y)zH)$$

on

$$(X, \mu) = (Y \times G/H, \nu \times m_{G/H}).$$

These are compact homogeneous extensions of a Kronecker system—the standard two-step measurably distal framework. No continuity assumption is imposed on ρ. arXiv

Equip the cocycle space

$$\mathscr{Z} = L^0(Y, G)$$

with convergence in measure. Here "comeagre" means containing a countable intersection of open dense subsets of this space; it is not a probability-almost-everywhere assertion.

Recall

$$a_k(E) = \sup\left\{\lambda(A) : \lambda\left(\bigcap_{j=0}^{k-1} R^{-jd}A\right) = 0 \text{ for every } d \in E\right\}, \qquad c(E) = \sup_{k\geq 2} a_k(E).$$

From our earlier arguments,

$$D \text{ is } \delta_0\text{-density large} \quad \Longleftrightarrow \quad c(D) \leq \delta_0, \qquad 0 < \delta_0 < 1. \tag{1}$$

For a neighborhood V of the identity in the strong Koopman topology, write

$$\mathcal{R}_\rho(V) = \{n \geq 1 : T_\rho^n \in V\}.$$

Theorem 1: Generic localization without density loss

Fix $D \subseteq \mathbb{N}$ with $c(D) < 1$. There is a dense G_δ set

$$\mathscr{G}_D \subseteq L^0(Y, G)$$

such that, for every $\rho \in \mathscr{G}_D$ and every strong identity neighborhood V,

$$\boxed{c\big(D \cap \mathcal{R}_\rho(V)\big) = c(D).} \tag{2}$$

Consequently, if D is δ_0-density large, then

$$\boxed{D \cap \mathcal{R}_\rho(V) \text{ is still } \delta_0\text{-density large for all systems.}} \tag{3}$$

Moreover, for every $\rho \in \mathscr{G}_D$, there is an increasing sequence

$$d_j \in D$$

such that

$$T_\rho^{d_j} \longrightarrow I \quad \text{strongly on } L^2(\mu). \tag{4}$$

In particular, for every measurable $A \subseteq X$ and every fixed $k \geq 2$,

$$\boxed{\mu\left(\bigcap_{\ell=0}^{k-1} T_\rho^{-\ell d_j} A\right) \longrightarrow \mu(A).} \tag{5}$$

Thus D is δ-density large for these systems for **every** $\delta \in (0,1)$, with a conclusion considerably stronger than positivity.

The quantifiers matter: ** D is fixed before the comeagre class $\mathscr{G}_D$ is chosen.**

2. Proof of the generic localization theorem

The argument combines our Bohr-localization theorem with finite recurrence certificates and approximation of cocycles by coboundaries.

2.1. A finite-certificate lemma

We need the following consequence of the finite extremal formulation.

Lemma. If

$$a_k(E) < \gamma,$$

then there is a finite $F \subseteq E$ such that

$$a_k(F) < \gamma. \tag{6}$$

Indeed, let $r_k(E,N)$ be the largest size of a subset of $[1,N]$ containing no k-term AP with difference in E. We established

$$a_k(E) = \inf_{N \geq 1} \frac{r_k(E,N)}{N}.$$

Choose N with

$$\frac{r_k(E,N)}{N} < \gamma,$$

and put $F = E \cap [1,N]$. Every relevant difference for a progression in $[1,N]$ belongs to this finite range, so

$$r_k(F,N) = r_k(E,N).$$

Consequently,

$$a_k(F) \leq \frac{r_k(F,N)}{N} < \gamma.$$

Thus a strict fixed-length density bound can be certified using finitely many differences.

2.2. Every cocycle cohomology class is dense

Here is the approximation fact we will use, including its proof.

Lemma. Fix a measurable cocycle $\rho_0 : Y \to G$. The cocycles

$$\rho_b(y) = b(Sy)\rho_0(y)b(y)^{-1}, \qquad b : Y \to G \text{ measurable}, \tag{7}$$

are dense in $L^0(Y, G)$.

In particular, coboundaries

$$\rho_b(y) = b(Sy)b(y)^{-1}$$

are dense.

Proof. Fix a target cocycle σ, a large integer q, and $\varepsilon > 0$. The Rokhlin lemma supplies a tower

$$B, SB, \dots, S^{q-1}B$$

whose levels are disjoint and whose union has measure greater than $1 - \varepsilon$. This is the standard probability-preserving form of the lemma. arXiv

Set $b = e_G$ on B, and recursively define, for $y \in B$,

$$b(S^{j+1}y) = \sigma(S^j y)\, b(S^j y)\, \rho_0(S^j y)^{-1}, \qquad 0 \le j < q - 1.$$

Define b arbitrarily off the tower. Then on every level except the top,

$$b(Sy)\rho_0(y)b(y)^{-1} = \sigma(y).$$

Therefore $\rho_b = \sigma$ outside a set of measure at most

$$\varepsilon + \frac{1}{q}.$$

Letting $\varepsilon \downarrow 0$ and $q \to \infty$ proves density. □

The corresponding change of fibre coordinates

$$J_b(y, zH) = (y, b(y)^{-1}zH)$$

conjugates T_{ρ_b} to T_{ρ_0}. In the coboundary case, it conjugates T_{ρ_b} to

$$S \times I_{G/H}. \tag{8}$$

2.3. Coboundaries already have the localization property

Suppose ρ is a coboundary. By (8), T_ρ has discrete spectrum, though it need not be ergodic.

For every strong identity neighborhood V, there is a Bohr neighborhood B of zero such that

$$n \in B \quad \Longrightarrow \quad T_\rho^n \in V. \tag{9}$$

To see this directly, a neighborhood V asks for finitely many functions to be approximately fixed in L^2. After applying the conjugacy, approximate those functions by finite sums of

$$\chi(y)\psi(zH),$$

where χ is a base eigenfunction. Requiring the finitely many corresponding eigenvalues to have their n-th powers close to 1 gives (9).

Now use our previously proved identity

$$c(D \cap B) = c(D) \qquad (B \text{ Bohr}_0). \tag{10}$$

Since

$$D \cap B \subseteq D \cap \mathcal{R}_\rho(V) \subseteq D,$$

monotonicity gives

$$c\big(D \cap \mathcal{R}_\rho(V)\big) = c(D). \tag{11}$$

So the desired property holds on a dense set of cocycles.

The remaining point is to express it through **open finite conditions**. Merely having a dense set would not suffice.

2.4. The Baire argument

Choose a dense sequence (f_i) in $L^2(X)$, and set

$$V_r = \left\{ R : \max_{1 \le i \le r} \|f_i \circ R - f_i\|_2 < \frac{1}{r} \right\}. \tag{12}$$

These form a neighborhood basis at the identity.

For each fixed n, the map

$$\rho \longmapsto T_\rho^n$$

is continuous from convergence in measure of cocycles to strong convergence of Koopman operators.

For $n = 1$, this follows first for functions $a(y)\psi(zH)$, with a bounded and ψ continuous, from uniform continuity of the compact group action. Density and unitarity then give the assertion for every L^2 function. Continuity of finite powers follows by telescoping.

Put

$$c = c(D), \qquad \gamma_s = c + \frac{1 - c}{s + 1},$$

so that

$$c < \gamma_s < 1, \qquad \gamma_s \downarrow c.$$

For $r, s \ge 1$ and $k \ge 2$, define

$$\mathcal{O}_{r,k,s} = \bigcup_{\substack{F \subseteq D \text{ finite} \\ a_k(F) < \gamma_s}} \bigcap_{d \in F} \{\rho : T_\rho^d \in V_r\}. \tag{13}$$

Each $\mathcal{O}_{r,k,s}$ is open. It is also dense: every coboundary belongs to it. Indeed, for a coboundary, (11) gives

$$c\big(D \cap \mathcal{R}_\rho(V_r)\big) = c < \gamma_s,$$

and the finite-certificate lemma supplies an appropriate finite

$$F \subseteq D \cap \mathcal{R}_\rho(V_r).$$

The space $L^0(Y, G)$ is completely metrizable; for example, a bounded complete metric on G induces the complete metric

$$d_0(\rho,\sigma)=\int_Y d_G(\rho(y),\sigma(y))\,d\nu(y).$$

Thus

$$\mathscr{G}_D=\bigcap_{r,k,s}\mathcal{O}_{r,k,s} \tag{14}$$

is dense G_δ.

Fix $\rho\in\mathscr{G}_D$, and write

$$E_r=D\cap\mathcal{R}_\rho(V_r).$$

For every k,s, membership in $\mathcal{O}_{r,k,s}$ provides a finite $F\subseteq E_r$ with

$$a_k(F)<\gamma_s.$$

Therefore

$$a_k(E_r)\le a_k(F)<\gamma_s.$$

Letting $s\to\infty$ gives

$$a_k(E_r)\le c.$$

Taking the supremum over k,

$$c(E_r)\le c.$$

But $E_r\subseteq D$ gives the reverse inequality, so

$$c(E_r)=c(D).$$

Finally, every strong identity neighborhood contains some V_r, and monotonicity proves (2) for arbitrary V. □

The finite certificates are essential: we transferred a fixed-length recurrence bound through an open condition, rather than trying to pass a possibly vanishing recurrence intersection through an uncontrolled limit.

3. Why this gives full-density recurrence

For $\rho\in\mathscr{G}_D$, each

$$E_r=D\cap\mathcal{R}_\rho(V_r)$$

has

$$c(E_r)=c(D)<1.$$

Such a set is unbounded. For instance, its Bohr-localizations $E_r\cap q\mathbb{N}$ still have critical density below 1, so they are nonempty for arbitrarily large q.

Choose increasingly

$$d_r\in E_r.$$

By (12) and density of (f_i),

$$T_\rho^{d_r} \to I \quad \text{strongly on } L^2.$$

For a measurable A, put

$$e_r = \mu(A \triangle T_\rho^{-d_r} A) \longrightarrow 0.$$

The triangle inequality and invariance give

$$\mu(A \triangle T_\rho^{-jd_r} A) \le j e_r.$$

Hence

$$\begin{aligned} 0 &\le \mu(A) - \mu\left(\bigcap_{j=0}^{k-1} T_\rho^{-jd_r} A\right) \\ &\le \sum_{j=1}^{k-1} \mu(A \triangle T_\rho^{-jd_r} A) \\ &\le \frac{k(k-1)}{2} e_r \longrightarrow 0. \end{aligned} \tag{15}$$

This proves (5).

The basic generic-rigidity conclusion can already be obtained from Bohr recurrence of D. **The stronger use of your fixed-density hypothesis is the exact preservation of $c(D)$ in (2): the selected near-identity times retain universal fixed-density regularity, not merely nonemptiness.**

4. The theorem reaches genuinely two-step, non-pro-nil systems

It is important to check that this result is not only detecting coboundaries, rotations, or systems already covered by your nilsystem theorem.

Specialize to

$$Sx = x + \alpha \quad \text{on } \mathbb{T}, \qquad \alpha \notin \mathbb{Q},$$

and circle-valued cocycles

$$T_\rho(x, z) = (x + \alpha, z + \rho(x)). \tag{16}$$

I claim that a dense G_δ set of these cocycles has Kronecker factor **exactly** the base $\sigma(x)$.

4.1. A dense class with no additional eigenfunctions

Decompose

$$L^2(\mathbb{T}^2) = \bigoplus_{q \in \mathbb{Z}} \mathcal{H}_q, \qquad \mathcal{H}_q = \{h(x)e(qz) : h \in L^2(\mathbb{T})\}.$$

On $\mathcal{H}_q$, the Koopman operator is identified with

$$V_{\rho,q} h(x) = e(q\rho(x)) h(x + \alpha). \tag{17}$$

Take

$$\rho_0(x) = x \pmod 1.$$

Then

$$V_{\rho_0,q} e(px) = e(p\alpha) e((p+q)x). \tag{18}$$

When $q \neq 0$, this operator has no eigenvectors: an eigenvector would have Fourier coefficients of constant magnitude along a nontrivial infinite progression of indices, contradicting square summability.

Every cocycle cohomologous to ρ_0 has the same property. Its conjugacy preserves each $\mathcal{H}_q$. By the tower argument above, this cohomology class is dense.

4.2. The no-extra-eigenfunctions condition is G_δ

For $q \neq 0$ and $p \in \mathbb{Z}$, define

$$A_N(\rho; q, p) = \frac{1}{N} \sum_{n=1}^{N} \left| \left\langle V_{\rho,q}^n e(px), e(px) \right\rangle \right|^2. \tag{19}$$

Every finite A_N is continuous in ρ.

If $\sigma_{\rho,q,p}$ is the spectral measure of $e(px)$, then

$$\lim_{N\to\infty} A_N(\rho; q, p) = \sum_{\lambda \in \mathbb{T}} \sigma_{\rho,q,p}(\{\lambda\})^2. \tag{20}$$

Indeed, express the average as a double integral of

$$\frac{1}{N} \sum_{n=1}^{N} (\lambda \overline{\zeta})^n$$

and use dominated convergence. This is the usual spectral-measure calculation. arXiv

Therefore $V_{\rho,q}$ has no eigenvectors precisely when the limits in (20) vanish for every p. If an eigenvector existed, it would have nonzero inner product with at least one Fourier basis vector.

The resulting condition is the G_δ set

$$\mathscr{W} = \bigcap_{\substack{q\neq 0,\, p\in\mathbb{Z} \\ r, M \geq 1}} \bigcup_{N \geq M} \{\rho : A_N(\rho; q, p) < 1/r\}. \tag{21}$$

It is dense because it contains the dense cohomology class of ρ_0.

For $\rho \in \mathscr{W}$, all eigenfunctions belong to $\mathcal{H}_0$, the base factor. Thus T_ρ is ergodic, its Kronecker factor is exactly $\sigma(x)$, and it has distal height exactly two.

Intersecting $\mathscr{W}$ with $\mathscr{G}_D$, we obtain a dense G_δ family of **ergodic, genuinely two-step distal systems rigid along a sequence inside** D.

4.3. These systems cannot be inverse limits of nilsystems

An ergodic nilsystem has the following property: for bounded f orthogonal to its Kronecker factor,

$$\langle T^n f, f \rangle \longrightarrow 0.$$

This is Theorem 1.6 of Ackelsberg–Richter–Shalom, following their discrete-plus-Lebesgue spectral decomposition. arXiv

It follows that an ergodic nilsystem which is rigid along an unbounded sequence must be a Kronecker system. Otherwise a nonzero bounded f orthogonal to the Kronecker factor would simultaneously satisfy

$$\langle T^{d_r} f, f\rangle \to \|f\|_2^2 \quad \text{and} \quad \langle T^{d_r} f, f\rangle \to 0.$$

Rigidity passes to factors. Consequently, if one of our systems were an inverse limit of nilsystems, every nilsystem factor would be Kronecker. Their inverse limit would then have discrete spectrum.

But for $\rho \in \mathscr{W}$, the nonzero spaces $\mathcal{H}_q$, $q \neq 0$, contain no eigenvectors. This is a contradiction.

We have therefore proved:

$$\boxed{\begin{array}{c}\text{For every fixed } D \text{ with } c(D) < 1, \text{ a comeagre family of} \\ \text{ergodic circle extensions of an irrational rotation} \\ \text{has exact distal height two, is not pro-nil,} \\ \text{and satisfies the desired recurrence conclusion at every density.}\end{array}} \tag{22}$$

This is the part that extends beyond the previously covered nilsystem and pro-nil classes.

5. Why approximation still does not prove the theorem for every cocycle

The generic theorem cannot be turned into the universal theorem just by taking limits.

For the skew shift

$$T_0(x,z) = (x+\alpha, z+x),$$

take $f(x,z) = e(z)$. Then

$$T_0^n f = e\left(z + nx + \binom{n}{2}\alpha\right),$$

so

$$\langle T_0^n f, f\rangle = 0 \qquad (n \geq 1).$$

Consequently,

$$\boxed{\|T_0^n f - f\|_2 = \sqrt{2} \qquad (n \geq 1).} \tag{23}$$

Thus T_0 has no rigidity sequence at all. Nevertheless, its cocycle can be approximated in measure by cocycles from $\mathscr{G}_D$, just as every cocycle can.

This does not make the skew shift a counterexample to your conjecture—it lies in the already-covered nilsystem class. It shows that **the stronger near-identity localization property is not closed**, and is not necessary for the desired multiple recurrence.

So a proof for all cocycles needs something beyond this Baire argument.

6. A further reduction for arbitrary circle cocycles

There is one more useful consequence of the generic theorem.

Because

$$L^0(\mathbb{T},\mathbb{T})$$

is an additive topological group, every circle cocycle ρ can be written

$$\boxed{\rho = \eta + \xi} \tag{24}$$

where **both** η and ξ belong to

$$\mathscr{G}_D \cap \mathscr{W}.$$

Indeed, both

$$\mathscr{G}_D \cap \mathscr{W} \quad \text{and} \quad \rho - (\mathscr{G}_D \cap \mathscr{W})$$

are comeagre, so they intersect.

Now form the relatively independent joining over the common base:

$$R(x, z_1, z_2) = (x + \alpha, z_1 + \eta(x), z_2 + \xi(x)). \tag{25}$$

The map

$$\Pi(x, z_1, z_2) = (x, z_1 + z_2)$$

is a measure-preserving factor map from R to T_ρ.

Therefore:

> **Every circle extension of an irrational rotation is a factor of a relatively independent joining of two exact two-step systems for which we have already proved the full conclusion—and even rigidity along D.**

This gives a specific remaining joining problem:

$$\boxed{\begin{array}{c}\text{Can the all-systems fixed-density hypothesis force}\\ \text{multiple recurrence along } D \text{ in the joining } (25),\\ \text{when its two component cocycles belong to } \mathscr{G}_D?\end{array}} \tag{26}$$

Individual rigidity sequences need not coincide. Nor do the separate identities

$$c(D \cap \mathcal{R}_\eta(V)) = c(D), \qquad c(D \cap \mathcal{R}_\xi(W)) = c(D)$$

justify the corresponding assertion for the intersection of these two sets of times. In fact, demanding simultaneous rigidity would be too strong: by taking $\rho(x) = x$, it would force rigidity of the skew-shift factor, contradicting (23).

The unrestricted two-step theorem remains unproved here. The new result removes the density-threshold restriction for a comeagre family with genuinely nonatomic, non-pro-nil dynamics. For arbitrary circle cocycles, it also replaces the general cocycle problem by the particular relative-joining problem (26), where both marginal systems already satisfy a much stronger recurrence conclusion.