PERIODIC COUNTEREXAMPLES TO THREE CONJECTURES ON KOLAKOSKI TRANSFORMS

WILL HOWE, GRANT HUFFMAN, JENSEN KOHLMEYER, AND LIAM KRUER

Abstract. The simple Kolakoski transform need not preserve the limiting mean of a periodic word on $\{1,2\}$. We give a counterexample with input period 24 and output period 12, and a balanced counterexample whose input has mean 3/2 while its output has mean 13/9. Both examples remain unchanged under every realization of a nonconstant summable perturbation of the weights. Their discrepancies from the input means are linear. Consequently they contradict Conjectures (1), (2), and (4) of Cloitre. These statements concern auxiliary transforms of periodic inputs; the classical Kolakoski frequency problem remains open.

hexagon:2611.00312v1

1. Definitions and scope

We use positive integer indexing. For a word $a = (a_i)_{i\ge1}$, its mean, when it exists, is

$$m(a) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} a_i.$$

Cloitre introduced the following transforms in [1, pp. 2–3]. For an input w on $\{1,2\}$ and positive weights $u = (u_i)_{i\ge1}$, write

$$F_{x,u}(0) = 0, \qquad F_{x,u}(i) = \sum_{j=1}^{i} u_j x_j, \qquad G_{x,u}(n) = \min\{i \ge 1 : F_{x,u}(i) \ge n\}.$$

With the initialization used here, a weighted transform is a word x satisfying

$$x_1 = 1, \quad x_2 = 2, \qquad x_n = w_{G_{x,u}(n)} \quad (n \ge 3). \tag{1.1}$$

The simple transform has $u_i = 1$ for every i. In that case write F_x and G_x for the two cumulative functions.

All weights in this paper satisfy $u_i \ge 1$. Consequently (1.1) is a causal recursion and determines x uniquely. Indeed, for $n \ge 3$ its first $n-1$ terms give

$$\sum_{j=1}^{n-1} u_j x_j \ge \sum_{j=1}^{n-1} x_j \ge n,$$

since $x_2 = 2$ and every other letter is at least one. Thus the least index in (1.1) is found among the already determined positions. This also proves existence by induction.

The three assertions addressed here are the following statements of [1, §§1, 2.2, 2.4]:

(i) The simple transform preserves every existing input mean.
(ii) The same mean preservation holds for random weights which are $\delta(i)$ with probability $1/f_\alpha(i)$ and one otherwise, where δ is bounded and positive, $0 < \alpha < 1$, and f_α is positive and regularly varying with index α.

Date: November 7, 2026.
Prepared with AI assistance from OpenAI Codex in mathematical exploration, checking, and manuscript preparation.

(iii) If $m(w) = \lambda$, the simple transform satisfies

$$\sum_{j=1}^{n} x_j = \lambda n + O(n^{1/2+\varepsilon}). \tag{1.2}$$

They are respectively Conjectures (1), (2), and (4). We disprove the last estimate for every fixed $0 < \varepsilon < 1/2$. For the random assertion, our counterexamples fail in every realization, so the interpretation of its probabilistic mode of convergence is immaterial.

2. A short periodic certificate

The following observation turns a finite block identity into an infinite transform identity. A block is a consecutive finite string; adjacent blocks may have the same letter.

Lemma 2.1 (Block criterion)**.** *Let x and w be words on $\{1, 2\}$, with $x_1 = 1$ and $x_2 = 2$. If for every $i \geq 1$ the positions*

$$F_x(i-1)+1, \ldots, F_x(i)$$

all carry the letter w_i, then x is the simple transform of w. Suppose, moreover, that x has period p, w has period q, $p \mid q$, and $p \mid F_x(q)$. Then it suffices to check these blocks for $1 \leq i \leq q$.

Proof. The blocks partition the positive integer positions. The block containing n has index $G_x(n)$, giving the required identity, even for $n = 1, 2$. For the periodic assertion, write $i = aq + r$ with $a \geq 0$ and $1 \leq r \leq q$. Since q is a period of x,

$$F_x(aq + r) = aF_x(q) + F_x(r).$$

Translation by $aF_x(q)$ preserves x, and translation by aq preserves w. The ith block therefore has the same color as the rth block. □

Theorem 2.2. *Let*

$$x = (122221122111)^{\infty}, \qquad w = (122121112221121122212111)^{\infty}.$$

Then x is the simple transform of w, and

$$m(w) = \frac{35}{24}, \qquad m(x) = \frac{3}{2}.$$

In particular, simple transforms do not preserve the means of periodic binary inputs.

Proof. For $i = 1, \ldots, 24$, the blocks formed by repeating w_i exactly x_i times are, in order,

$$\begin{array}{cccccccccccc} 1 & 22 & 22 & 11 & 22 & 1 & 1 & 11 & 22 & 2 & 2 & 1 \\ 1 & 22 & 11 & 11 & 22 & 2 & 2 & 11 & 22 & 1 & 1 & 1. \end{array}$$

Their concatenation is exactly $(122221122111)^3$. Here $p = 12$, $q = 24$, and $F_x(q) = 36$, so Lemma 2.1 applies. The displayed output period has six twos, while the input period has eleven. The respective period sums are therefore 18 and 35, which give the stated means. □

Lemma 2.3 (Periodic discrepancy)**.** *If a bounded real word a has period p, then*

$$\sum_{j=1}^{n} a_j = m(a)n + O(1), \qquad m(a) = \frac{1}{p}\sum_{j=1}^{p} a_j.$$

Proof. Write $n = kp + r$, where $0 \leq r < p$. The difference on the left is $\sum_{j=1}^{r} a_j - m(a)r$, one of finitely many values. □

Corollary 2.4. *For the input in Theorem 2.2,*

$$\sum_{j=1}^{n} x_j - \frac{35}{24}n = \frac{n}{24} + O(1).$$

Thus (1.2) *fails whenever* $0 < \varepsilon < 1/2$.

Proof. Apply Lemma 2.3. Dividing by n yields a nonzero limit $1/24$, whereas an $O(n^{1/2+\varepsilon})$ discrepancy would give limit zero. □

3. A balanced input

The discrepancy can also occur when the input has equal frequencies of one and two. Define the periodic binary word y by

$$y = (101010101001101001101001000)^{\infty}. \tag{3.1}$$

Set

$$x_1 = 1, \qquad x_{2i} = x_{2i+1} = 1 + y_i \quad (i \geq 1). \tag{3.2}$$

In this section x denotes this new word.

Theorem 3.1. *The word* (3.2) *is the simple transform of a periodic word* w *of period* 486. *This input has* 243 *twos in one period. Consequently*

$$m(w) = \frac{3}{2}, \qquad m(x) = \frac{13}{9}, \qquad \sum_{j=1}^{n} x_j - \frac{3}{2}n = -\frac{n}{18} + O(1).$$

Proof. Put

$$t_i = \sum_{j<i}(1 + y_j), \qquad F_x(2i-1) = 1 + 2t_i.$$

After its first singleton, the proposed source lengths occur in equal pairs. The two source blocks corresponding to $2i$ and $2i+1$ together cover the $1 + y_i$ target pairs beginning at positions

$$2(t_i + 1),\ 2(t_i + 2),\ \ldots,\ 2(t_i + 1 + y_i).$$

Every target pair has equal letters by (3.2). If $y_i = 1$, each source block is a whole target pair. If $y_i = 0$, the two source blocks are the two singleton halves of one target pair. All blocks are therefore monochromatic. Let w_i be the color of the ith block. This defines the input and proves the transform identity by Lemma 2.1.

The period in (3.1) contains twelve ones. Since $y_{27} = 0$, the word x is periodic from its first position, with period 54. That period has 24 twos and sum 78. In particular,

$$F_x(i + 486) = F_x(i) + 702, \qquad 702 = 13 \cdot 54.$$

Translating a source block by 486 thus translates its target block by a period of x. Hence w has period 486.

It remains to count the twos in this input period. Count the 243 pairs w_{2i}, w_{2i+1} for $1 \leq i \leq 243$; these are also 486 consecutive input letters. If $y_i = 1$, their two colors read the target bits y_{t_i+1}, y_{t_i+2} once each. If $y_i = 0$, they read y_{t_i+1} twice. Thus every target bit from 1 through

$$t_{244} = 9\sum_{i=1}^{27}(1 + y_i) = 9 \cdot 39 = 351$$

is read once, with one additional reading of y_{t_i+1} for each zero source bit. The first readings contribute $13 \cdot 12 = 156$ twos.

For the additional readings, let

$$A_r = \#\{1 \le j \le 27 : y_j = 1,\ j - 1 \equiv r \pmod 3\},$$
$$B_r = \#\{1 \le i \le 27 : y_i = 0,\ t_i \equiv r \pmod 3\}.$$

The period (3.1) gives the following small table:

r	0	1	2
A_r	4	1	7
B_r	4	1	10

For completeness, the t_i at zero positions i in this first period are

$$2, 5, 8, 11, 14, 15, 20, 23,$$
$$24, 29, 32, 33, 36, 37, 38,$$

which verify the second row. For a fixed source position i, moving to $i + 27k$ adds $39k$ to t_i. The nine residues $t_i + 39k$ modulo 27, $0 \le k \le 8$, are exactly the residue class of t_i modulo 3, since $\gcd(39, 27) = 3$. Its additional readings therefore contribute A_r twos when $t_i \equiv r \pmod 3$. The total number of additional twos is

$$\sum_{r=0}^{2} B_r A_r = 4 \cdot 4 + 1 \cdot 1 + 10 \cdot 7 = 87.$$

There are $156 + 87 = 243$ input twos, as asserted. The two means and the linear discrepancy now follow from the period sums and Lemma 2.3. □

4. Weights that leave every cut unchanged

Lemma 4.1 (Stability of integer cuts)**.** *Let x be a word on $\{1, 2\}$, and suppose*

$$u_i \ge 1, \qquad \sum_{i=1}^{\infty} 2(u_i - 1) < 1.$$

Then for every pair of positive integers i, n,

$$F_{x,u}(i) \ge n \quad \Longleftrightarrow \quad F_x(i) \ge n.$$

In particular $G_{x,u}(n) = G_x(n)$ for all n.

Proof. For each i,

$$0 \le F_{x,u}(i) - F_x(i) = \sum_{j=1}^{i} x_j(u_j - 1) < 1.$$

Since $F_x(i)$ and n are integers, this perturbation cannot change the comparison with n. □

Theorem 4.2. *For either input of Theorems 2.2 and 3.1, put*

$$\delta(i) = 1 + 2^{-i-2}.$$

For every choice of weights $u_i \in \{1, \delta(i)\}$, the weighted transform is the same word x as the simple transform. In particular, the random weights in Conjecture (2) do not preserve the input mean.

Proof. The function δ is positive, bounded, and nonconstant. For every choice of its permitted weights,

$$0 \le \sum_{i=1}^{\infty} 2(u_i - 1) \le \sum_{i=1}^{\infty} 2^{-i-1} = \frac{1}{2} < 1.$$

Lemma 4.1 shows that all cut indices are unchanged. Thus x satisfies (1.1) for these weights. Uniqueness of the causal recursion identifies it with the weighted transform.

For a concrete admissible probability law, fix any $0 < \alpha < 1$ and set $f_\alpha(t) = t^\alpha$ for $t \geq 1$. It is positive and regularly varying of index α, and $1/f_\alpha(i)$ is a probability. Give u_i the value $\delta(i)$ with that probability and one otherwise. Every resulting weight sequence meets the preceding deterministic bound; no independence assumption is needed. The output mean is therefore $3/2$ instead of $35/24$ for the short example, or $13/9$ instead of $3/2$ for the balanced example, in every realization. □

5. The classical question

The periodic examples disprove the three universal assertions stated in Section 1. They do not replace the alternating input $121212\cdots$ by itself: that input gives the classical Kolakoski word. Nothing in the argument determines its limiting frequency of either letter. Nor do the counterexamples address the specifically prescribed alphabet perturbations in Cloitre's Conjectures (6)–(8). The classical frequency problem remains open.

References

[1] Benoit Cloitre, *The Kolakoski transform of words*, manuscript, October 6, 2013, 27 pp., `https://oeis.org/A000002/a000002.pdf`.

Email address: `howew@purdue.edu`

Email address: `huffma97@purdue.edu`

Email address: `jenwin@purdue.edu`

Email address: `kruerl@purdue.edu`