Finite-memory Collatz parity dynamics: Fibonacci support, Bernoulli branch drift, and the pointwise obstruction

William Goodfellow*

May 4, 2026

Abstract

We study finite-memory parity histories for the odd-to-odd shortcut maps

$$S_a(n) = \frac{an+1}{2^{\nu_2(an+1)}}$$

with a odd, focusing on the Collatz case $a = 3$. Sampling raw parity histories at odd arrivals collapses the apparent 2^m-state history space to a Fibonacci recurrent class

$$\Omega_m = \{h \in \{0,1\}^m : h_0 = 1,\ h_i h_{i+1} = 0\}, \qquad |\Omega_m| = F_m,$$

where F_m is the m-th Fibonacci number. Under normalized Haar branch measure on odd 2-adic integers, the next branch exponent $K_a = \nu_2(an+1)$ is independent of every finite history state $h \in \Omega_m$ and satisfies $\mathbb{P}(K_a = k) = 2^{-k}$. Consequently, the conditional branch drift is uniform on Ω_m:

$$\mathbb{E}[\log_2 a - K_a \mid h] = \log_2 a - 2.$$

Thus the $(3n+1)$ shortcut has uniform negative finite-memory drift $\log_2 3 - 2 \approx -0.415$, while the $(5n+1)$ control has uniform positive drift $\log_2 5 - 2 \approx +0.322$. We also derive exact stationary weights on Ω_m, verify the laws numerically on integer trajectories, and identify a finite-height survival bias in descending $(3n+1)$ samples that is absent in the growing $(5n+1)$ control. Finally, we show that no finite-history correction can turn this supermartingale into a pointwise one-step Lyapunov function: the alternating state has a $K = 1$ self-transition with positive log increment $\log_2 3 - 1 > 0$, so any potential $J(n,h) = \log_2 n + g(h)$ fails on this transition. The result isolates the finite-memory stochastic structure of Collatz parity dynamics while leaving the pointwise Collatz conjecture untouched.

1 Introduction

The Collatz conjecture asks whether every positive integer eventually reaches 1 under the iteration $T(n) = n/2$ if n is even and $T(n) = 3n + 1$ if n is odd. The conjecture is widely believed and has been verified computationally to $n \leq 2^{71}$ [6], but a proof has eluded all approaches.

A standard heuristic argument suggests the iteration should descend on average: per odd-to-odd transition, the iterated state is multiplied by 3 and divided by 2 a geometrically-distributed number of times, giving an average log-multiplier of $\log_2 3 - 2 \approx -0.415 < 0$. This heuristic underlies almost

*Independent researcher. Contact: werner85@gmail.com.

all rigorous partial results, from Terras' density-1 stopping-time theorem [9] through Krasikov-Lagarias' $x^{0.84}$ density bound [3] to Tao's recent log-density-1 almost-bounded theorem [8]. Yet the heuristic is not, by itself, a proof: negative average drift does not exclude thin pointwise-divergent trajectories or non-trivial cycles.

The present paper isolates a clean finite-memory layer of this story. We work with the odd-to-odd shortcut map $S(n) = (3n+1)/2^{\nu_2(3n+1)}$ and consider the rolling m-bit parity history of the underlying ordinary Collatz iteration, sampled at odd arrivals. We prove three results:

(1) **Fibonacci support** (Theorem 7): the reachable history set is exactly $\Omega_m = \{h \in \{0,1\}^m : h_0 = 1,\ h_i h_{i+1} = 0\}$, of size F_m.

(2) **Bernoulli branch independence** (Theorem 8): under the natural 2-adic Haar branch measure, the next branch exponent is Geom(1/2)-distributed and independent of every finite history state.

(3) **Uniform conditional drift** (Corollary 9): $\mathbb{E}[\log_2 a - K_a \mid h] = \log_2 a - 2$ for every $h \in \Omega_m$, giving a uniform finite-memory supermartingale structure on $(3n+1)$ and a uniform positive-drift analogue for $(5n+1)$.

We also derive exact stationary weights $\pi_m(h)$ on Ω_m (Theorem 11), and show that no finite-history correction (bounded or unbounded) can convert the uniform conditional drift into pointwise one-step descent (Corollary 13). The latter is structural rather than negative: it precisely localizes the boundary between finite-memory probabilistic descent and pointwise Lyapunov descent.

These results are proved under the natural Bernoulli branch measure, and their integer-trajectory recovery is verified by numerical experiments on 80-bit and 200-bit starts. A finite-height survival bias is identified for descending $(3n+1)$ trajectories and shown to vanish on the growing $(5n+1)$ control.

Scope and what this paper does not claim. This paper does *not* prove the pointwise Collatz conjecture, nor does it claim to bridge Tao's almost-bounded theorem to pointwise convergence. The finite-memory supermartingale we identify is exact under the Bernoulli branch measure but, as Corollary 13 shows, cannot be promoted to a pointwise descent theorem by any finite-history correction. The pointwise Collatz residue lies in the possibility of non-generic long-range branch-exponent sequences that are visible to neither the finite history automaton nor any history-only correction.

2 Definitions

Definition 1 (Odd-to-odd shortcut)**.** Let a be a positive odd integer. The $(an+1)$ *odd-to-odd shortcut map* is

$$S_a : \{n \in \mathbb{N} : n \text{ odd}\} \to \{n \in \mathbb{N} : n \text{ odd}\}, \qquad S_a(n) = \frac{an+1}{2^{K_a(n)}}, \qquad K_a(n) = \nu_2(an+1),$$

where ν_2 is the 2-adic valuation. The original Collatz shortcut corresponds to $a = 3$. We write $S = S_3$ and $K = K_3$ when the coefficient is unambiguous.

Remark 2. Since a is odd, $an+1$ is even whenever n is odd, so $K_a(n) \geq 1$ and $S_a(n)$ is well-defined and odd.

Definition 3 (Raw parity sequence and history window)**.** Let T denote the ordinary (non-shortcut) Collatz map, $T(n) = n/2$ if n even, $T(n) = an + 1$ if n odd. The *raw parity sequence* of a starting value n_0 is the sequence $\sigma_t = n_t \bmod 2$, where $n_t = T^t(n_0)$. The *m-bit raw parity history at time t* is the rolling window

$$h_t = (\sigma_{t-m+1}, \sigma_{t-m+2}, \ldots, \sigma_t) \in \{0,1\}^m,$$

which we encode as an integer $h_t \in \{0, \ldots, 2^m - 1\}$ with σ_t as the lowest-order bit.

Definition 4 (Sampling at odd arrivals)**.** We sample h_t at the times t such that n_t is odd. Each odd arrival n_t is followed by an excursion through one or more even states $T(n_t), T^2(n_t), \ldots$ ending at the next odd state $S_a(n_t)$. The *shortcut step* from n_t to $S_a(n_t)$ updates the history h_t by appending the parities of the intermediate states, then the parity of $S_a(n_t)$.

Definition 5 (Fibonacci recurrent set)**.** For $m \geq 1$ define

$$\Omega_m = \big\{ h \in \{0,1\}^m : h_0 = 1,\ \ h_i h_{i+1} = 0 \text{ for all } i = 0, \ldots, m-2 \big\}.$$

Equivalently, Ω_m is the set of m-bit patterns whose lowest bit is 1 and which contain no two adjacent 1s.

3 Fibonacci support

Lemma 6 (Finite branch words are realized)**.** *Let a be odd and let $k_1, \ldots, k_r \geq 1$ be any finite branch word. Then the cylinder*

$$C(k_1, \ldots, k_r) = \{ n \in \mathbb{Z}_2^\times : K_a(S_a^{j-1}(n)) = k_j,\ j = 1, \ldots, r \}$$

has positive normalized Haar measure

$$\mu(C(k_1, \ldots, k_r)) = \prod_{j=1}^{r} 2^{-k_j}.$$

In particular, the cylinder is a nonempty compact-open subset of $\mathbb{Z}_2^\times$ and contains arbitrarily large positive odd integers.

Proof. For $k \geq 1$, define the inverse branch map

$$L_k : \mathbb{Z}_2^\times \to \mathbb{Z}_2^\times, \qquad L_k(y) = \frac{2^k y - 1}{a}.$$

Since a is odd, division by a is a homeomorphism of $\mathbb{Z}_2$. The map L_k sends $\mathbb{Z}_2^\times$ homeomorphically onto the branch

$$B_k = \{ n \in \mathbb{Z}_2^\times : K_a(n) = k \},$$

and $S_a \circ L_k = \mathrm{id}$.

For a finite branch word $k_1, \ldots, k_r$, the cylinder is

$$C(k_1, \ldots, k_r) = L_{k_1} \circ L_{k_2} \circ \cdots \circ L_{k_r}(\mathbb{Z}_2^\times).$$

If $A \subseteq \mathbb{Z}_2^\times$ is compact-open, then $L_k(A)$ occupies the corresponding residue classes modulo an additional k powers of 2, so $\mu(L_k(A)) = 2^{-k}\mu(A)$. By induction,

$$\mu(C(k_1, \ldots, k_r)) = \prod_{j=1}^{r} 2^{-k_j}.$$

The cylinder is a nonempty compact-open subset of $\mathbb{Z}_2^\times$, hence contains infinitely many, indeed arbitrarily large, positive odd integers (since $\mathbb{N} \cap \{\text{odd}\}$ is dense in $\mathbb{Z}_2^\times$). □

Theorem 7 (Fibonacci support)**.** *Let a be odd. Sampling the m-bit raw parity history at odd arrivals of the orbit of any odd starting value $n_0 \in \mathbb{N}$ produces histories $h \in \Omega_m$. Conversely, every $h \in \Omega_m$ arises as such a history for some odd starting value. Consequently the reachable history set has size*

$$|\Omega_m| = F_m,$$

where F_m is the m-th Fibonacci number with the convention $F_1 = 1, F_2 = 1, F_3 = 2, \dots, F_8 = 21, F_{12} = 144$.

Proof. Forward direction. At an odd arrival n_t, the sampled history h_t has lowest bit $\sigma_t = n_t \bmod 2 = 1$. Suppose for contradiction that h_t contains adjacent 1s in positions $i, i+1$ for some $0 \le i \le m-2$. This corresponds to two consecutive states n_{t-i-1}, n_{t-i} in the orbit being odd. But for a odd, $T(n_{t-i-1}) = an_{t-i-1} + 1$ is even, so $n_{t-i} = T(n_{t-i-1})$ is even and $\sigma_{t-i} = 0$. Contradiction. Hence $h_t \in \Omega_m$.

Reverse direction. Given $h \in \Omega_m$, let $0 = p_0 < p_1 < \cdots < p_r \le m-1$ be the positions of its 1s. Define the visible branch gaps

$$K_j = p_j - p_{j-1} - 1, \qquad j = 1, \dots, r,$$

and choose any older gap

$$K_{r+1} \ge \max(m - p_r - 1, 1).$$

By Lemma 6, there are positive odd integers realizing this finite branch word, and the associated raw parity window at the corresponding odd arrival is exactly h.

Cardinality. The condition $h_0 = 1$ fixes the lowest bit. The remaining $m-1$ bits at positions $1, \dots, m-1$ form a binary string with $h_1 = 0$ (forced by $h_0 h_1 = 0$) and no adjacent 1s among positions $1, \dots, m-1$. Equivalently, we count subsets of $\{2, 3, \dots, m-1\}$ with no two consecutive integers. The number of such subsets is the Fibonacci number F_m (using the standard combinatorial identity $\#\{S \subseteq [n] : \text{no two consecutive}\} = F_{n+2}$). For $m = 8$, $|\Omega_8| = F_8 = 21$. ☐

4 Bernoulli branch theorem

We prove the inline branch law directly. The result is also a consequence of the broader 2-adic conjugacy of the $3x+1$ map established by Bernstein and Lagarias [1], who showed that the $(3x+1)$ map on $\mathbb{Z}_2$ is measure-preserving, strongly mixing, and metrically conjugate to the 2-adic shift, hence Bernoulli.

Theorem 8 (Bernoulli branch independence)**.** *Let a be odd. Under the normalized Haar measure μ on the odd 2-adic integers $\mathbb{Z}_2^\times$, the branch exponent $K_a(n) = \nu_2(an+1)$ is geometrically distributed with parameter $1/2$:*

$$\mu\big(\{n \in \mathbb{Z}_2^\times : K_a(n) = k\}\big) = 2^{-k}, \qquad k \ge 1.$$

Moreover, the branch map $S_a(n) = (an+1)/2^{K_a(n)}$ acts as a Bernoulli shift on $\mathbb{Z}_2^\times$ with respect to the branch sequence: the successive branch exponents $K_a(n), K_a(S_a(n)), K_a(S_a^2(n)), \dots$ are independent and identically Geom(1/2)*-distributed under μ. In particular, for any history state $h \in \Omega_m$ generated by a finite past initial segment of the branch sequence, the next branch exponent is independent of h.*

Proof. Branch law. Fix $k \ge 1$. The condition $K_a(n) = k$ is equivalent to $an + 1 \equiv 2^k \pmod{2^{k+1}}$, i.e.,

$$n \equiv \frac{2^k - 1}{a} \pmod{2^{k+1}},$$

where the inverse a^{-1} exists in $\mathbb{Z}/2^{k+1}\mathbb{Z}$ because a is odd. So $\{n \in \mathbb{Z}_2^\times : K_a(n) = k\}$ is a single residue class modulo 2^{k+1} intersected with the odd residues mod 2. The odd residues form a measure-$\frac{1}{2}$ subset of $\mathbb{Z}/2^{k+1}\mathbb{Z}$, and the further condition pins down a single one. Hence

$$\mu\big(\{K_a = k\}\big) = \frac{1}{2^k}.$$

Independence via cylinder identity. The inverse-branch maps L_k from Lemma 6 satisfy $\mu(L_k(A)) = 2^{-k}\mu(A)$ for every compact-open $A \subseteq \mathbb{Z}_2^\times$. Equivalently,

$$\mu(B_k \cap S_a^{-1}A) = 2^{-k}\mu(A).$$

Iterating, for every finite branch word $k_1, \dots, k_r$ and every Borel A,

$$\mu(C(k_1, \dots, k_r) \cap S_a^{-r}A) = \left(\prod_{j=1}^{r} 2^{-k_j}\right)\mu(A).$$

This identity says the successive branch exponents under μ are iid with $\Pr(K_a = k) = 2^{-k}$. Since any finite history state $h \in \Omega_m$ is determined by finitely many past branch exponents, the next exponent is independent of h.

Bernstein-Lagarias context. For $a = 3$, this branch independence is part of the broader 2-adic conjugacy of the $3x + 1$ map established in [1]: the map on $\mathbb{Z}_2$ is measure-preserving, strongly mixing, and metrically conjugate to the 2-adic shift, hence Bernoulli. The cylinder argument above gives the branch-law statement directly for all odd a without invoking that conjugacy. □

Corollary 9 (Uniform conditional drift)**.** *Let a be odd. Under the natural Bernoulli branch measure on $\mathbb{Z}_2^\times$, for every $h \in \Omega_m$:*

$$\mathbb{E}\big[\log_2 a - K_a \;\big|\; h\big] = \log_2 a - 2.$$

For integer trajectories, the asymptotic shortcut log-step is

$$\Delta \log_2 n_t = \log_2 a + \log_2(1 + 1/(an_t)) - K_a(n_t) = \log_2 a - K_a(n_t) + O(1/n_t),$$

so the branch-measure conditional drift is exact for the symbolic step $\log_2 a - K_a$ and asymptotically exact for $\Delta \log_2 n_t$.

Proof. By Theorem 8, K_a given h is Geom(1/2)-distributed independently of h, with $\mathbb{E}[K_a] = \sum_{k\geq 1} k \cdot 2^{-k} = 2$. Hence $\mathbb{E}[\log_2 a - K_a \mid h] = \log_2 a - 2$. □

Example 10 ($a = 3$ vs $a = 5$)**.** For the Collatz coefficient $a = 3$:

$$\log_2 3 - 2 \approx -0.41504 < 0,$$

giving uniform negative conditional drift on every $h \in \Omega_m$. For the $(5n + 1)$ analogue $a = 5$:

$$\log_2 5 - 2 \approx +0.32193 > 0,$$

giving uniform positive conditional drift on every $h \in \Omega_m$. The branch law and Fibonacci support are identical in both cases; the drift sign is determined entirely by the coefficient. This is the cleanest available coefficient-specificity statement and shows that any proof of finite-memory descent must use $\log_2 a < 2$ (equivalently, $a < 4$).

5 Exact stationary weights

Theorem 11 (Exact stationary weights). *Let a be odd. Under the Bernoulli branch measure on $\mathbb{Z}_2^\times$, the stationary distribution π_m on Ω_m induced by the shortcut map S_a is given as follows. For $h \in \Omega_m$ with 1-positions $0 = p_0 < p_1 < \cdots < p_r \leq m-1$, set $K_j = p_j - p_{j-1} - 1 \geq 1$ for $j = 1, \ldots, r$ (the recent branch exponents recoverable from h). Then*

$$\pi_m(h) = \left(\prod_{j=1}^{r} 2^{-K_j}\right) \cdot \mathbb{P}\big(K_a \geq m - p_r - 1\big),$$

where $\mathbb{P}(K_a \geq L) = 2^{-(L-1)}$ for $L \geq 1$ and $\mathbb{P}(K_a \geq L) = 1$ for $L \leq 0$. Equivalently,

$$\pi_m(h) = 2^{-\sum_{j=1}^{r} K_j} \cdot 2^{-\max(m-p_r-2,0)}.$$

The total mass on Ω_m sums to 1:

$$\sum_{h \in \Omega_m} \pi_m(h) = 1.$$

Proof. The history h records the positions of past odd arrivals within the rolling m-bit window. The r visible branch exponents $K_1, \ldots, K_r$ are recoverable from h as $K_j = p_j - p_{j-1} - 1$. The unseen older exponent K_{r+1} must satisfy $K_{r+1} \geq m - p_r - 1$ to keep the next-older odd arrival outside the window. By Theorem 8, the K_j are iid Geom(1/2), so

$$\pi_m(h) = \prod_{j=1}^{r} \mathbb{P}(K_a = K_j) \cdot \mathbb{P}(K_a \geq m - p_r - 1) = \prod_{j=1}^{r} 2^{-K_j} \cdot \mathbb{P}(K_a \geq m - p_r - 1).$$

For total mass: the events indexed by $h \in \Omega_m$ partition the backward renewal process according to which odd arrivals fall inside the length-m window and where the first older arrival outside the window occurs. Since the branch gaps are iid Geom(1/2) and almost surely finite, these events have total probability 1. □

Example 12. For $m = 8$ and $h = 01010101$, the 1-positions are $0, 2, 4, 6$, giving $K_1 = K_2 = K_3 = 1$ and the unseen exponent $K_4 \geq 1$. Hence $\pi_8(01010101) = 2^{-3} \cdot 1 = 1/8 = 0.125$. For $h = 00000001$, the only 1-position is 0, giving no recent exponents and $K_1 \geq 7$. Hence $\pi_8(00000001) = 2^{-6} = 1/64 \approx 0.0156$.

6 No finite-history pointwise Lyapunov correction

Corollary 13 (No finite-history Lyapunov correction). *Let $a = 3$ and $m \geq 2$. There does not exist any function $g : \Omega_m \to \mathbb{R}$ such that the augmented potential*

$$J(n, h) = \log_2 n + g(h)$$

satisfies $J(S_3(n), h') < J(n, h)$ for every odd n and every history transition $h \to h'$ generated by the orbit. The obstruction holds whether or not g is bounded.

Proof. Define the alternating history $h \in \Omega_m$ by $h_i = 1$ for even i and $h_i = 0$ for odd i (for $m = 8$, $h =$`01010101`; for general m, h has 1s exactly at the even positions).

By Lemma 6, there are arbitrarily large positive odd integers whose current history is h and whose next branch exponent is $K = 1$. For such n, the shortcut step appends two parities to the rolling window: a 0 (from $T(n) = 3n + 1$, even) followed by a 1 (from $S_3(n) = (3n + 1)/2$, odd, since $K = 1$). Left-shifting the window by two bits and appending `01` preserves the alternating pattern: $h' = h$.

The corresponding log-step is

$$\Delta \log_2 n = \log_2((3n + 1)/2) - \log_2 n = \log_2(3 + 1/n) - 1 = \log_2 3 - 1 + O(1/n).$$

For sufficiently large n, $\Delta \log_2 n \geq (\log_2 3 - 1)/2 > 0$. For any candidate g,

$$J(S_3(n), h') - J(n, h) = \Delta \log_2 n + (g(h') - g(h)) = \Delta \log_2 n + 0 > 0,$$

since $h' = h$ forces $g(h') - g(h) = 0$. Hence J fails to be a strict pointwise Lyapunov on this transition. The obstruction depends only on $h' = h$ on this self-loop, not on any boundedness or regularity of g. □

Remark 14 (Structural interpretation)**.** Corollary 13 is not a negative result about Collatz; it precisely localizes the boundary between the finite-memory probabilistic descent established in Corollary 9 and the pointwise Lyapunov descent that would resolve Collatz. The uniform conditional drift theorem gives a supermartingale at the level of finite-history measure, but the alternating self-loop with $K = 1$ exhibits a deterministic obstruction to promoting any history-only correction to pointwise descent. The pointwise Collatz residue lies in non-generic long-range branch-exponent sequences invisible to any finite history window.

7 Numerical experiments

We verify the theorems on integer trajectories and identify a finite-height survival bias.

7.1 Setup

For each experiment, we draw N random odd starting values uniformly from $[1, 2^B]$, run a burn-in of β shortcut steps, and then sample σ shortcut transitions per trajectory. We track the rolling 8-bit raw parity history h and the branch exponent K at each transition. Trajectories that fall below 2^{20} during sampling are discarded for that trajectory. For $(3n+1)$, we use $B = 80$, $\beta = 100$, $\sigma = 300$, $N = 10{,}000$. For $(5n + 1)$, we use $B = 200$ to sample the same high-altitude branch regime while avoiding finite-height absorption effects, with $\beta = 100$, $\sigma = 300$, $N = 10{,}000$.

All computations used arbitrary-precision integer arithmetic (Python). Random starts were sampled uniformly among odd integers below 2^B using the standard pseudorandom generator with fixed seeds (42–45 across the experiments). The pseudorandom generator, seed values, source code, and full 21-cell tables are available in the accompanying computational supplement; the appendix contains the complete cell-by-cell tables.

7.2 Experiment 1: aggregate branch distribution

The aggregate K distribution across all sampled transitions is compared to the theoretical Geom(1/2) law $\mathbb{P}(K = k) = 2^{-k}$. Both $(3n + 1)$ and $(5n + 1)$ data match Geom(1/2) within 1.5% for $k \leq 8$:

k	ideal 2^{-k}	$(3n+1)$ empirical	$(5n+1)$ empirical
1	0.5000	0.4992	0.5000
2	0.2500	0.2497	0.2501
3	0.1250	0.1266	0.1250
4	0.0625	0.0626	0.0626
5	0.0312	0.0311	0.0312
6	0.0156	0.0155	0.0157
7	0.0078	0.0078	0.0076
8	0.0039	0.0039	0.0038

This numerically recovers the Bernoulli branch law on the sampled integer trajectories.

7.3 Experiment 2: per-cell conditional drift

For each $h \in \Omega_8$, we compute $\mathbb{E}[\Delta \log_2 n \mid h]$ and $\mathbb{E}[K \mid h]$ across sampled transitions. The linear prediction $\mathbb{E}[\Delta \log_2 n \mid h] \approx \log_2 a - \mathbb{E}[K \mid h]$ holds across all 21 cells within sampling noise. For $(3n+1)$ at $\sim 314{,}000$ samples, per-cell $\mathbb{E}[K \mid h]$ ranges from 1.9834 to 2.0334 (within 1.7% of the theoretical value 2.0); per-cell drift correspondingly ranges from -0.448 to -0.398, all matching $1.585 - \mathbb{E}[K \mid h]$ exactly within sampling noise. For $(5n+1)$ at $\sim 3{,}000{,}000$ samples, per-cell drift ranges from $+0.316$ to $+0.327$, all matching $\log_2 5 - \mathbb{E}[K \mid h] \approx 2.322 - 2.0$.

7.4 Experiment 3: stationary weights

We compare empirical cell weights to the predicted π_8 from Theorem 11.

h	π_8 predicted	$(3n+1)$ empirical	ratio	$(5n+1)$ empirical	ratio
00000001	0.01562	0.01329	0.851	0.01564	1.001
00010101	0.06250	0.06201	0.992	0.06250	1.000
01010101	0.12500	0.13086	1.047	0.12512	1.001
10000001	0.01562	0.01407	0.901	0.01561	0.999
10010101	0.06250	0.06476	1.036	0.06245	0.999

(Full 21-cell tables in the supplementary material.)

Survival bias. The $(3n+1)$ empirical weights deviate from π_8 by up to 14.9% in either direction, with single-1 boundary cells (recent $K \geq 7$) systematically undercounted by 10-15%. The same diagnostic on $(5n+1)$ recovers π_8 to within 0.7% on every cell:

	mean ratio	std	max deviation
$(3n+1)$ (descent)	0.9797	0.0536	14.9%
$(5n+1)$ (growth)	1.0000	0.0030	0.7%

This isolates the discrepancy as finite-height survival conditioning rather than a failure of the Bernoulli branch model: in the descending $(3n+1)$ dynamics, recent large-K transitions produce large downward jumps and increase the chance that a trajectory enters the absorbing small-value region before sampling; in the growing $(5n+1)$ control, this censoring mechanism is absent. The 2-adic Bernoulli branch model is exact under the Haar branch measure (Theorem 8); the integer-trajectory sampling on descending dynamics introduces a finite-height survival correction whose sign and localization are explained by descent.

8 Discussion

8.1 Relation to Tao's almost-bounded theorem

The present finite-memory result is orthogonal to Tao's almost-bounded theorem [8], which proves that for any $f : \mathbb{N} \to \mathbb{N}$ with $f(N) \to \infty$, the set $\{N : \mathrm{Col}_{\min}(N) \le f(N)\}$ has logarithmic density 1. Tao's theorem operates at a much deeper level via Syracuse random variables on 3-adic cyclic groups, characteristic-function estimates, and an approximate transport property. Our theorem is exact under the 2-adic Bernoulli branch measure but does not bridge to logarithmic density on $\mathbb{N}$, much less pointwise.

What our theorem provides that Tao's does not: an explicit finite-memory phase-space description (the Fibonacci recurrent class and its stationary weights), a uniform finite-history conditional drift, and a structural impossibility result for finite-history Lyapunov corrections. What Tao's theorem provides that ours does not: a rigorous descent statement on integer orbits in logarithmic density.

8.2 Pointwise Collatz residue

The pointwise Collatz conjecture remains outside the reach of the finite-memory theorem proved here. By Corollary 13, no history-only correction (bounded or unbounded) can convert the uniform conditional drift to pointwise descent. The pointwise residue lies in the possibility of integer trajectories whose long-range branch-exponent sequences are statistically non-generic: trajectories whose empirical branch distribution deviates persistently from Geom(1/2), or whose long-range correlations between K_t and the history h_t deviate from independence. Our experiments find no detectable deviation at 80-bit and 200-bit starts on 10^5-10^6 samples, but generic statistical agreement at finite samples does not exclude a thin pointwise exceptional set.

8.3 Connection to fixed-window Lyapunov programs

Several Collatz proof attempts have sought finite-window Lyapunov certificates on the parity history or on a related block graph. Our Corollary 13 explains why these attempts must fail to give pointwise descent: the alternating-state self-loop with $K = 1$ produces a positive log-step from any state back to itself, making any finite-history correction useless on this transition. The negative drift is real but measure-level, not pointwise.

8.4 Coefficient-specificity and the Conway-Kurtz-Simon barrier

Building on Conway's undecidability work [2], Kurtz and Simon [4] showed that a natural generalized Collatz problem is Π_2^0-complete: there exist Collatz-shaped iterations encoding the halting problem at the family level. Our coefficient-specificity result (Example 10) shows that $\log_2 a - 2 < 0$ is a necessary condition for finite-memory descent and distinguishes original Collatz ($a = 3$) from $5n + 1$ ($a = 5$). Any proof of original Collatz must use this coefficient-specific arithmetic; finite-memory uniform descent does not transfer to general Collatz-shaped iterations.

8.5 Open questions

1. Does the empirical branch distribution on integer trajectories agree with Geom(1/2) to all orders, or are there higher-moment deviations detectable at large N and large samples?

2. Is there a longer-range memory structure (window size $m \to \infty$, or memory beyond the rolling window) under which a sharper supermartingale or even a pointwise Lyapunov could be constructed?
3. How does the survival-bias correction quantify the gap between the asymptotic Bernstein-Lagarias measure and finite-integer-trajectory empirical distributions? Can it be made into a rigorous correction term with computable bounds?

Disclosure of AI-assisted research

During exploratory development, the author used AI-assisted systems to propose diagnostics, critique formulations, assist with proof organization, and revise explanatory text. The author independently reviewed the proofs, computations, citations, and final claims, and takes responsibility for the manuscript.

Acknowledgments

The author thanks the Velisyl Constellation collaboration network for sustained dialogue on this work.

References

[1] D. J. Bernstein and J. C. Lagarias. The $3x + 1$ conjugacy map. *Canadian Journal of Mathematics*, 48(6):1154–1169, 1996.

[2] J. H. Conway. Unpredictable iterations. In *Proc. 1972 Number Theory Conference*, pages 49–52. University of Colorado, 1972.

[3] I. Krasikov and J. C. Lagarias. Bounds for the $3x + 1$ problem using difference inequalities. *Acta Arithmetica*, 109(3):237–258, 2003.

[4] S. A. Kurtz and J. Simon. The undecidability of the generalized Collatz problem. In *Theory and Applications of Models of Computation*, volume 4484 of *Lecture Notes in Computer Science*, pages 542–553. Springer, 2007.

[5] C. Hercher. There are no Collatz m-cycles with $m \leq 91$. *Journal of Integer Sequences*, 26(3):Article 23.3.5, 2023. arXiv:2201.00406.

[6] D. Barina. Improved verification limit for the convergence of the Collatz conjecture. *The Journal of Supercomputing*, 81:Article 810, 2025. doi:10.1007/s11227-025-07337-0.

[7] R. P. Steiner. A theorem on the Syracuse problem. In *Proceedings of the 7th Manitoba Conference on Numerical Mathematics and Computing, 1977*, Congressus Numerantium 20, pages 553–559. Utilitas Mathematica, 1978.

[8] T. Tao. Almost all orbits of the Collatz map attain almost bounded values. *Forum of Mathematics, Pi*, 10:E12, 2022. arXiv:1909.03562.

[9] R. Terras. A stopping time problem on the positive integers. *Acta Arithmetica*, 30(3):241–252, 1976.

A Complete 21-cell tables

We list the full $\Omega_8 = 21$ cell tables for the three integer-trajectory experiments.

A.1 Per-cell conditional drift and $\mathbb{E}[K \mid h]$ for $(3n + 1)$

From 314,261 post-burn-in samples at 80-bit starts (one consistent run):

h	samples	$\mathbb{E}[K \mid h]$	empirical drift	predicted ($\log_2 3 - \mathbb{E}[K \mid h]$)
00000001	4101	2.0334	−0.4484	−0.4484
00000101	8832	1.9908	−0.4059	−0.4059
00001001	9230	2.0179	−0.4329	−0.4329
00010001	9401	2.0047	−0.4197	−0.4197
00010101	19456	2.0093	−0.4243	−0.4243
00100001	9218	2.0114	−0.4264	−0.4264
00100101	19759	1.9980	−0.4130	−0.4130
00101001	19832	1.9849	−0.3999	−0.3999
01000001	9161	2.0031	−0.4181	−0.4181
01000101	19959	1.9980	−0.4130	−0.4130
01001001	19750	1.9995	−0.4145	−0.4145
01010001	20200	2.0011	−0.4161	−0.4161
01010101	40597	1.9834	−0.3985	−0.3985
10000001	4447	2.0040	−0.4191	−0.4191
10000101	9586	2.0187	−0.4337	−0.4337
10001001	9909	1.9841	−0.3991	−0.3991
10010001	9866	1.9924	−0.4074	−0.4074
10010101	20298	1.9959	−0.4109	−0.4109
10100001	9780	2.0076	−0.4226	−0.4226
10100101	20467	2.0139	−0.4289	−0.4289
10101001	20412	2.0035	−0.4186	−0.4186
Σ	314261	—	—	—

The drift-prediction column matches the empirical column to four decimal places throughout (the $O(1/n)$ Archimedean correction is negligible at 80-bit altitude), confirming Corollary 9 on integer trajectories.

A.2 Stationary weights $\pi_8(h)$: predicted vs empirical

h	π_8 predicted	$(3n+1)$ empirical	ratio	$(5n+1)$ empirical	ratio
00000001	0.01562	0.01329	0.851	0.01564	1.001
00000101	0.03125	0.02815	0.901	0.03135	1.003
00001001	0.03125	0.02930	0.938	0.03104	0.993
00010001	0.03125	0.02914	0.932	0.03141	1.005
00010101	0.06250	0.06201	0.992	0.06250	1.000
00100001	0.03125	0.02927	0.937	0.03130	1.002
00100101	0.06250	0.06217	0.995	0.06230	0.997
00101001	0.06250	0.06400	1.024	0.06253	1.001
01000001	0.03125	0.02890	0.925	0.03131	1.002
01000101	0.06250	0.06288	1.006	0.06260	1.002
01001001	0.06250	0.06274	1.004	0.06239	0.998
01010001	0.06250	0.06327	1.012	0.06243	0.999
01010101	0.12500	0.13086	1.047	0.12512	1.001
10000001	0.01562	0.01407	0.901	0.01561	0.999
10000101	0.03125	0.03098	0.991	0.03117	0.997
10001001	0.03125	0.03083	0.986	0.03132	1.002
10010001	0.03125	0.03133	1.003	0.03141	1.005
10010101	0.06250	0.06476	1.036	0.06245	0.999
10100001	0.03125	0.03134	1.003	0.03105	0.993
10100101	0.06250	0.06575	1.052	0.06249	1.000
10101001	0.06250	0.06495	1.039	0.06257	1.001
Σ	1.00000	1.00000	—	1.00000	—

The $(3n+1)$ column shows systematic undercount on small-π_8 cells (single-1 boundary cells) and slight overcount on large-π_8 cells. The $(5n+1)$ column matches predictions to within 0.7% across all cells. The asymmetry confirms the survival-bias diagnosis: long-K transitions in descending dynamics produce large downward jumps that increase the chance of trajectory absorption before sampling, undercounting the corresponding boundary cells.

A.3 Aggregate K distribution comparison

k	ideal 2^{-k}	$(3n+1)$ empirical	$(5n+1)$ empirical
1	0.5000	0.4992	0.5000
2	0.2500	0.2497	0.2501
3	0.1250	0.1266	0.1250
4	0.0625	0.0626	0.0626
5	0.0312	0.0311	0.0312
6	0.0156	0.0155	0.0157
7	0.0078	0.0078	0.0076
8	0.0039	0.0039	0.0038
9	0.0020	0.0019	0.0020
10	0.0010	0.0008	0.0010
11	0.0005	0.0005	0.0005