

Polylogarithmic Descent for Almost All Collatz Orbits in Natural Density

Idris Ali Shaik

Independent researcher, Dubai, United Arab Emirates

Email: shaikidris@apache.org

ORCID: 0009-0009-9699-9712

Date: 4 October 2026

Version: 4.0.2

Preprint DOI: 10.5281/zenodo.23132901

2020 Mathematics Subject Classification: Primary 11B83; Secondary 37P99, 60G40

Keywords: Collatz map, almost all Collatz orbits, natural density, first passage, stopping time, parity vectors, polylogarithmic descent

Abstract

Korec proved that, for each fixed $\theta > \log 3 / \log 4$, almost every positive integer n has a Collatz iterate below n^θ in natural density. Tao proved descent below every prescribed diverging threshold in logarithmic density. We obtain polylogarithmic descent in ordinary natural density within logarithmically many steps, with bounds on the preceding orbit values.

The shortcut map T sends even n to $n/2$ and odd n to $(3n+1)/2$. The unaccelerated map Col sends even n to $n/2$ and odd n to $3n+1$. For any prescribed positive function $f(n) \rightarrow \infty$, however slowly, we prove that almost every positive integer n , in natural density, has nonnegative integers $k = k(n)$ and $\ell = \ell(n)$ satisfying

$$T^k(n) = \text{Col}^\ell(n) < (\log n)^{A_{\text{FP}}} f(n),$$

where $A_{\text{FP}} \approx 9.9911$ is an explicit exponent. The shortcut time is $k = c_* \log n + O_f(\sqrt{\log n \log \log n})$, where $c_* = 2/\log(4/3)$. The leading time $c_* \log n$ is predicted by the average logarithmic contraction of the equally weighted multipliers $1/2$ and $3/2$. The unaccelerated time is $\ell = \frac{3}{2}c_* \log n + O_f(\sqrt{\log n \log \log n})$. Both orbits are bounded by $nf(n)$ up to their respective times.

For each fixed $A > A_{\text{FP}}$, put $q = A/(2A_{\text{FP}}) - \frac{1}{2}$. For sufficiently large X , all but at most $C_A X/(\log X)^q$ integers $n \leq X$ satisfy the same conclusions with target $(\log n)^A$ and bound $n(\log n)^q$ on both orbits. The constant C_A and the time-error constants depend only on A .

The proof counts finite parity sequences exactly and follows first passages through decreasing thresholds. We count later failures among the original starting values, without assuming that values reached after a descent are uniformly distributed. For starting values $2^M \leq n < 2^{M+1}$, only $O(\sqrt{M \log M})$ cumulative passage times need to be counted at each threshold. A finite random-walk survival estimate supplies the additional saving needed at the critical exponent.

Density-one descent below a fixed multiple of $(\log n)^{A_{\text{FP}}}$ is not established, and the pointwise Collatz conjecture remains open. The theorem for arbitrary f is formalized in Lean 4 with Mathlib and registered as PALOMAR-2026-08-26-000005, version 4.

1. Introduction and main results

We study how far a Collatz orbit descends for almost every starting value, how long this takes, and how large the preceding orbit values can be. The main result reaches a polylogarithmic value in ordinary natural density. It gives two-sided time estimates for both Collatz maps and a prescribed bound on all orbit values up to those times.

The shortcut map T sends even n to $n/2$ and odd n to $(3n+1)/2$. The unaccelerated map Col sends even n to $n/2$ and odd n to $3n+1$. We use $\mathbb{N} = \{1, 2, \dots\}$, write $\log$ for the natural logarithm and $\log_2$ for the base-two logarithm, and put $T_{\min}(n) = \min_{j \geq 0} T^j(n)$. A set $S \subseteq \mathbb{N}$ has natural density one if $\#(S \cap [1, X])/X \rightarrow 1$.

Let $H_2(p) = -p \log_2 p - (1-p) \log_2(1-p)$ be binary entropy and put

$$p_* = \log_3 2, \quad \kappa_* = 1 - H_2(p_*), \quad A_{\text{FP}} = \frac{1}{2\kappa_*} \approx 9.9911, \quad c_* = \frac{2}{\log(4/3)} \approx 6.9521. \quad (1.1)$$

Theorem 1.1 (polylogarithmic descent)

Let $U, R : \mathbb{N} \rightarrow (0, \infty)$ be prescribed functions tending to infinity. There are a set S of natural density one and constants $K, n_0 > 0$, depending only on U and R , such that for every $n \in S$ with $n \geq n_0$, there are nonnegative integers $k = k(n)$ and $\ell = \ell(n)$ satisfying

$$T^k(n) = \text{Col}^\ell(n) < (\log n)^{A_{\text{FP}}} U(n),$$

$$|k - c_* \log n| + \left| \ell - \frac{3}{2} c_* \log n \right| \leq K \sqrt{\log n \log \log n},$$

and

$$\max_{0 \leq j \leq k} T^j(n) \leq nR(n), \quad \max_{0 \leq i \leq \ell} \text{Col}^i(n) \leq nR(n).$$

Neither function is required to be monotone.

The abstract uses one function f , with $U = R = f$. Conversely, the choice $f(n) = \min\{U(n), R(n)\}$ gives the independent bounds above. The proof constructs a value below n for sufficiently large n , even when the prescribed target is larger than n . A fixed multiple of $(\log n)^{A_{\text{FP}}}$ is not covered by this theorem.

The fixed-exponent case has an explicit bound on the number of exceptions.

Corollary 1.2 (quantitative fixed-exponent descent)

For every fixed $A > A_{\text{FP}}$, put $q = \kappa_* A - 1/2 > 0$. There are constants $C_A, K_A, n_A > 0$ such that, for every sufficiently large X , all but at most $C_A X / (\log X)^q$ integers $1 \leq n \leq X$ have $n \geq n_A$ and nonnegative integers k, ℓ satisfying

$$T^k(n) = \text{Col}^\ell(n) < (\log n)^A,$$

$$|k - c_* \log n| + \left| \ell - \frac{3}{2} c_* \log n \right| \leq K_A \sqrt{\log n \log \log n},$$

$$\max_{0 \leq j \leq k} T^j(n) \leq n(\log n)^q, \quad \max_{0 \leq i \leq \ell} \text{Col}^i(n) \leq n(\log n)^q.$$

The constants depend only on A .

The count concerns failure to satisfy these conditions together. Its order is sharp for this joint conclusion: the height bound itself forces at least a constant times $X/(\log X)^q$ exceptions, as shown in Proposition 7.1. This is not a lower bound for failure to descend when no height bound is imposed.

The two results follow directly from the shell estimate in Proposition 6.4. Their proofs are given at the end of Section 6.

Relation to earlier results

Terras [11], [12] and Everett [2] proved that almost every starting value eventually falls below its initial value. Korec [5] obtained descent below n^θ in natural density for every fixed $\theta > \log 3 / \log 4$. Inslemann [3] obtains every fixed power n^ε , with shortcut time $\lfloor c_* \log n \rfloor$. His Theorems 1.1 and 1.9 also bound the shortcut and unaccelerated orbits throughout the stated time interval. For each fixed $\varepsilon > 0$, his Theorem 2.18 supplies a power-saving exceptional count $O_\varepsilon(X^{1-d_\varepsilon})$ for some $d_\varepsilon > 0$.

Tao [10] proves descent below every prescribed diverging threshold in logarithmic density. The preprints of Allikvere [14] and Mazur [15] state descent below every such threshold in natural density. Their stated unaccelerated times are eventually less than $12 \log n$ and $436 \log n$, respectively.

We give a proof using finite parity words and first passages. For polylogarithmic targets, it gives explicit exceptional counts and two-sided estimates of both times. The bound on the preceding orbits can be prescribed independently. We do not claim descent below an arbitrary diverging absolute threshold.

Their quantitative fixed-target statements concern a different event. For a threshold $N_0 \geq 2$, Allikvere's Theorem 1.3 bounds the proportion of odd inputs up to X that fail to reach N_0 within the stated Syracuse time by $O((\log N_0)^{-1/29} + X^{-1/2000})$. Mazur's Theorem 1.2 gives $O_d((\log N_0)^{-d})$ for each $0 < d < 5/143$. A Syracuse step sends odd n to $(3n+1)/2^{\nu_2(3n+1)}$, where $\nu_2(m)$ is the exponent of 2 dividing m . These statements do not assert the two-sided times and prescribed height bounds of Theorem 1.1.

The earlier endpoint-transport result [9] concerns a stretched-logarithmic target; the proof here is independent of it. For surveys and complementary approaches, see [1], [6], [7], and [13].

The main estimate and its proof

The constant c_* is the reciprocal of the mean logarithmic decrease of the equally weighted multipliers $1/2$ and $3/2$. Put

$$a_0 = \frac{\log_2 3}{2}, \quad c_*^{-1} = (1 - a_0) \log 2 = \frac{1}{2} \log(4/3). \quad (1.2)$$

For the drift heuristics, see [4] and [8].

The obstacle is repeating a descent estimate. Values reached after one descent need not be uniformly distributed. We count the starting values whose orbits first enter a specified set, then repeat this argument along decreasing thresholds. Every later passage is still a first passage of the original orbit. This also keeps track of the additive terms in the Collatz map.

Let $I_M = [2^M, 2^{M+1}) \cap \mathbb{N}$ be the M -th dyadic shell. For a final threshold 2^L , with $L \asymp \log M$, and an orbit bound nR_M , where $R_M \geq 4$, the principal estimate in Proposition 6.4 bounds the exceptional proportion by

$$C \left(\sqrt{\frac{(M+2) \log(M+2)}{L+1}} 2^{-\kappa_* L} + (M+2)^{-N} + R_M^{-1} \right).$$

Here $N > 0$ is fixed. The constant depends on N and the fixed upper and lower bounds for $L/\log(M+2)$. It is uniform in L and R_M subject to those bounds.

Three estimates produce the first term. Prefix bounds leave only $O(\sqrt{M \log M})$ possible cumulative passage times to each threshold. Reversing a passage contributes a factor of order L . A finite random-walk survival estimate contributes $O(L^{-3/2} 2^{-\kappa_* L})$. When $L \sim A \log_2 M$, their product has order $M^{-(\kappa_* A - 1/2)}$. This explains both $A_{\text{FP}} = 1/(2\kappa_*)$ and the exponent q in Corollary 1.2. At the critical exponent, every divergent factor is allowed, but a fixed factor is not established.

Sections 2–3 establish the parity identities and bounds on one orbit block. Sections 4–5 count first passages through decreasing thresholds. Section 6 combines these estimates with the finite survival estimate in Appendix A and proves Theorem 1.1 and Corollary 1.2. Section 7 proves the matching lower bound forced by the orbit bound.

Theorem 1.1 is formalized in Lean 4 with Mathlib and registered as PALOMAR-2026-08-26-000005, version 4 [17]. Section 8 specifies the formalization scope. The use of AI assistance is described in the research disclosure.

2. Density, parity, and affine iterates

The argument uses three common tools: a shell-to-prefix summation rule, exact parity coding on each dyadic shell, and an affine formula for the iterate. This section establishes all three before they are used. Lemma 2.1 converts shell exceptional estimates into global natural-density bounds. Proposition 2.2 makes parity words exactly uniform on a complete shell, supplying the counting law used in Sections 3 and 6. Proposition 2.3 separates each iterate into its multiplicative term and affine correction, which is used to bound every iterate in Section 3.

A dyadic threshold is a value 2^q . A first passage is the first time the orbit is at or below that value. A block is the orbit segment up to such a crossing; successive blocks start at the values reached by the preceding crossings. The shell rank of n is $\lfloor \log_2 n \rfloor$. The times k and ℓ count shortcut and unaccelerated steps, respectively. If s_h is the number of odd states in the first h shortcut steps, the same value is reached after $h + s_h$ unaccelerated steps.

For an exceptional set $E \subseteq \mathbb{N}$, shell estimates imply the following bounds for the whole initial interval.

Lemma 2.1 (dyadic summation)

Suppose that $e_M \geq 0$ and

$$\#(E \cap I_M) \leq e_M 2^M \tag{2.1}$$

for all sufficiently large M . If $e_M \rightarrow 0$, then E has natural density zero. If, for some fixed $q > 0$, one has $e_M = O(M^{-q})$, then

$$\#(E \cap [1, X]) = O\left(\frac{X}{(\log X)^q}\right). \tag{2.2}$$

Proof Fix $\epsilon > 0$ and choose M_0 so that $e_M \leq \epsilon$ for $M \geq M_0$. If $2^J \leq X < 2^{J+1}$, these shells contribute at most $\epsilon \sum_{M=M_0}^J 2^M \leq 2\epsilon X$. The earlier shells contain a fixed finite number of integers. Divide by X , then let $\epsilon \downarrow 0$, to obtain the qualitative assertion.

For the quantitative assertion, the shells with $M < \lfloor J/2 \rfloor$ contain fewer than $2^{J/2} \leq \sqrt{X}$ integers. On every remaining shell, $e_M \ll J^{-q}$, so their total contribution is $O(XJ^{-q})$. Since $J \asymp \log X$ and $\sqrt{X} = o(X/(\log X)^q)$, this proves (2.2). $\square$

For $n \geq 1$, define

$$p_i(n) = \mathbf{1}_{\{T^i(n) \text{ odd}\}}, \quad s_k(n) = \sum_{i=0}^{k-1} p_i(n). \quad (2.3)$$

The exact parity coding below is classical in the stopping-time literature; see, for example, Terras [11]. Its short proof is included because no external result is needed by the argument.

Proposition 2.2 (parity-vector bijection)

For every $M \geq 0$,

$$n \bmod 2^M \mapsto (p_0(n), \dots, p_{M-1}(n)) \quad (2.4)$$

is a bijection from $\mathbb{Z}/2^M\mathbb{Z}$ to $\{0, 1\}^M$.

Proof The relation

$$n \equiv n' \pmod{2^{k+1}} \implies T(n) \equiv T(n') \pmod{2^k}$$

shows inductively that the first M parity bits depend only on $n \bmod 2^M$. For the converse, induct on the word length. Suppose the first bit is e , and use the induction hypothesis to recover the residue $m = T(n) \bmod 2^k$ from the remaining k bits. If $e = 0$, the unique compatible residue is

$$n \equiv 2m \pmod{2^{k+1}}.$$

If $e = 1$, it is

$$n \equiv (2m - 1)3^{-1} \pmod{2^{k+1}},$$

where 3 is a unit modulo 2^{k+1} . The first residue is even, the second is odd, and in either case applying T recovers $m \bmod 2^k$. Thus every word has exactly one preimage residue. $\square$

Consequently, as n ranges over the complete shell I_M , its length- M parity word is exactly uniform on $\{0, 1\}^M$.

Proposition 2.3 (exact affine iterate)

For every $n \geq 1$ and $k \geq 0$,

$$T^k(n) = \frac{3^{s_k(n)}}{2^k} n + \sum_{i=0}^{k-1} \frac{p_i(n) 3^{s_k(n) - s_{i+1}(n)}}{2^{k-i}}. \quad (2.5)$$

Proof The assertion is trivial at $k = 0$. Passing from k to $k + 1$, an even letter divides every existing term by two. An odd letter multiplies every existing term by $3/2$ and adds $1/2$. This is exactly the displayed update. $\square$

3. Parity-prefix bounds on an orbit block

The parity bijection identifies a complete dyadic shell with the Boolean cube. We now discard the sparse words whose prefixes deviate too far from their mean. The resulting bounds hold at every step of the block and will guarantee the first passages used later.

For a parity word of length M , put

$$Y_k = 2s_k - k, \quad H_M = \max_{0 \leq k \leq M} \left| s_k - \frac{k}{2} \right| = \frac{1}{2} \max_{0 \leq k \leq M} |Y_k|. \quad (3.1)$$

For $0 < p, q < 1$, write

$$D(p||q) = p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q}$$

for binary relative entropy. For $0 \leq t < 1/2$, put

$$\mathcal{I}(t) = D\left(\frac{1}{2} + t \parallel \frac{1}{2}\right) = \left(\frac{1}{2} + t\right) \log(1+2t) + \left(\frac{1}{2} - t\right) \log(1-2t). \quad (3.2)$$

Lemma 3.1 (maximal deviation of parity prefixes)

If $M \geq 1$ and $0 \leq h < M/2$, then

$$2^{-M} \#\{w \in \{0,1\}^M : H_M(w) > h\} \leq 2 \exp(-M\mathcal{I}(h/M)). \quad (3.3)$$

In particular, the right side is at most $2 \exp(-2h^2/M)$.

Proof Fix $\theta > 0$ and stop the Boolean tree at the first prefix u with $|Y(u)| > 2h$. For a frontier node of depth k , set

$$\Phi_\theta(u) = 2^{-k} \cosh(\theta Y(u)) (\cosh \theta)^{M-k}.$$

The identity

$$\cosh(\theta(y-1)) + \cosh(\theta(y+1)) = 2 \cosh(\theta y) \cosh \theta$$

shows that replacing an unexpanded node by its two children preserves the total potential. The frontier formed by first-crossing prefixes and noncrossing depth- M words therefore has total $(\cosh \theta)^M$. Each crossing prefix contributes at least $2^{-|u|} \cosh(2\theta h)$, whence

$$\Pr(H_M > h) \leq 2 \exp(M \log \cosh \theta - 2\theta h). \quad (3.4)$$

Take $u = 2h/M$ and $\theta = \operatorname{arctanh} u$. The elementary Legendre identity

$$u\theta - \log \cosh \theta = D\left(\frac{1+u}{2} \parallel \frac{1}{2}\right) = \mathcal{I}(h/M)$$

proves (3.3). Pinsker's inequality in this binary case, or the direct convexity estimate $\mathcal{I}(t) \geq 2t^2$, gives the quadratic assertion.

□

Set

$$\rho := \frac{\sqrt{3}}{2} = 2^{a_0-1}.$$

Define

$$r_k(n) = \sum_{i=0}^{k-1} \frac{p_i(n)}{3^{s_{i+1}(n)} 2^{k-i}}, \quad d_k = s_k - \frac{k}{2}. \quad (3.5)$$

By Proposition 2.3,

$$T^k(n) = \rho^k n 3^{d_k} + (r_k(n) 3^{k/2}) 3^{d_k}. \quad (3.6)$$

Lemma 3.2 (uniform affine correction)

If $H_M \leq h$, then for every $1 \leq k \leq M$,

$$r_k(n)3^{k/2} \leq (2 + \sqrt{3}) \left(1 - \rho^k\right) 3^h < (2 + \sqrt{3})3^h. \quad (3.7)$$

Proof For $0 \leq i < k$, the barrier gives $s_{i+1} \geq (i+1)/2 - h$. Hence

$$r_k 3^{k/2} \leq 3^h \sum_{i=0}^{k-1} \frac{3^{(k-i-1)/2}}{2^{k-i}} = \frac{3^h}{2} \sum_{j=0}^{k-1} \rho^j,$$

which is (3.7). $\square$

For $0 < \eta \leq 1$, let W_η be the set of integers n such that

$$\rho^k n^{1-\eta} \leq T^k(n) \leq \rho^k n^{1+\eta} \quad (0 \leq k \leq \lfloor \log_2 n \rfloor). \quad (3.8)$$

Put

$$c_0 = \frac{1}{2(\log_2 3)^2}, \quad C_0 = 2e^{4c_0}.$$

The following estimate is uniform in the tolerance:

$$\frac{\#(I_m \setminus W_t)}{2^m} \leq C_0 e^{-c_0 t^2 m} \quad (0 < t \leq 1). \quad (3.9)$$

Proof When $m \geq 4$ and $tm \geq 2$, apply Lemma 3.1 at height

$$h = \frac{tm}{2 \log_2 3}, \quad 3^h = 2^{tm/2}.$$

Here $h < m/2$, since $t \leq 1 < \log_2 3$, so the height is in the range of Lemma 3.1. If $H_m \leq h$, then $|d_k| \leq h$ for every $k \leq m$. Since $n \geq 2^m$,

$$3^{-h} = 2^{-tm/2} \geq n^{-t}, \quad 3^h = 2^{tm/2} \leq 2^{-tm/2} n^t.$$

Thus the multiplicative term in (3.6) lies above $\rho^k n^{1-t}$ and below $2^{-tm/2} \rho^k n^{1+t}$. The unused difference from the upper bound is at least

$$\left(1 - 2^{-tm/2}\right) \rho^m 2^{(1+t)m} \geq \frac{1}{2} 2^{(a_0+t)m}.$$

Meanwhile Lemma 3.2 and $d_k \leq h$ bound the additive term in (3.6) by

$$(2 + \sqrt{3})3^{2h} = (2 + \sqrt{3})2^{tm}.$$

For $m \geq 4$, this is at most the displayed slack because $2^{a_0 m}/2 \geq 2^{4a_0}/2 = 9/2 > 2 + \sqrt{3}$. Hence for $n \in I_m$, $H_m(n) \leq h$ implies $n \in W_t$, or equivalently $I_m \setminus W_t \subseteq \{x \in I_m : H_m(x) > h\}$. By Proposition 2.2, Lemma 3.1, and $\mathcal{I}(u) \geq 2u^2$, the complementary proportion is at most

$$2 \exp\left(-\frac{t^2 m}{2(\log_2 3)^2}\right) = 2e^{-c_0 t^2 m}.$$

When $m < 4$ or $tm < 2$, $t^2 m \leq 4$, so the trivial shell bound is absorbed by $C_0 = 2e^{4c_0}$. Thus (3.9) is uniform in t . $\square$

4. Counting first-passage preimages

The barrier identifies a sparse set of bad orbit values. To count the starting values whose orbits reach such values, we reverse their first passage to a lower threshold. The estimate below bounds the number reaching any specified subset of the crossing interval.

For real $Y > 1$, define

$$\tau_Y(n) = \min\{h \geq 0 : T^h(n) \leq Y\} \quad (4.1)$$

when the set is nonempty.

Lemma 4.1 (first-passage band and reverse product)

Let $1 < Y < 2^M$, let $n \in I_M$, and suppose $\tau_Y(n) = h \geq 1$. Write $x_j = T^j(n)$, $y = x_h$, and let s be the number of odd values among $x_0, \dots, x_{h-1}$. Then

$$\frac{Y}{2} < y \leq Y \quad (4.2)$$

and

$$n = \frac{2^h y}{3^s} \prod_{\substack{0 \leq j < h \\ x_j \text{ odd}}} \left(1 - \frac{1}{2x_{j+1}}\right). \quad (4.3)$$

Consequently, whenever $h < 2Y$,

$$\left(1 - \frac{h}{2Y}\right) \frac{2^h y}{3^s} \leq n \leq \frac{2^h y}{3^s}. \quad (4.4)$$

Proof The final crossing cannot be odd: otherwise $x_h = (3x_{h-1} + 1)/2 > x_{h-1} > Y$. Hence $x_{h-1} = 2y$, proving (4.2). Reversing an even step gives $x_j = 2x_{j+1}$, while reversing an odd step gives

$$x_j = \frac{2x_{j+1} - 1}{3} = \frac{2}{3}x_{j+1} \left(1 - \frac{1}{2x_{j+1}}\right).$$

Multiplication proves (4.3). Every odd factor occurs before the final crossing, so its following state is greater than Y . Therefore $\prod_i (1 - u_i) \geq 1 - \sum_i u_i$ gives a lower product bound $1 - s/(2Y) \geq 1 - h/(2Y)$; the upper bound is one. $\square$

Write the reverse product in (4.3) as $\prod_{j < h} (1 - u_j)$, where

$$u_j(n) = \begin{cases} \frac{1}{2T^{j+1}(n)}, & T^j(n) \text{ odd}, \\ 0, & T^j(n) \text{ even}, \end{cases} \quad E_Y(n) = Y \sum_{j=0}^{h-1} u_j(n). \quad (4.5)$$

The factor Y records the additive reverse loss at the target scale: E_Y/Y is an additive upper bound for the defect of the reverse product. It also makes concatenation compatible with changing thresholds. If a segment is measured locally at $Y' \geq Y$, then its contribution after rescaling to the final threshold Y is exactly Y/Y' times its local contribution to $E_{Y'}$. Thus segment losses add after rescaling, which is the coordinate used below to bound the total loss along a decreasing threshold chain. The elementary product inequality gives

$$1 - \frac{E_Y(n)}{Y} \leq \prod_{j=0}^{h-1} (1 - u_j(n)) \leq 1. \quad (4.6)$$

Lemma 4.2 (starting values reaching one value at a fixed time)

Let $D \geq 0$, and suppose

$$\frac{D}{Y} \leq \frac{1}{3}. \quad (4.7)$$

For a fixed pair (h, y) , the number of starting values $n \in I_M$ satisfying

$$\tau_Y(n) = h, \quad T^h(n) = y, \quad E_Y(n) \leq D$$

is at most

$$1 + 3D \frac{2^M}{Y}. \quad (4.8)$$

Proof Put $P(n) = \prod_{j < h} (1 - u_j(n))$. By Lemma 4.1, if two such starting values had odd counts $s_1 < s_2$, put $A_i = 2^{h_i} y / 3^{s_i}$. Then $A_1 \geq 3A_2$, while (4.3) and (4.6) give

$$n_1 = A_1 P(n_1) \geq 3A_2 \left(1 - \frac{D}{Y}\right) \geq 2A_2 \geq 2n_2.$$

This is impossible in the half-open shell I_M . The argument uses only positive products and introduces no division by a reverse product. Fix the common odd count and put $A = 2^h y / 3^s$. All such starting values lie in $[(1 - D/Y)A, A]$. Nonemptiness, (4.7), and $n < 2^{M+1}$ imply $A < (3/2)2^{M+1}$, so this interval has length at most

$$3D \frac{2^M}{Y}.$$

Any additional congruence or residue restriction can only reduce the number of starting values in this interval. This proves (4.8). $\square$

Proposition 4.3 (starting values reaching a specified set)

Put $J_Y = (Y/2, Y] \cap \mathbb{N}$. Let $1 < Y < 2^M$ and $H \in \mathbb{Z}_{\geq 0}$. Under (4.7), every $B \subseteq J_Y$ satisfies the exact bound

$$\# \left\{ n \in I_M : \begin{array}{l} \tau_Y(n) \leq H, \\ E_Y(n) \leq D \end{array} \quad T^{\tau_Y(n)}(n) \in B, \right\} \leq H \left(1 + 3D \frac{2^M}{Y}\right) |B|. \quad (4.9)$$

Consequently,

$$\# \left\{ n \in I_M : \begin{array}{l} \tau_Y(n) \leq H, \\ E_Y(n) \leq D \end{array} \quad T^{\tau_Y(n)}(n) \in B, \right\} \leq H(1 + 3D) \frac{2^M}{Y} |B|. \quad (4.10)$$

The estimate remains valid when the starting values are restricted to any subset.

Proof Since $Y < 2^M$, every passage time is positive. Sum Lemma 4.2 over the at most $H|B|$ pairs to obtain (4.9). The contribution $H|B|$ is at most $H(2^M/Y)|B|$, which gives (4.10). $\square$

5. Decreasing thresholds and accumulated reverse loss

To apply Section 4 repeatedly, we need to identify each successive crossing as a first passage of the original orbit. We also need a bound on its reverse product that depends on the final threshold, rather than on the full elapsed time. The following two lemmas provide these facts.

Consider a finite sequence of thresholds with strictly decreasing integer exponents $q_0 > \dots > q_i \geq 1$. Set $n_0 = n$, $t_0 = 0$, and suppose that $n_j > 2^{q_j}$ and the first-passage time $h_j = \tau_{2^{q_j}}(n_j)$ exists. Define

$$n_{j+1} = T^{h_j}(n_j), \quad t_{j+1} = t_j + h_j. \quad (5.1)$$

Lemma 5.1 (successive crossings are direct first passages)

For every block in this sequence,

$$t_{j+1} = \tau_{2^{q_j}}(n), \quad n_{j+1} = T^{\tau_{2^{q_j}}(n)}(n). \quad (5.2)$$

Proof The assertion is the definition for $j = 0$. Inductively, every state before t_j exceeds the preceding threshold and therefore also the smaller threshold 2^{q_j} . The orbit from n_j stays above 2^{q_j} until its local first-passage time h_j . Thus $t_j + h_j$ is the original orbit's first crossing of that threshold. $\square$

Lemma 5.2 (reverse loss along decreasing thresholds)

Suppose that, for a fixed $r_* > 0$, every block has duration $h_j \leq m_j$, where $m_j = \lfloor \log_2 n_j \rfloor$, and $q_j + 1 > r_* m_j$. Then the direct first passage to the final threshold satisfies

$$E_{2^{q_i}}(n) < \frac{q_i + 2}{r_*}. \quad (5.3)$$

Proof The final step of each block is even. Every state following an odd step in block j therefore exceeds 2^{q_j} , and each corresponding reverse-product decrement is less than 2^{-q_j-1} . Expressing all decrements at the final scale 2^{q_i} gives

$$E_{2^{q_i}}(n) < \frac{1}{2} \sum_{j \leq i} m_j 2^{q_i - q_j} < \frac{1}{2r_*} \sum_{j \leq i} (q_j + 1) 2^{q_i - q_j}.$$

The exponents q_j are distinct and at least q_i . Hence this last sum is at most

$$\frac{1}{2r_*} \sum_{d \geq 0} (q_i + d + 1) 2^{-d} = \frac{q_i + 2}{r_*},$$

using $\sum_{d \geq 0} 2^{-d} = \sum_{d \geq 0} d 2^{-d} = 2$. $\square$

In Section 6, a block starting at a value in I_m lasts at most m steps. Their threshold exponents satisfy the hypotheses of Lemma 5.2 with one fixed r_* . No condition on the value reached at the final crossing is needed to bound the reverse loss accumulated before it.

6. Shrinking prefix bounds and compressed passage times

Section 4 bounds the number of starting values whose first passage reaches a specified set. Section 5 bounds the accumulated reverse loss. To obtain a polylogarithmic target, we now bound the number of possible passage times and the proportion that fail to cross the next threshold. At large ranks, each descent reduces the rank by a fixed ratio. Shrinking prefix tolerances then keep the sum of the time errors small. Near the final rank, these tolerances no longer give the required exceptional-count bound. We instead decrease the rank by a bounded amount at each passage and use a finite survival estimate to count failures to cross. The change occurs at a rank of order $\log M$. Proposition 6.4 combines the two estimates.

For each shell I_M , prescribe an integer final threshold exponent $L_M \asymp \log M$. The choices yielding Theorem 1.1 and Corollary 1.2 are made after Proposition 6.4. The switch exponent S_M defined below will be chosen so that $L_M < S_M$ for all sufficiently large M .

First passages at large shell ranks

Fix parameters

$$a_0 < r_{\text{hi}} < 1, \quad 0 < \tau < r_{\text{hi}} - a_0, \quad (6.1)$$

and constants $D_{\text{hi}}, C_{\text{sw}} > 0$. Put

$$S_M = \lceil C_{\text{sw}} \log(M+2) \rceil, \quad (6.2)$$

and, for a block starting in I_m with $m \geq S_M$, use the tolerance

$$\eta_{M,m} = \min \left\{ \tau, D_{\text{hi}} \sqrt{\frac{\log(M+2)}{m}} \right\}. \quad (6.3)$$

Since $\eta_{M,m} \leq \tau < r_{\text{hi}} - a_0$, membership in $W_{\eta_{M,m}}$ guarantees a crossing of $2^{\lfloor r_{\text{hi}} m \rfloor}$ within m steps for all sufficiently large m . Indeed,

$$T^m(x) \leq \rho^m x^{1+\eta_{M,m}} < 2^{(a_0+\tau)m+1+\tau} \leq 2^{\lfloor r_{\text{hi}} m \rfloor}.$$

Proposition 6.2 will estimate how many values fail this condition.

Fix M , write $S = S_M$, and define the initial exceptional set $\mathcal{F}_M^{\text{ini}} = I_M \setminus W_{\eta_{M,M}}$. For $n \in I_M \setminus \mathcal{F}_M^{\text{ini}}$, let $n_0 = n$, $m_0 = M$, and $H_0 = 0$. Define a finite sequence recursively. At each $n_i \in I_{m_i}$ with $m_i \geq S$ and $n_i \in W_{\eta_{M,m_i}}$, put

$$q_i = \lfloor r_{\text{hi}} m_i \rfloor, \quad h_i = \tau 2^{q_i}(n_i), \quad n_{i+1} = T^{h_i}(n_i), \quad H_{i+1} = H_i + h_i. \quad (6.4)$$

The value n_{i+1} at the crossing is called the *landing*. An upper endpoint $n_{i+1} = 2^{q_i}$ ends this sequence. When $q_i \leq S$, repeated halving reaches a shell below L_M in at most S further steps. Every nonendpoint landing belongs to I_{q_i-1} . If $q_i - 1 < S$, it is the initial value of the small-rank sequence defined below. If $q_i - 1 \geq S$, the recursion continues precisely when $n_{i+1} \in W_{\eta_{M,q_i-1}}$, with $m_{i+1} = q_i - 1$.

Let $\mathcal{F}_M^{\text{hi}}$ consist of those starting values in $I_M \setminus \mathcal{F}_M^{\text{ini}}$ whose sequence ends at either an upper endpoint with $q_i > S$, or a nonendpoint landing outside $W_{\eta_{M,q_i-1}}$ with $q_i - 1 \geq S$. Every earlier large-rank value satisfies its prefix condition. Thus this set counts the first failure along the sequence.

Lemma 6.1 (possible passage times at large ranks)

There is a constant $K > 0$, depending only on the fixed parameters, such that, for every shell I_M and every threshold rank q reached by the sequence defined by (6.4), all cumulative first-passage times to its landing band $J_{2^q} = (2^{q-1}, 2^q] \cap \mathbb{N}$ belong to a finite set $\mathcal{H}_{M,q}^{\text{hi}}$ satisfying

$$\#\mathcal{H}_{M,q}^{\text{hi}} \leq K \sqrt{(M+2) \log(M+2)}. \quad (6.5)$$

Proof Write $g = 1 - a_0$. We bound each block's duration, then sum the errors over the decreasing shell ranks.

Consider one high block satisfying this condition from the shell I_m to threshold 2^q , and let h be its duration and $t = \eta_{M,m}$. The two bounds in (3.8) and the first-passage band give

$$(1-t)m - q \leq gh < (1+t)(m+1) - q + 1. \quad (6.6)$$

Thus

$$|gh - (m - q)| \leq tm + t + 2. \quad (6.7)$$

The next computation makes this telescoping explicit.

Suppose high blocks $0, \dots, j$ have been executed. Every landing before the last starts another block and lies in $(2^{q_i-1}, 2^{q_i}]$. By the endpoint convention above, a landing equal to 2^{q_i} terminates the chain: it is followed by deterministic halving when $q_i \leq S$, and is a value contributing to $\mathcal{F}_M^{\text{hi}}$ when $q_i > S$. Thus no continuing block starts at an upper endpoint. Each lies in the unique shell I_{q_i-1} , and therefore

$$m_{i+1} = q_i - 1 \quad (0 \leq i < j). \quad (6.8)$$

This removes the apparent rank branching. In particular,

$$\sum_{i=0}^j (m_i - q_i) = M - q_j - j. \quad (6.9)$$

Summing (6.7), using (6.9), and paying one additional unit per block to change the center from $M - q_j - j$ to $M + 1 - q_j$, gives the following bound on cumulative time

$$|gH_{j+1} - ((M + 1) - q_j)| \leq \sum_{i=0}^j (t_i m_i + t_i + 3). \quad (6.10)$$

For every block,

$$t_i m_i \leq D_{\text{hi}} \sqrt{m_i \log(M + 2)}, \quad t_i \leq \tau,$$

and (6.8) gives $m_{i+1} \leq r_{\text{hi}} m_i$. The resulting geometric square-root sum is $O(\sqrt{M \log(M + 2)})$, so the right side of (6.10) has the same order.

For fixed M and $q_j = q$, the center in (6.10) is fixed. Consequently every feasible integer H_{j+1} lies in one interval of length $O(\sqrt{M \log(M + 2)})$; counting its integer points proves (6.5). No union over intermediate high-rank histories remains. $\square$

The fixed-time, fixed-landing count also has the following restricted-time form. If $\mathcal{H}$ is any finite set of positive times, $B \subseteq J_Y$, and $D_q \geq 0$ satisfies $D_q/Y \leq 1/3$, then

$$\# \left\{ n \in I_M : \begin{array}{l} \text{the first passage at or below } Y \text{ occurs at some } h \in \mathcal{H}, \\ T^h(n) \in B, \ E_Y(n) \leq D_q \end{array} \right\} \leq \# \mathcal{H} \left(1 + 3D_q \frac{2^M}{Y} \right) \# B. \quad (6.11)$$

This is Proposition 4.3 summed only over the declared times. No interval structure or density hypothesis on $\mathcal{H}$ is used.

Proposition 6.2 (density of large-rank failures)

Assume

$$\frac{D_{\text{hi}}}{\sqrt{C_{\text{sw}}}} \leq \tau \quad \text{and} \quad 1 \leq S_M \leq M.$$

For a high threshold rank q , define the complete landing-band target

$$B_{M, S_M, q}^{\text{sh}} = \{2^q\} \cup \left[(I_{q-1} \cap J_{2^q}) \setminus W_{\eta_M, q-1} \right].$$

Then the cap in (6.3) is inactive at every $m \geq S_M$, and

$$\eta_{M,m}^2 m = D_{\text{hi}}^2 \log(M+2). \quad (6.12)$$

Moreover,

$$\frac{\#(I_M \setminus W_{\eta_{M,M}})}{2^M} \leq C_0(M+2)^{-c_0 D_{\text{hi}}^2}, \quad (6.13)$$

and, for every q with $S_M \leq q-1$,

$$\frac{|B_{M,S_M,q}^{\text{sh}}|}{2^q} \leq 2^{-q} + \frac{C_0}{2}(M+2)^{-c_0 D_{\text{hi}}^2}. \quad (6.14)$$

Proof From $m \geq S_M \geq C_{\text{sw}} \log(M+2)$,

$$D_{\text{hi}} \sqrt{\frac{\log(M+2)}{m}} \leq \frac{D_{\text{hi}}}{\sqrt{C_{\text{sw}}}} \leq \tau.$$

This proves that the cap is inactive and gives (6.12).

Apply (3.9) with $t = \eta_{M,M}$ and use (6.12) to obtain (6.13).

For $q \geq 1$, the landing band J_{2^q} lies in I_{q-1} apart from its single upper endpoint 2^q . Apply (3.9) in I_{q-1} , divide by 2^q , and use (6.12) at $m = q-1$. This gives (6.14). $\square$

Consequently the initial exceptional set and every large-rank landing target have a common density bound

$$d_{\text{hi}}(M) \ll (M+2)^{-p_{\text{hi}}}, \quad p_{\text{hi}} = \min\{C_{\text{sw}} \log 2, c_0 D_{\text{hi}}^2\}, \quad (6.15)$$

where $c_0 > 0$ is the fixed constant in the quadratic prefix-deviation bound. The first term in the minimum pays for the dyadic endpoint boundary at the switch.

Critical timeout estimate at small shell ranks

For a nonendpoint landing in a shell below S_M , the construction continues at smaller ranks. There the proof no longer imposes the large-rank prefix condition. A failure to cross still gives information about the orbit: every prefix of the actual parity word has kept the orbit above the next threshold. The exact multiplicative recurrence converts this into survival above a fixed logarithmic barrier. The finite ballot estimate proved in Appendix A then gains a factor m^{-1} over the terminal-only binomial bound.

For a block starting in I_m , success means that its first passage to the next threshold occurs within m steps.

Let $L = L_M \asymp \log M$, fix $K_0 > 0$, and put

$$r_L = 1 - \frac{K_0}{2L}, \quad q_L(m) = \lfloor r_L m \rfloor. \quad (6.16)$$

For sufficiently large L , one has $0 < r_L < 1$ and $q_L(m) < m$ whenever $m \geq L$. For $x \in I_m$, define the low timeout event by

$$\text{To}_{L,m}(x) \iff T^j(x) > 2^{q_L(m)} \quad \text{for every } 0 \leq j \leq m. \quad (6.17)$$

If this event fails, the least $0 \leq h \leq m$ with $T^h(x) \leq 2^{q_L(m)}$ is an actual successful first-passage block.

Let x_0 be the first nonendpoint landing in a shell below $S = S_M$. No additional prefix condition is imposed on x_0 . Define the small-rank sequence recursively, writing $x_i \in I_{m_i}$. A rank $m_i < L$ ends the sequence with the required target value. For $m_i \geq L$, the sequence ends if $\text{To}_{L,m_i}(x_i)$ holds. Otherwise, put $h_i = \tau_{2^{q_L(m_i)}}(x_i)$ and $x_{i+1} = T^{h_i}(x_i)$. Each such passage has duration at most m_i , and the next shell rank is at most $q_L(m_i) < m_i$. Thus the sequence ends after at most $S + 1$ blocks.

Let $\mathcal{F}_{M,L}^{\text{low}}$ be the set of starting values $n \in I_M$ whose small-rank sequence ends at a timeout, all its preceding blocks having crossed their thresholds. Define $\text{Fail}_{M,L}^{\text{to}} = \mathcal{F}_M^{\text{ini}} \cup \mathcal{F}_M^{\text{hi}} \cup \mathcal{F}_{M,L}^{\text{low}}$.

Lemma 6.3 (timeout targets and feasible times)

Fix $C > 1$. Uniformly for $L \leq m \leq CL$, for all sufficiently large L ,

$$\boxed{\frac{\#\{x \in I_m : \text{To}_{L,m}(x)\}}{2^m} \ll_{C,K_0} m^{-3/2} 2^{-\kappa_* m}.} \quad (6.18)$$

Moreover, for every sufficiently large shell rank M and every possible first-timeout landing band J_{2^p} , the cumulative times at which that band is reached lie in a finite set $\mathcal{H}_{M,p}^{\text{to}}$ satisfying

$$\#\mathcal{H}_{M,p}^{\text{to}} \ll \sqrt{(M+2) \log(M+2)}. \quad (6.19)$$

Proof Write $x_j = T^j(x)$, let b_i be the parity of x_i , and put

$$s_j = \sum_{i < j} b_i, \quad Z_j = s_j \log 3 - j \log 2.$$

Iteration of the shortcut recurrence gives the exact product

$$x_j = x e^{Z_j} \prod_{\substack{0 \leq i < j \\ x_i \text{ odd}}} \left(1 + \frac{1}{3x_i}\right). \quad (6.20)$$

On a timeout every $x_i > Y := 2^{q_L(m)}$. The logarithm of the product in (6.20) is therefore between zero and $m/(3Y)$. Since $x < 2^{m+1}$ and $x_j > 2^{q_L(m)}$, every prefix satisfies

$$Z_j > -(m+1 - q_L(m)) \log 2 - \frac{m}{3Y}. \quad (6.21)$$

Integer rounding gives

$$m - q_L(m) = \left\lceil \frac{K_0 m}{2L} \right\rceil \leq \frac{CK_0}{2} + 1.$$

Also $q_L(m) \geq m - CK_0/2 - 1$, so $m/(3Y) \leq 1$ uniformly for sufficiently large L . Thus every timeout word satisfies $Z_j > -B$ for every $0 \leq j \leq m$, with the fixed choice

$$B = \left(\frac{CK_0}{2} + 2 \right) \log 2 + 1.$$

Lemma A.3, followed by the exact parity-word bijection of Proposition 2.2, now proves (6.18).

It remains to identify the target set and its possible passage times. Suppose the first timeout occurs at a low starting value x_i . Let H_i denote the total shortcut time from n to x_i , including the preceding large-rank blocks. The value x_i is the landing of the preceding successful block; let p be that block's threshold rank. The identity in Lemma 5.1 collapses the preceding threshold chain, making H_i the direct first-passage time of the starting value n at or below 2^p , with

$$x_i \in (2^{p-1}, 2^p].$$

The upper endpoint cannot time out, since a power of two crosses every lower dyadic threshold by repeated halving. Hence $x_i \in I_{p-1}$. Since the sequence has no further blocks below rank L , and timeouts occur only below the switch,

$$L + 1 \leq p \leq S.$$

Thus the first-timeout target in the landing band J_{2^p} is

$$\mathcal{C}_{L,p}^{\text{to}} = \{y \in (2^{p-1}, 2^p) \cap \mathbb{N} : \text{To}_{L,p-1}(y)\}, \quad \frac{|\mathcal{C}_{L,p}^{\text{to}}|}{2^p} \ll p^{-3/2} 2^{-\kappa_* p}. \quad (6.22)$$

The last estimate is (6.18) with $m = p - 1$, after changing only its fixed constant and absorbing the harmless factor between 2^{p-1} and 2^p .

For a fixed high entry threshold, Lemma 6.1 places the cumulative high entry times in an interval of length $O(\sqrt{M \log(M+2)})$. The possible entry thresholds vary over only $O(S)$ ranks, shifting the common center by $O(S)$. Low ranks decrease strictly, there are at most $S + 1$ successful low stages, and each duration is at most S ; hence

$$\sum_{\text{successful low stages}} h_i \leq S(S+1). \quad (6.23)$$

Taking the Minkowski sum enlarges the high interval by $O(S^2)$. Since $S = O(\log(M+2))$, this is $o(\sqrt{M \log M})$, proving (6.19). $\square$

Assembly of the shell estimate

The times k, ℓ below are constructed witnessing times. They need not be the first times below a prescribed target.

Proposition 6.4 (critical shell estimate with time and relative-height bounds)

Fix constants $0 < c_- \leq c_+ < \infty$, $K_0 > 0$, and $N > 0$. Let L_M be an integer terminal-rank sequence satisfying, eventually,

$$c_- \log(M+2) \leq L_M \leq c_+ \log(M+2),$$

and let $R_M \geq 4$ be any prescribed real sequence. The high parameters and switch constant may be chosen so that there are fixed constants C, K such that, for all sufficiently large M ,

$$\boxed{\frac{\#\text{Fail}_{M,L_M,R_M}^{\text{str}}}{2^M} \leq C \left[\sqrt{\frac{(M+2) \log(M+2)}{L_M+1}} 2^{-\kappa_* L_M} + (M+2)^{-N} + R_M^{-1} \right]}. \quad (6.24)$$

Outside this failure set, every $n \in I_M$ has corresponding shortcut and unaccelerated times k, ℓ and one common value y such that

$$T^k(n) = \text{Col}^\ell(n) = y < 2^{L_M}, \quad (6.25)$$

$$|k - c_* \log n| + \left| \ell - \frac{3}{2} c_* \log n \right| \leq K \sqrt{\log n \log \log n}, \quad (6.26)$$

and

$$\max_{j \leq k} T^j(n) \leq n R_M, \quad \max_{i \leq \ell} \text{Col}^i(n) \leq n R_M. \quad (6.27)$$

The constants C, K depend only on the fixed data and not on the detailed sequences L_M, R_M ; the lower bound on M also includes the point from which the displayed bounds on L_M hold. The failure set Fail^{str} is the union of $\text{Fail}_{M, L_M}^{\text{to}}$ with the initial multiplicative-maximum exception used for (6.27).

Proof Estimating $\mathcal{F}_M^{\text{ini}} \cup \mathcal{F}_M^{\text{hi}}$.

Choose a fixed $r \in (a_0, 1)$ and

$$a_0 < r < 1, \quad 0 < \tau < \min\{r - a_0, (1 - r)/r, a_0, 1\}. \quad (6.28)$$

Thus $r(1 + \tau) < 1$. Use $r_{\text{hi}} = r$, choose D_{hi} so that $c_0 D_{\text{hi}}^2 > N + 5$, and then choose C_{sw} large enough that

$$C_{\text{sw}} \log 2 > N + 5, \quad C_{\text{sw}} \geq D_{\text{hi}}^2 / \tau^2, \quad r C_{\text{sw}} > 2c_+. \quad (6.29)$$

With $S_M = \lceil C_{\text{sw}} \log(M + 2) \rceil$, one has $L_M < S_M < M$ eventually. The final high threshold entering below the switch is still at least $L_M + 1$, since its starting shell rank is at least S_M and its threshold rank is at least $r S_M - 1$. The proportion of $\mathcal{F}_M^{\text{ini}} \cup \mathcal{F}_M^{\text{hi}}$ is then $O((M + 2)^{-N})$. Indeed, Proposition 6.2, the direct-passage identity, (6.5), and (6.11) bound their total proportion by

$$d_{\text{hi}}(M) + O\left(\sqrt{M \log M} M^2 d_{\text{hi}}(M)\right),$$

where the transported contribution is estimated as follows. For a high target rank q , let B_q be its bad landing target. Every producing block has duration at most its starting shell rank m and $q + 1 > r_{\text{hi}} m$, so the bound in Lemma 5.2 gives $E_{2q}(n) < D_q := (q + 2)/r_{\text{hi}}$. Take M sufficiently large that $D_q/2^q = (q + 2)/(r_{\text{hi}} 2^q) \leq 1/3$ for every high target rank. Thus (6.11) applies. Since $|B_q| \leq 2^q d_{\text{hi}}(M)$, it gives, after division by 2^M , a band proportion $\ll \sqrt{M \log M} (2^{q-M} + D_q) d_{\text{hi}}(M) \ll \sqrt{M \log M} M d_{\text{hi}}(M)$. Summing over at most $M + 1$ possible ranks proves the displayed total. The exponent in $d_{\text{hi}}(M)$ exceeds $N + 5$ by (6.29), so the polynomial transport factors are absorbed with room to spare.

Estimating $\mathcal{F}_{M, L_M}^{\text{low}}$.

Fix $r_* = 1/2 < r_{\text{hi}}$, which is permitted because $r_{\text{hi}} > a_0 > 1/2$. Eventually $r_* < r_{L_M}$. Every successful high or low block from starting shell rank m to threshold rank q then has duration at most m and $q + 1 > r_* m$. The bound in Lemma 5.2 therefore gives, at a first-timeout target rank p ,

$$E_{2p}(n) < \frac{p + 2}{r_*}.$$

Since $L_M \asymp \log M$ and $S_M = C_{\text{sw}} \log(M + 2) + O(1)$, a fixed $C > 1$, depending only on the fixed bounds c_- and C_{sw} , satisfies $p - 1 \leq S_M - 1 \leq C L_M$ eventually; for example, any

$$C > \max\left\{1, \frac{2C_{\text{sw}}}{c_-}\right\}$$

works for all sufficiently large M . Thus Lemma 6.3 applies uniformly to every timeout band $L_M + 1 \leq p \leq S_M$. For all sufficiently large L_M , this loss satisfies the smallness condition (4.7).

Apply the restricted-time form (6.11) with target $\mathcal{C}_{L_M, p}^{\text{to}}$ and time set $\mathcal{H}_{M, p}^{\text{to}}$. Equations (6.19) and (6.22) give

$$\frac{\#\{\text{first timeout in band } p\}}{2^M} \ll \sqrt{M \log M} (p+1) p^{-3/2} 2^{-\kappa_* p}.$$

The geometric tail satisfies

$$\sum_{p=L_M+1}^{S_M} (p+1) p^{-3/2} 2^{-\kappa_* p} \ll L_M^{-1/2} 2^{-\kappa_* L_M}. \quad (6.30)$$

This proves the first term of (6.24). The contribution from $\mathcal{F}_M^{\text{ini}} \cup \mathcal{F}_M^{\text{hi}}$ supplies the second term.

Shortcut and unaccelerated time estimates. Write $g = 1 - a_0$, and let H be the cumulative time at the last high landing, whose threshold rank q_{ent} is $O(S_M)$. Equation (6.10) and the geometric decay of the high blocks' starting shell ranks give

$$|gH - ((M+1) - q_{\text{ent}})| \ll \sqrt{M \log(M+2)}.$$

The remaining low part costs at most $S_M(S_M+1)$, and an endpoint-halving tail costs at most S_M . Both are absorbed by the same error. Since $\log n = M \log 2 + O(1)$, the resulting shortcut witness satisfies

$$k = c_* \log n + O(\sqrt{\log n \log \log n}). \quad (6.31)$$

Let q_f be the final threshold rank. The small-rank sequence and the endpoint-halving alternative give

$$L_M - K_0/2 - 1 \leq q_f \leq L_M. \quad (6.32)$$

Indeed, if the final block is low, its starting shell rank is at least L_M , so $q_f = \lfloor r_{L_M} m \rfloor \geq L_M - K_0/2 - 1$. A threshold rank above L_M cannot be final, whereas a nonendpoint landing at threshold rank L_M already has shell rank $L_M - 1$; in the endpoint-halving case, $q_f = L_M - 1$. This accounts for every successful terminal case. Treating every endpoint-halving step as a one-step first passage, the complete successful chain is the direct first passage of the starting value n to its last threshold. If s is the number of odd shortcut steps, the exact reverse product of Section 4 gives

$$n = \frac{2^k y}{3^s} P, \quad 1 - \frac{q_f + 2}{r_* 2^{q_f}} \leq P \leq 1. \quad (6.33)$$

Hence $\log P = O(L_M 2^{-L_M})$, while $0 \leq \log y \leq L_M \log 2$. Taking logarithms and using (6.31) yields

$$s - \frac{k}{2} = \frac{gk \log 2 - \log n + \log y + \log P}{\log 3} = O(\sqrt{\log n \log \log n}).$$

Since $\text{Col}^{k+s}(n) = T^k(n) = y$, this proves the two-sided unaccelerated-time estimate in (6.26).

Relative height. It remains to establish the common relative-height bound. For the first M parity prefixes of the starting value n , put $F_j = 3^{s_j}/2^j$. Lemma A.4 and the exact shell parity law show that, outside a set of proportion at most $4/R_M$,

$$\max_{0 \leq j \leq M} F_j \leq R_M/4. \quad (6.34)$$

During the first high block, the multiplicative correction in (6.20) is at most $\exp(M/2^{\lfloor r_M \rfloor + 1}) \leq 2$ for all sufficiently large M . Thus every shortcut state there is at most $nR_M/2$, and every intervening

unaccelerated odd state is at most twice the following shortcut state. Every later high block starts below $2^{\lfloor rM \rfloor}$, so the bound $T^j(x) \leq x^{1+\tau}$ from (3.8) and $r(1+\tau) < 1$ keep its shortcut states below n and its unaccelerated states below $2n$. Low blocks and endpoint halving remain below $2^{2S_M+2} = M^{O(1)} < n$ eventually. Since $R_M \geq 4$, these estimates prove (6.27). Adding the initial maximum exception contributes the final term R_M^{-1} in (6.24). $\square$

Deductions of the main results

Proof of Theorem 1.1 For each shell I_M , take the minima

$$u_M = \min_{n \in I_M} U(n), \quad v_M = \min_{n \in I_M} R(n).$$

Both sequences tend to infinity. Put

$$w_M = \min\{u_M, \log(M+2)\}, \quad L_M = \lfloor A_{\text{FP}} \log_2(M \log 2) + \log_2 w_M \rfloor, \quad R_M = \min\{v_M, M+2\}. \quad (6.35)$$

For all sufficiently large M , one has $R_M \geq 4$ and $L_M \asymp \log M$, with fixed upper and lower constants. Moreover, for $n \in I_M$,

$$2^{L_M} \leq (M \log 2)^{A_{\text{FP}}} w_M \leq (\log n)^{A_{\text{FP}}} U(n), \quad R_M \leq R(n).$$

Since $\kappa_* A_{\text{FP}} = 1/2$, the first term of (6.24) satisfies

$$\sqrt{\frac{(M+2) \log(M+2)}{L_M+1}} 2^{-\kappa_* L_M} \ll w_M^{-\kappa_*} = o(1).$$

Apply Proposition 6.4 with $N = 1$. The other two terms in its exceptional proportion also tend to zero. Lemma 2.1 therefore shows that the union of these exceptional sets has natural density zero. On its complement, the target, time estimates, and both orbit bounds are exactly those asserted in Theorem 1.1. The finitely many earlier integers can be omitted from the density-one set. $\square$

Proof of Corollary 1.2 Fix $A > A_{\text{FP}}$ and put $q = \kappa_* A - 1/2 > 0$. Choose

$$L_M = \lfloor A \log_2(M \log 2) \rfloor, \quad R_M = (M \log 2)^q, \quad (6.36)$$

and fix $N > q$. For sufficiently large M , these sequences satisfy the hypotheses of Proposition 6.4. For $n \in I_M$, they give $2^{L_M} \leq (\log n)^A$ and $R_M \leq (\log n)^q$. The first term of (6.24) is

$$O_A(M^{1/2-\kappa_* A}) = O_A(M^{-q}),$$

and the other two terms have this order or smaller. Thus the exceptional proportion on I_M is $O_A(M^{-q})$. The quantitative assertion of Lemma 2.1 gives $O_A(X/(\log X)^q)$ exceptions in $[1, X]$. The time-error constant in Proposition 6.4 depends only on the fixed data, which here depend only on A . Enlarging the exceptional-count constant absorbs the finitely many integers below n_A . $\square$

7. The cost of the orbit bound

The exceptional count in Corollary 1.2 has a matching lower bound. This comes from the prescribed height of the orbit, not from the difficulty of reaching the target alone.

Proposition 7.1 (sharpness of the joint exceptional count)

Fix $A > A_{\text{FP}}$, put $q = \kappa_* A - 1/2$, and let E_A be the set of integers for which the complete conclusion of Corollary 1.2 fails, with its constants K_A, n_A . Then

$$\#(E_A \cap [1, X]) = \Theta_A\left(\frac{X}{(\log X)^q}\right). \quad (7.1)$$

Proof The upper bound is Corollary 1.2. For the lower bound, work on a large shell I_M and put

$$h = \lfloor M/2 \rfloor, \quad u_M = ((M+1) \log 2)^q.$$

For all sufficiently large M ,

$$((M+1) \log 2)^A < 2^{M-h}, \quad \log u_M \leq \varepsilon_0 h \log 3,$$

where $\varepsilon_0 > 0$ is defined in Lemma A.5. Every shortcut step is at least a halving, so, for $n \in I_M$ and $0 \leq j \leq h$,

$$T^j(n) \geq 2^{-j} n \geq 2^{M-h} > (\log n)^A.$$

Thus no value below the target can be reached in the first h steps.

For the first h parity letters, set $F_j = 3^{s_j(n)}/2^j$. If $\max_{j \leq h} F_j > u_M$, the positive affine correction in Proposition 2.3 gives, at some $j \leq h$,

$$T^j(n) \geq n F_j > n u_M \geq n (\log n)^q.$$

Every later time reaching the target therefore violates the shortcut orbit bound. Such an integer is in E_A , regardless of the time or unaccelerated-orbit requirements.

Since $h \leq M$, each length- h parity word occurs exactly 2^{M-h} times on I_M , by Proposition 2.2. Lemma A.5 now yields

$$\frac{\#(E_A \cap I_M)}{2^M} \geq \frac{2}{3u_M} \left(1 - \frac{K_{\max}}{h}\right) \gg_A M^{-q}.$$

For an arbitrary large X , choose M with $2^{M+1} \leq X < 2^{M+2}$. The complete shell I_M lies in $[1, X]$, has more than $X/4$ integers, and satisfies $M \asymp \log X$. Retaining this shell gives the required lower bound for every sufficiently large X . $\square$

This argument proves sharpness only for the joint conclusion with its stated orbit bound. It gives no lower bound for descent failures when that bound is removed, and does not establish optimality of A_{FP} .

8. Formal verification

Theorem 1.1, with the shared value, both time estimates, and both orbit bounds, is formalized in Lean 4 with Mathlib and registered as PALOMAR-2026-08-26-000005, version 4 [17], using only `propext`, `Classical.choice`, and `Quot.sound`. The source is archived at doi:10.5281/zenodo.23043600. The related registered fixed-exponent upper bound permits a constant $C > 0$ in the target $C(\log n)^A$ and uses the orbit bound $n(\lfloor \log_2 n \rfloor + 2)^q$; the registered lower bound counts failure of the shortcut target-and-height conditions with those normalizations. These differ from the printed formulations of Corollary 1.2 and Proposition 7.1. The supplement `formalization-supplement.md` gives the exact statements, proof locations, and mathematical definitions used in the code.

Research and software disclosure

OpenAI Codex assisted with mathematical exploration, Lean formalization, and manuscript preparation. The author reviewed the AI-assisted material and takes full responsibility for the manuscript.

Appendix A. Finite ballot estimates

This appendix proves the finite-word estimate used in Lemma 6.3. Probability notation below denotes an exact finite product weighting of words.

Lemma A.1 (centered finite survival)

Let $X_1, \dots, X_n$ be independent copies of a random variable on a finite weighted alphabet such that

$$\mathbb{E}X_1 = 0, \quad |X_1| \leq 1, \quad \mathbb{E}X_1^2 = v > 0.$$

Put $S_0 = 0$ and $S_j = \sum_{i=1}^j X_i$. For every $h \geq 0$ and integer $n \geq 1$,

$$\begin{aligned} \mathbb{P}\{S_j \geq -h \text{ for } 0 \leq j \leq n\} &\leq \frac{h+1}{\sqrt{n}+h+1} + \frac{(h+1)(\sqrt{n}+1)}{vn} \\ &\leq \left(1 + \frac{2}{v}\right) \frac{h+1}{\sqrt{n}}. \end{aligned} \tag{A.1}$$

Proof Set $H = \sqrt{n}$, and let τ be the smaller of n and the first index at which $S_j < -h$ or $S_j \geq H$. If no such index occurs, put $\tau = n$. The stopped value has the finite expansion

$$Y = S_\tau = \sum_{j=1}^n \mathbf{1}_{\{\tau \geq j\}} X_j.$$

The indicator in the j -th summand is determined by the first $j-1$ letters. Independence and centering therefore give $\mathbb{E}Y = 0$. Conditioning on the letters before the later increment makes every cross term in $\mathbb{E}Y^2$ vanish, while the diagonal terms give

$$\mathbb{E}Y^2 = v \sum_{j=1}^n \mathbb{P}(\tau \geq j) = v\mathbb{E}\tau. \tag{A.2}$$

Bounded overshoot gives $-h-1 \leq Y \leq H+1$. Hence

$$0 \leq \mathbb{E}[(Y+h+1)(H+1-Y)] = (h+1)(H+1) - v\mathbb{E}\tau,$$

so $\mathbb{E}\tau \leq (h+1)(H+1)/v$. If $U = \{Y \geq H\}$, then $Y \geq -h-1$ off U , and $\mathbb{E}Y = 0$ gives

$$\mathbb{P}(U) \leq \frac{h+1}{H+h+1}.$$

A path surviving above $-h$ through time n either stops at the upper boundary or has $\tau = n$. Markov's inequality for τ therefore gives

$$\mathbb{P}(\text{survival}) \leq \mathbb{P}(U) + \mathbb{P}(\tau = n) \leq \frac{h+1}{H+h+1} + \frac{\mathbb{E}\tau}{n},$$

which proves (A.1). $\square$

Lemma A.2 (uniform binomial atom)

For fixed $p \in (0, 1)$, there is $C_p < \infty$ such that, for every integer $n \geq 1$,

$$\sup_{k \in \mathbb{Z}} \binom{n}{k} p^k (1-p)^{n-k} \leq C_p n^{-1/2}, \quad (\text{A.3})$$

where masses outside $0 \leq k \leq n$ are zero.

Proof The ratio of consecutive positive masses is

$$\frac{n-k}{k+1} \frac{p}{1-p}.$$

Thus a mode may be chosen as $k_0 = \lfloor (n+1)p \rfloor$, with $|k_0 - np| \leq 1$. For all sufficiently large n , $u = k_0/n$ remains in a fixed compact subinterval of $(0, 1)$. The elementary factorial Stirling bounds give

$$\binom{n}{k_0} p^{k_0} (1-p)^{n-k_0} \leq \frac{C}{\sqrt{nu(1-u)}} e^{-nD(u||p)} \leq C_p n^{-1/2},$$

because $D(u||p) \geq 0$. Enlarge C_p for the finitely many smaller values of n . Every other atom is no larger than a mode. $\square$

Lemma A.3 (fair-word mass above a fixed barrier)

For a binary word $b_0, \dots, b_{m-1}$, put

$$s_j = \sum_{i < j} b_i, \quad Z_j = s_j \log 3 - j \log 2.$$

For every fixed $B > 0$, there is $C_B < \infty$ such that, for every integer $m \geq 1$,

$$2^{-m} \#\{b : Z_j > -B \text{ for every } 0 \leq j \leq m\} \leq C_B m^{-3/2} 2^{-\kappa_* m}. \quad (\text{A.4})$$

The constant may be chosen uniformly for B in a bounded positive interval.

Proof Give a letter 1 weight p_* and a letter 0 weight $1 - p_*$, and write $\mathbb{P}_*$ for the resulting finite product law. The increments of Z then have mean zero, absolute value less than one, and variance

$$v = p_*(1-p_*)(\log 3)^2 = \log 2 \log(3/2) > 0.$$

The negated increments have the same properties. Put

$$\lambda = \frac{\log(p_*/(1-p_*))}{\log 3} > 0.$$

For every individual word, direct simplification gives

$$\frac{2^{-m}}{p_*^{s_m} (1-p_*)^{m-s_m}} = 2^{-\kappa_* m} e^{-\lambda Z_m}. \quad (\text{A.5})$$

For $m \geq 8$, split the word into a first block of length $a = \lfloor m/4 \rfloor$, a middle block of length $b = m - 2a$, and a last block of length a . Let W be the event on the left of (A.4), and partition its terminal values into

$$Z_m + B \in [q, q+1), \quad q = 0, 1, 2, \dots$$

On W , the first block stays above $-B$. On W in bin q , the reversed, negated last block has partial sums

$$Z_{m-j} - Z_m > -(q+1), \quad 0 \leq j \leq a.$$

These two necessary outer-block events depend on disjoint letters. Applying Lemma A.1 to each bounds their product probability by

$$\frac{C(B+1)(q+2)}{a}.$$

After the two outer blocks are fixed, the terminal-bin condition asks that the middle-block sum lie in an interval of length one. That sum is $K \log 3 - b \log 2$, where K has the $\text{Bin}(b, p_*)$ law. Its lattice spacing is $\log 3 > 1$, so the interval contains at most one atom. Lemma A.2 bounds the conditional probability by $Cb^{-1/2}$, uniformly in the outer blocks. Since $a \geq m/8$ and $b \geq m/2$,

$$\mathbb{P}_*(W, Z_m + B \in [q, q+1)) \leq C(B+1)(q+2)m^{-3/2}. \quad (\text{A.6})$$

The power $m^{-3/2}$ has three square-root sources: survival of the first outer block, survival of the reversed-negated last outer block, and the single middle-block binomial atom. Only the two disjoint outer-block events have been multiplied here; no independence of the full survival and terminal events is asserted.

On bin q , $e^{-\lambda Z_m} \leq e^{\lambda B} e^{-\lambda q}$. Summing (A.6) with the weights in (A.5) yields

$$2^{-m} \#W \leq C(B+1)e^{\lambda B} \left(\sum_{q \geq 0} (q+2)e^{-\lambda q} \right) m^{-3/2} 2^{-\kappa_* m}.$$

The series converges, and the finitely many $m < 8$ are absorbed into the constant. The same expression is bounded uniformly when B ranges over a bounded positive interval. $\square$

This is a ballot-type estimate. General ballot estimates for finite-variance walks are classical; see Addario-Berry and Reed [16]. Their theorem is not a premise here: the finite argument above supplies the precise estimate used in Section 6.

Lemma A.4 (finite multiplicative maximum)

For fair binary words of length M , put $F_j = 3^{s_j}/2^j$. For every $u > 0$,

$$2^{-M} \#\{b : \max_{0 \leq j \leq M} F_j > u\} \leq u^{-1}. \quad (\text{A.7})$$

Proof Given a prefix, the next multiplicative factor is $1/2$ or $3/2$, whose fair average is one. Stop at the first index for which $F_j > u$, or at M if there is none; if $u < 1$, stop at zero. Repeatedly replacing each pair of children by its parent in this finite word tree shows that the stopped terminal value has mean one. It is nonnegative everywhere and exceeds u on the crossing event. Hence $1 \geq u \mathbb{P}(\text{crossing})$, proving (A.7). $\square$

Lemma A.5 (reciprocal finite multiplicative maximum)

For an integer $h \geq 1$ and $u \geq 1$, let

$$A_h(u) = \left\{ \max_{0 \leq j \leq h} F_j > u \right\}, \quad \varepsilon_0 = \frac{3/4 - p_*}{2} > 0, \quad K_{\max} = \frac{3}{16\varepsilon_0^2}.$$

If $\log u \leq \varepsilon_0 h \log 3$, then

$$\frac{2}{3u} \left(1 - \frac{K_{\max}}{h}\right) \leq 2^{-h} \#A_h(u) \leq \frac{1}{u}. \quad (\text{A.8})$$

For small h , a negative lower bound is understood as vacuous.

Proof Stop each fair word at its first crossing of u , or at h if it does not cross. The finite-tree identity used in Lemma A.4 gives $\mathbb{E}F_\sigma = 1$. At a crossing, the previous value is at most u , and only an upward multiplier can cross, so $u < F_\sigma \leq 3u/2$. Consequently

$$1 \leq \frac{3}{2}u \mathbb{P}_{1/2}(A_h(u)) + \mathbb{E}_{1/2}[F_h; A_h(u)^c]. \quad (\text{A.9})$$

For each individual word there is the exact change of finite word weights

$$2^{-h}F_h = \left(\frac{3}{4}\right)^{s_h} \left(\frac{1}{4}\right)^{h-s_h}. \quad (\text{A.10})$$

Therefore the final term in (A.9) equals $\mathbb{P}_{3/4}(A_h(u)^c)$, and is at most $\mathbb{P}_{3/4}(F_h \leq u)$.

Under the latter finite product weighting, s_h has mean $3h/4$ and variance $3h/16$. The identity $\log F_h = s_h \log 3 - h \log 2$, the definition $p_* = \log_3 2$, and the hypothesis on u imply

$$F_h \leq u \implies s_h \leq p_* h + \frac{\log u}{\log 3} \leq \left(\frac{3}{4} - \varepsilon_0\right) h.$$

Chebyshev's inequality bounds this probability by $K_{\max}/h$. Substitution in (A.9) proves the lower bound; Lemma A.4 gives the upper bound. $\square$

References

- [1] R. E. Crandall, *On the “3x+1” problem*, Math. Comp. **32** (1978), no. 144, 1281–1292, doi:10.1090/S0025-5718-1978-0480321-3.
- [2] C. J. Everett, *Iteration of the number-theoretic function $f(2n) = n$, $f(2n+1) = 3n+2$* , Adv. Math. **25** (1977), no. 1, 42–45, doi:10.1016/0001-8708(77)90087-1.
- [3] M. Inslemann, *An approximation of the Collatz map and a lower bound for the average total stopping time*, arXiv:2402.03276v3 (2024), doi:10.48550/arXiv.2402.03276.
- [4] A. V. Kontorovich and J. C. Lagarias, *Stochastic models for the $3x+1$ and $5x+1$ problems and related problems*, in *The Ultimate Challenge: The $3x+1$ Problem* (J. C. Lagarias, ed.), American Mathematical Society, Providence, RI, 2010, pp. 131–188.
- [5] I. Korec, *A density estimate for the $3x+1$ problem*, Math. Slovaca **44** (1994), no. 1, 85–89, DML-CZ record.
- [6] I. Krasikov and J. C. Lagarias, *Bounds for the $3x+1$ problem using difference inequalities*, Acta Arith. **109** (2003), 237–258, doi:10.4064/aa109-3-4.
- [7] J. C. Lagarias, *The $3x+1$ problem and its generalizations*, Amer. Math. Monthly **92** (1985), no. 1, 3–23, doi:10.2307/2322189.
- [8] J. C. Lagarias and A. Weiss, *The $3x+1$ problem: two stochastic models*, Ann. Appl. Probab. **2** (1992), no. 1, 229–261, doi:10.1214/aoap/1177005779.

- [9] I. A. Shaik, *Quantitative Collatz Descent to Stretched-Logarithmic Scale in Natural Density, with a Lean 4 Formalization*, version 2.0.2, preprint (2026), doi:10.5281/zenodo.21851173.
- [10] T. Tao, *Almost all orbits of the Collatz map attain almost bounded values*, Forum Math. Pi **10** (2022), e12, doi:10.1017/fmp.2022.8.
- [11] R. Terras, *A stopping time problem on the positive integers*, Acta Arith. **30** (1976), 241–252.
- [12] R. Terras, *On the existence of a density*, Acta Arith. **35** (1979), 101–102, doi:10.4064/aa-35-1-101-102.
- [13] G. J. Wirsching, *The Dynamical System Generated by the $3n + 1$ Function*, Lecture Notes in Mathematics, vol. 1681, Springer-Verlag, Berlin, 1998, doi:10.1007/BFb0095985.
- [14] J. Allikvere, *Almost all Collatz orbits attain almost bounded values in natural density*, version 2, preprint (22 July 2026), doi:10.5281/zenodo.21499244.
- [15] L. Mazur, *Natural-density almost-bounded Collatz orbits in logarithmic time*, version 2, preprint (21 July 2026), public manuscript.
- [16] L. Addario-Berry and B. A. Reed, *Ballot theorems for random walks with finite variance*, arXiv:0802.2491v2 (2008), doi:10.48550/arXiv.0802.2491.
- [17] I. A. Shaik, *Polylogarithmic Descent for Almost All Collatz Orbits in Natural Density*, Lean 4 formalization, Palomar Registry, PALOMAR-2026-08-26-000005, version 4 (2026), registry record.