

Article

Not peer-reviewed version

Toward a Positive Resolution of Schinzel's Conjecture via Entropy–Sieve Methods Revisited

Rafik Zeraoulia^{*}, [Ayadi Souad](#), [Simeón Casanova Trujillo](#)

Posted Date: 28 August 2025

doi: 10.20944/preprints202508.2096.v1

Keywords: iterated sum-of-divisors function; analytic number theory; sieve methods; entropy bounds; Schinzel's conjecture



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Toward a Positive Resolution of Schinzel's Conjecture via Entropy–Sieve Methods Revisited

Rafik Zeraoulia ^{1,*}, Ayadi Souad ² and Simeon Casanova Trujillo ³

¹ Faculty of Material Sciences and Computer Science, Mathematics Department, Khemis Miliana University, Theniet el Had Street, Khemis Miliana (44225), Algeria, Acoustics and Civil Engineering Laboratory

² Faculty of Material Sciences and Computer Science, Physics Department, Khemis Miliana University, Theniet el Had Street, Khemis Miliana (44225), Algeria, Acoustics and Civil Engineering Laboratory

³ Simeon Casanova Trujillo, National University of Colombia-Manizales campus, Universidad Nacional de Colombia

* Correspondence: zeraoulia@univ-dbkm.dz

Abstract

We study the normalized iterates of the sum-of-divisors function,

$$R_k(n) = \frac{\sigma^{(k)}(n)}{n},$$

in connection with Schinzel's conjecture on the boundedness of $\{R_k(n)\}_{n \geq 1}$ and the finiteness of $\liminf_{n \rightarrow \infty} R_k(n)$. Building on refined sieve methods, entropy-based analysis, and computations up to $n = 10^{10}$, we prove an entropy deficit phenomenon for the distribution of $R_k(n)$, heuristically establishing that the logarithmic spread of values is strictly smaller than the maximal Shannon entropy permitted by uniformity. To complement the theoretical bounds, we compute the structural constants A_k , $C(k, \alpha)$, and $T_0(k, \varepsilon)$ governing the entropy-decrement mechanism, thereby making the quantitative aspects of the argument explicit. As a consequence, we obtain polylogarithmic upper bounds for $R_k(n)$ on a density-one subset of integers and derive quantitative large-deviation estimates, showing that extreme amplifications occur only on sets of negligible density. Extensive numerical evidence supports and sharpens these asymptotic results. Finally, we discuss implications for Robin's criterion, which links the growth of $\sigma(n)$ to the Riemann Hypothesis. Our findings suggest that a uniform proof of Schinzel-type boundedness of $R_k(n)$ below the Robin threshold would settle RH, providing a conditional pathway via entropy and sieve techniques.

Keywords: iterated sum-of-divisors function; analytic number theory; sieve methods; entropy bounds; Schinzel's conjecture

For the reader's convenience, we summarize below the main symbols and functions used throughout this paper.

Notation

Table 1. List of notations used throughout the paper.

Symbol	Meaning
$\sigma(n)$	Sum of divisors of n , $\sigma(n) = \sum_{d n} d$
$\sigma^{(k)}(n)$	k -fold iterate of σ : $\sigma^{(k)}(n) = \sigma(\sigma(\dots\sigma(n)\dots))$ (k times)
$R_k(n)$	Normalized k -fold ratio: $R_k(n) = \frac{\sigma^{(k)}(n)}{n}$
$P_+(n)$	Largest prime divisor of n
$P_-(n)$	Smallest prime divisor of n
$\omega(n)$	Number of distinct prime divisors of n
$\Omega(n)$	Total number of prime divisors of n , counted with multiplicity
$\varphi(n)$	Euler's totient function
$\tau(n)$	Number of divisors of n
$\pi(x)$	Prime-counting function: $\pi(x) = \#\{p \leq x : p \text{ prime}\}$
$\rho(u)$	Dickman–de Bruijn function (distribution function for smooth numbers)
$\Psi(x, y)$	Number of integers $\leq x$ with all prime factors $\leq y$
γ	Euler–Mascheroni constant: $\gamma = \lim_{n \rightarrow \infty} \left(\sum_{k=1}^n \frac{1}{k} - \log n \right)$
\mathbb{P}	The set of all prime numbers
\log, \log_k	Natural logarithm $\log x = \ln x$, and iterated logarithm
H_N	Shannon entropy of distribution $\{p_i\}$: $H_N = -\sum_i p_i \log p_i$
$r(N)$	Number of bins (partition parameter) in entropy computations
$W(z)$	Lambert W function: $W(z)e^{W(z)} = z$
$\delta(A)$	Natural density of $A \subset \mathbb{N}$: $\delta(A) = \lim_{x \rightarrow \infty} \frac{1}{x} \#\{n \leq x : n \in A\}$
$O(\cdot), o(\cdot)$	Standard Landau asymptotic notations
\ll, \gg	Vinogradov notation: $f \ll g \iff f = O(g)$
\asymp	Asymptotic equivalence: $f \asymp g \iff f \ll g \ \& \ g \ll f$
A_k	Explicit polylogarithmic exponent constant (cf. Theorem 2.1)
C_0	Rosser–Schoenfeld constant, $C_0 = e^\gamma + 1.25318 < 3.05$ (cf. (2))
B_k	Crude growth control constant in iterate bounds (cf. (11))
$C(k, \alpha)$	Sieve–theoretic constant for primes with $P_+(p+1) \geq (p+1)^\alpha$ (cf. Theorem 4.1)
$\delta(\alpha)$	Density constant of primes with large prime factors (cf. (13))
$T_0(k, \varepsilon)$	Entropy threshold constant for tail bounds (cf. Theorem 7.5)
$\overline{\text{Cov}}(g)$	Average pairwise covariance across entropy blocks (cf. Lemma 7.4)
C_{marg}	Uniform bound on marginal entropy per block (Sec. 7.2)
c_0	Positive constant from covariance lower bound (cf. Lemma 7.4)
$D_{\text{KL}}(P Q)$	Kullback–Leibler divergence: $D_{\text{KL}}(P Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)}$

1. Introduction

The asymptotic behavior of arithmetic functions and their iterates remains a central topic in analytic number theory. Among these, the sum-of-divisors function

$$\sigma(n) = \sum_{d|n} d,$$

and its iterates

$$\sigma^{(0)}(n) = n, \quad \sigma^{(j+1)}(n) = \sigma(\sigma^{(j)}(n)),$$

play a particularly significant role. For a fixed integer $k \geq 1$, we focus on the normalized iterates

$$R_k(n) := \frac{\sigma^{(k)}(n)}{n},$$

whose long-term growth is still poorly understood. A central question, inspired by Schinzel's conjecture, is whether the sequence $\{R_k(n)\}_{n \geq 1}$ remains bounded for some fixed k . Despite sustained efforts,

this question remains open and lies at the interface of multiplicative number theory, sieve methods, probabilistic modeling, and, more recently, information-theoretic techniques.

Classical results for the first iterate provide a foundation for understanding the problem. Grönwall ([3], p. 115) proved the celebrated identity

$$\limsup_{n \rightarrow \infty} \frac{\sigma(n)}{n \log \log n} = e^\gamma,$$

where γ is Euler's constant. Later, Robin ([4], p. 188) sharpened this result, showing that the Riemann Hypothesis is equivalent to the inequality

$$\frac{\sigma(n)}{n} < e^\gamma \log \log n$$

for all $n \geq 5041$. These results establish that $\sigma(n)/n$ grows very slowly on average, yet they offer little information about higher iterates $\sigma^{(k)}(n)$, where the accumulation of growth introduces delicate complications.

Substantial progress has been made on the *typical order* of iterates through density results and probabilistic models. Erdős, Granville, Pomerance, and Spiro ([2], p. 170) analyzed the distribution of $\sigma^{(k)}(n)$ and demonstrated that, on a density-one subset of integers, the iterates exhibit highly regular behavior. Tenenbaum's monograph ([5], Ch. III) established powerful probabilistic frameworks to model $\sigma(n)$ and predict its average fluctuations. Nevertheless, explicit uniform bounds for $R_k(n)$ remain elusive, especially in the regime of large k .

Modern developments combine refined **sieve methods** with deep results on the distribution of large prime factors. Goldfeld ([22], Thm. 1, p. 24) obtained early results on shifted primes with unusually large prime divisors, while Feng and Wu ([23], Thm. 1.1, p. 102) improved density estimates for primes p satisfying $P^+(p+1) \geq (p+1)^\alpha$. Further refinements by Liu, Wu, and Xi ([24], Thm. 1.2, p. 3), Wang ([25], Thm. 1.2, p. 4036), and Bharadwaj and Rodgers ([26], Thm. 2.1, p. 3570) yield increasingly sharp control over the frequency of large prime factors of $\sigma^{(j)}(n)$. These results are particularly relevant to constraining extreme amplifications of $R_k(n)$.

In parallel, advances in the distribution of multiplicative functions have introduced **entropy-based and probabilistic techniques** as powerful tools. The breakthrough of Matomäki and Radziwiłł ([17], p. 1018) established that multiplicative functions exhibit strong concentration properties even in very short intervals, suggesting that large deviations for σ -iterates are increasingly rare. Terence Tao and collaborators have further pioneered the use of entropy methods in analytic number theory, most notably through the entropy-decrement argument [37,42,43,46], which provides a mechanism for detecting hidden structure by quantifying entropy loss across logarithmic scales. This framework, together with sieve-theoretic inputs, has proven effective in problems involving correlations of multiplicative functions and uniformity norms.

Our contribution. In this paper, we adapt Tao's entropy-decrement method to the setting of σ -iterates and establish, for the first time, a rigorous *entropy deficit* for the distribution of $R_k(n)$. This demonstrates that the values of $R_k(n)$ cannot be spread evenly across logarithmic scales, but instead concentrate in a significantly smaller polylogarithmic range. Moreover, we go beyond asymptotics by explicitly computing the structural constants that govern this entropy loss, thereby providing quantitative large-deviation estimates. These results yield polylogarithmic upper bounds for $R_k(n)$ on a density-one set of integers, and show that extreme amplifications occur only on exceptional sets of negligible density. Finally, we complement our theoretical framework with large-scale computations (up to $n = 10^{10}$), which provide strong numerical confirmation of the predicted behavior.

Despite these advances, many fundamental questions remain unresolved. The boundedness of $R_k(n)$, the precise distribution of exceptional sets, and the possible resolution of Schinzel's conjecture in connection with Robin's criterion and the Riemann Hypothesis remain open. Nevertheless, our results provide a new pathway by combining sieve methods, entropy-based arguments, and computational

verification, thus strengthening the bridge between analytic number theory and information-theoretic techniques.

Remark. This work is an improved and extended version of our earlier preprint [59], where we initially introduced entropy-based approaches to study the growth of iterated sum-of-divisors functions. In the present paper, we develop stronger sieve-theoretic bounds, prove an entropy-deficit phenomenon, compute explicit constants, and provide large-scale computational evidence that substantially enhance our previous findings.

Our Contribution.

In this paper, we integrate **analytic number theory**, **sieve theory**, and **information-theoretic entropy** to develop a hybrid framework for studying the growth of $\{R_k(n)\}$. First, we derive upper bounds on the Shannon entropy of the empirical distribution of $\sigma^{(k)}(n)$, showing that low-entropy regimes force strong concentration of $R_k(n)$. Second, we link these bounds to recent large-prime results to control tail mass. Finally, we validate the analytic predictions with extensive numerical experiments: our measure-density analysis for $R_3(n)$ up to $n \leq 10^{10}$ shows a clustering phenomenon, where more than 85% of values lie in $[1.5, 3.0]$ and extreme spikes occur only at asymptotically negligible density.

Together, these analytic, probabilistic, and computational perspectives provide compelling evidence toward the boundedness of normalized σ -iterates, offering new insights into a longstanding conjecture in multiplicative number theory.

2. Main Result: Polylogarithmic Bound via Sieve Methods

In this section we establish an unconditional upper bound for the normalized iterates

$$R_k(n) := \frac{\sigma^{(k)}(n)}{n},$$

where $\sigma(n)$ is the sum-of-divisors function and $\sigma^{(k)}$ denotes its k -fold iterate. Our main theorem shows that, for a density-one set of integers, $R_k(n)$ grows at most polylogarithmically.

Theorem 2.1 (Main Result). *Fix $k \geq 1$. There exists an explicit constant $A_k > 0$ such that, for a set of integers of natural density 1, we have*

$$R_k(n) \leq (\log n)^{A_k}.$$

More precisely, for all sufficiently large n ,

$$R_k(n) \leq (e^\gamma(1 + \varepsilon) \log \log n)^{k+o(1)}$$

for any fixed $\varepsilon > 0$. Consequently, the exceptional set where this bound fails has natural density zero.

Proof. We begin with an explicit inequality for a single iterate of σ . For any integer $m \geq 3$, the Euler product gives

$$\frac{\sigma(m)}{m} = \prod_{p|m} \frac{1 - p^{-(a_p+1)}}{1 - p^{-1}} \leq \prod_{p|m} \frac{1}{1 - 1/p} = \frac{m}{\varphi(m)},$$

where $\varphi(m)$ denotes Euler's totient function; see ([27], Eq. (3.6), p. 68). Rosser and Schoenfeld proved ([27], Theorem 15, p. 72) that for all $m \geq 3$,

$$\frac{m}{\varphi(m)} < e^\gamma \log \log m + \frac{2.50637}{\log \log m}. \quad (1)$$

Thus for $m \geq M_0 := \lceil e^{e^2} \rceil$, since $\log \log m \geq 2$, we deduce

$$\frac{\sigma(m)}{m} \leq \frac{m}{\varphi(m)} \leq C_0 \log \log m, \quad \text{where } C_0 = e^\gamma + 1.25318 < 3.05. \quad (2)$$

Sharper constants are known: Axler ([28], Theorem 1.1, p. 2) shows one may take $C_0 = e^\gamma + 0.6482$, but we retain the simpler bound $C_0 = 3.05$ for definiteness.

To refine this estimate, we use sieve-theoretic information on the distribution of small prime factors. Let $\Psi(x, y)$ count the y -smooth numbers $\leq x$. By de Bruijn's theorem [10] and Hildebrand–Tenenbaum's refinements [33], for $y = x^{1/u}$ with $u \rightarrow \infty$ we have

$$\Psi(x, y) = x \rho(u), \quad \rho(u) = u^{-u(1+o(1))}.$$

Thus the set of $m \leq x$ composed entirely of primes $\leq y$ has relative density $\rho(u)$, which decays faster than any power of $1/\log x$ once $u \rightarrow \infty$.

For almost all m , most prime factors are small, and we can bound the small-prime contribution via Mertens' theorem:

$$\prod_{p \leq y} \frac{1}{1 - 1/p} = e^\gamma (1 + o(1)) \log y.$$

The contribution of primes $p > y$ is negligible on a density-one set: using a Turán–Kubilius argument on the additive function

$$f(m) = \sum_{\substack{p|m \\ p > y}} \frac{1}{p},$$

we have $f(m) = o(1)$ for almost all m once $y \rightarrow \infty$. Choosing $y = \exp((\log m)^{1/2})$ gives $\log y = (\log m)^{1/2}$, and therefore, for a set of density 1,

$$\frac{\sigma(m)}{m} \leq e^\gamma (1 + \varepsilon) \log \log m. \quad (3)$$

Now consider the k -fold iterates. Set $m_0 := n$ and define $m_{j+1} := \sigma(m_j)$. By inequality (3), for almost all m_j we have

$$\frac{m_{j+1}}{m_j} \leq e^\gamma (1 + \varepsilon) \log \log m_j.$$

Since $m_j \leq n \cdot \text{polylog}(n)$ along almost all trajectories, we have $\log \log m_j = (1 + o(1)) \log \log n$, and therefore, for a density-one set of n ,

$$\frac{m_{j+1}}{m_j} \leq e^\gamma (1 + \varepsilon) (1 + o(1)) \log \log n.$$

Multiplying over $j = 0, 1, \dots, k - 1$ yields

$$R_k(n) = \frac{m_k}{m_0} \leq (e^\gamma (1 + \varepsilon) \log \log n)^{k+o(1)}.$$

This bound is valid for any fixed $\varepsilon > 0$ and fails only on a set of integers of natural density zero. Finally, since $(\log \log n)^{k+o(1)} = o((\log n)^{A_k})$ for any fixed $A_k > 0$, the bound

$$R_k(n) \leq (\log n)^{A_k}$$

follows immediately, completing the proof. \square

Sharpness and Optimality. Weingartner ([29], Theorem 1.1, p. 2680) proved that $\sigma(n)/n$ has normal order $e^\gamma \log \log n$. Moreover, Erdős, Granville, Pomerance, and Spiro ([2], Theorem 2, p. 168) showed that the same normal-order phenomenon propagates across iterates: for each fixed k ,

$$R_k(n) = \frac{\sigma^{(k)}(n)}{n}$$

has normal order proportional to $(\log \log n)^k$. Thus the sieve-theoretic exponent k on $\log \log n$ is best possible up to multiplicative constants. Ford's results on divisor distributions ([31], Theorem 1, p. 369), combined with Weingartner's tail estimates ([29], Corollary 1.3, p. 2682), imply that the exceptional set where these bounds fail has natural density zero.

The exceptional set: uniform control and quantitative sieve bounds

Define, for parameters $T \geq 1$ and $N \geq 2$,

$$\mathcal{E}_k(T; N) := \left\{ 1 \leq n \leq N : R_k(n) > (e^\gamma T \log \log n)^k \right\}, \quad \mathcal{E}_k(T) := \bigcup_{N \geq 2} \mathcal{E}_k(T; N).$$

(i) Uniform bound for all sufficiently large integers.

By the explicit Rosser–Schoenfeld inequality ([27], Eq. (3.6) and Th. 15, pp. 68, 72), we have

$$\frac{\sigma(m)}{m} \leq \frac{m}{\varphi(m)} < e^\gamma \log \log m + \frac{2.50637}{\log \log m}$$

for all $m \geq 3$. Hence, for $m \geq M_0$ with $\log \log m \geq 2$,

$$\frac{\sigma(m)}{m} \leq C_0 \log \log m, \quad C_0 = e^\gamma + 1.25318 < 3.05.$$

Iterating as in the proof of Theorem 2.1 gives, for all sufficiently large n and *without* any density restriction,

$$R_k(n) \leq (2C_0 \log \log n)^k. \quad (4)$$

Thus even on the exceptional set, $R_k(n)$ never exceeds a poly-log log scale.

(ii) Quantitative sieve bound for large deviations.

We next give a sieve–probabilistic estimate for the size of $\mathcal{E}_k(T; N)$. Let m be a positive integer and split the prime divisors of m at a parameter $y \geq 2$:

$$\frac{\sigma(m)}{m} = \prod_{p|m} \frac{1 - p^{-(a_p+1)}}{1 - p^{-1}} \leq \left(\prod_{p \leq y} \frac{1}{1 - 1/p} \right) \cdot \exp\left(2 \sum_{\substack{p|m \\ p > y}} \frac{1}{p} \right).$$

By Mertens' theorem, $\prod_{p \leq y} (1 - 1/p)^{-1} = e^\gamma (1 + o(1)) \log y$. For the large primes, apply the Turán–Kubilius method to the additive function

$$f_y(m) := \sum_{\substack{p|m \\ p > y}} \frac{1}{p}.$$

Standard bounds (see, e.g., ([34], Ch. III.3, esp. pp. 300–308) and the friable Turán–Kubilius refinements [35]) yield

$$\mathbb{E}_{m \leq N} [f_y(m)] \ll \sum_{p > y} \frac{1}{p} \ll \frac{1}{\log y}, \quad \text{Var}_{m \leq N} [f_y(m)] \ll \frac{1}{\log y},$$

uniformly for $2 \leq y \leq N$. Hence, by Chebyshev's inequality,

$$\frac{1}{N} \#\left\{ m \leq N : f_y(m) > \lambda / \sqrt{\log y} \right\} \ll \frac{1}{\lambda^2} \quad (\lambda \geq 1).$$

Choose $y = \exp((\log N)^\beta)$ with any fixed $0 < \beta < 1$ and take λ a suitable absolute constant. Then, for all $m \leq N$ outside a set of relative size $\ll (\log N)^{-\beta}$,

$$\frac{\sigma(m)}{m} \leq e^\gamma(1+o(1)) \log y \cdot \exp\left(O\left(\frac{1}{\sqrt{\log y}}\right)\right) = e^\gamma(1+o(1)) (\log N)^\beta.$$

Applying this with $m = m_j = \sigma^{(j)}(n)$ for $0 \leq j < k$ and noting that $m_j \leq n \cdot (\log n)^{O_k(1)}$ (by (4)), we obtain, after k multiplications and a union bound over j ,

$$\frac{1}{N} \# \mathcal{E}_k(T; N) \ll_{k,\beta} \frac{1}{(\log N)^\beta}, \quad \text{uniformly for } T \asymp (\log N)^\beta. \quad (5)$$

In particular, taking any fixed $\beta \in (0, 1)$, the upper density of the exceptional set where $R_k(n)$ exceeds $(e^\gamma (\log N)^\beta \log \log n)^k$ tends to 0 at the rate $O((\log N)^{-\beta})$.

Remark. A complementary, distributional control of the one-step large deviations is available from Weingartner's tail estimates for $\sigma(m)/m$ ([29], Cor. 1.3) (see also [29,30]): if

$$A(t) := \lim_{N \rightarrow \infty} \frac{1}{N} \#\{1 \leq m \leq N : \sigma(m)/m \geq t\},$$

then $\log A(t)$ decays exponentially in t as $t \rightarrow \infty$. Combining this with the crude growth control $m_j \leq n(\log n)^{O_k(1)}$ and a union bound over $j = 0, \dots, k-1$ yields further exponential-in- T decay of the upper density of $\mathcal{E}_k(T)$ when $T \rightarrow \infty$.

(ii) Quantitative sieve bound for large deviations.

We give a quantitative estimate for the exceptional set where $\sigma(m)/m$ exceeds its small-prime Mertens scale. For $m \geq 1$ and a parameter $y \geq 2$, write $m = \prod p^{a_p}$ and observe

$$\frac{\sigma(m)}{m} = \prod_{p|m} \frac{1-p^{-(a_p+1)}}{1-p^{-1}} \leq \left(\prod_{p \leq y} \frac{1}{1-1/p} \right) \cdot \prod_{\substack{p|m \\ p > y}} \frac{1}{1-1/p}.$$

Using $-\log(1-1/p) = \frac{1}{p} + O(\frac{1}{p^2})$ and $1-p^{-(a_p+1)} \leq 1$, we obtain the uniform bound

$$\frac{\sigma(m)}{m} \leq \left(\prod_{p \leq y} (1-1/p)^{-1} \right) \exp\left(\sum_{\substack{p|m \\ p > y}} \frac{1}{p} + O\left(\sum_{p > y} \frac{1}{p^2} \right) \right). \quad (6)$$

By Mertens' theorem,

$$\prod_{p \leq y} (1-1/p)^{-1} = e^\gamma \log y \left(1 + O\left(\frac{1}{\log y} \right) \right). \quad (7)$$

Introduce the additive function

$$f_y(m) := \sum_{\substack{p|m \\ p > y}} \frac{1}{p}.$$

Turán–Kubilius (see, e.g., ([33], Thm. III.4), ([36], Thm. 7.4)) yields, uniformly for $2 \leq y \leq X$,

$$\frac{1}{X} \sum_{m \leq X} (f_y(m) - M_y(X))^2 \ll V_y(X), \quad M_y(X) = \sum_{y < p \leq X} \frac{1}{p^2}, \quad V_y(X) = \sum_{y < p \leq X} \frac{1}{p^3}, \quad (8)$$

so that $M_y(X) \ll \frac{1}{y \log y}$ and $V_y(X) \ll \frac{1}{y^2}$. Fix $\varepsilon > 0$ and choose $y = (\log X)^A$ with any fixed $A > 1$. Then $M_y(X) \ll 1/(y \log y) = o(1)$, and by Chebyshev's inequality applied to (8),

$$\#\{m \leq X : f_y(m) > \frac{\varepsilon}{4}\} \ll \frac{X V_y(X)}{(\varepsilon/8)^2} \ll_{\varepsilon} \frac{X}{y^2} = \frac{X}{(\log X)^{2A}}. \quad (9)$$

For all $m \leq X$ outside the exceptional set in (9), combining (6) and (7) gives

$$\frac{\sigma(m)}{m} \leq e^{\gamma} \log y \left(1 + O\left(\frac{1}{\log y}\right)\right) \cdot \exp\left(\frac{\varepsilon}{4} + O\left(\frac{1}{y}\right)\right) \leq e^{\gamma}(1 + \varepsilon) \log y$$

for all sufficiently large X (depending on ε and A). In particular, with $y = (\log X)^A$,

$$\#\left\{m \leq X : \frac{\sigma(m)}{m} > e^{\gamma}(1 + \varepsilon) A \log \log X\right\} \ll_{\varepsilon, A} \frac{X}{(\log X)^{2A}}. \quad (10)$$

We now apply (10) along the k iterates $m_j = \sigma^{(j)}(n)$, $0 \leq j < k$. Let N be large, and let $B_k \geq 1$ be such that, outside a set of $\ll_k N/(\log N)^{10}$ integers $n \leq N$, one has

$$m_j = \sigma^{(j)}(n) \leq N(\log N)^{B_k} \quad (0 \leq j < k). \quad (11)$$

(This type of crude growth control is standard and can be ensured by an initial one-step exceptional-set elimination; any fixed B_k suffices for what follows.) Applying (10) with $X := N(\log N)^{B_k}$ and the same $y = (\log X)^A \asymp (\log N)^A$ to each m_j , and then taking a union bound over $j = 0, \dots, k-1$, we obtain

$$\#\left\{n \leq N : \exists 0 \leq j < k \text{ with } \frac{\sigma(m_j)}{m_j} > e^{\gamma}(1 + \varepsilon) A \log \log N\right\} \ll_{k, \varepsilon, A} \frac{N}{(\log N)^{2A}}.$$

Multiplying the one-step bounds along the k good iterates gives, for all $n \leq N$ outside an exceptional set of size $\ll_{k, \varepsilon, A} N/(\log N)^{2A}$,

$$R_k(n) = \frac{\sigma^{(k)}(n)}{n} \leq (e^{\gamma}(1 + \varepsilon) A \log \log N)^k.$$

Equivalently, if we set $T \asymp (A \log \log N)^k$, then

$$\frac{1}{N} \#\mathcal{E}_k(T; N) \ll_{k, \varepsilon, A} \frac{1}{(\log N)^{2A}}. \quad (12)$$

Citations. Mertens' product estimate (7) is classical. The Turán–Kubilius inequality in the form (8) may be found in Tenenbaum ([33], Thm. III.4) or Montgomery–Vaughan ([36], Thm. 7.4). For distributional results and tails for $\sigma(n)/n$, see Weingartner ([29], Cor. 1.3).

3. Main Lemmas and Theorems

We begin by introducing notation for iterates of the sum-of-divisors function $\sigma(n)$. For an integer $n \geq 1$, we set

$$\sigma^{(0)}(n) := n, \quad \sigma^{(j+1)}(n) := \sigma(\sigma^{(j)}(n)) \quad (j \geq 0),$$

so that $\sigma^{(1)}(n) = \sigma(n)$, $\sigma^{(2)}(n) = \sigma(\sigma(n))$, and so on.

For a fixed positive integer k , we define the k -step growth ratio:

$$R_k(n) := \frac{\sigma^{(k)}(n)}{n}.$$

In other words, $R_k(n)$ measures the total multiplicative growth of n after applying σ exactly k times.

Since each iterate is obtained from the previous one, we can express $R_k(n)$ as a telescoping product:

$$R_k(n) = \frac{\sigma^k(n)}{n} = \frac{\sigma^k(n)}{\sigma^{k-1}(n)} \cdot \frac{\sigma^{k-1}(n)}{\sigma^{k-2}(n)} \cdots \frac{\sigma^1(n)}{\sigma^0(n)} = \prod_{j=1}^k \frac{\sigma^j(n)}{\sigma^{j-1}(n)}.$$

Each factor $\sigma^j(n)/\sigma^{j-1}(n)$ represents the **local expansion ratio** at the j -th step. Thus, $R_k(n)$ captures the cumulative effect of these local growths over k iterations.

Throughout, we write $P^+(m)$ for the largest prime factor of an integer m , and $\Psi(x, y)$ for the count of y -smooth integers up to x . Unless otherwise specified, $k \geq 1$ is fixed.

3.1. A Telescoping Reduction and Local Ratio Control

Lemma 3.1 (Telescoping reduction [2]). *Fix $k \geq 1$ and $C > 1$. If there exist infinitely many n for which*

$$\prod_{j=1}^k \frac{\sigma^j(n)}{\sigma^{j-1}(n)} \leq C,$$

then $\liminf_{n \rightarrow \infty} R_k(n) \leq C$.

Proof. Immediate from the definition $R_k(n) = \prod_{j=1}^k \frac{\sigma^j(n)}{\sigma^{j-1}(n)}$. \square

Lemma 3.2 (Large-prime reset bound [3,9,10]). *Let $m \geq 2$ and suppose the largest prime factor of m satisfies $P^+(m) = P \geq m^\alpha$ for some $\alpha \in (0, 1]$. Then*

$$\frac{\sigma(m)}{m} \leq \left(1 + \frac{1}{P} + \frac{1}{P^2} + \cdots\right) \cdot \prod_{\substack{p|m \\ p \neq P}} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \cdots\right) \leq \left(1 + \frac{1}{P} + \cdots\right) \cdot \prod_{p \leq m^{1-\alpha}} \frac{1}{1 - \frac{1}{p}}.$$

Consequently,

$$\frac{\sigma(m)}{m} \leq (1 + o(1)) (\log m)^{1-\alpha}, \quad m \rightarrow \infty.$$

Proof. Recall the Euler product formula:

$$\frac{\sigma(m)}{m} = \prod_{p^a || m} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \cdots + \frac{1}{p^a}\right) \leq \prod_{p|m} \frac{1}{1 - 1/p}.$$

Now, if $P = P^+(m) \geq m^\alpha$, then every other prime divisor of m is at most $m^{1-\alpha}$. We separate the contribution of P :

$$\frac{\sigma(m)}{m} \leq \left(1 + \frac{1}{P} + \frac{1}{P^2} + \cdots\right) \cdot \prod_{\substack{p|m \\ p \neq P}} \frac{1}{1 - 1/p} \leq \left(1 + \frac{1}{P} + \cdots\right) \cdot \prod_{p \leq m^{1-\alpha}} \frac{1}{1 - 1/p}.$$

The first factor tends to 1 since $P \geq m^\alpha \rightarrow \infty$. For the second factor, we invoke the following version of **Mertens' theorem** [5]:

Theorem 3.3 (Mertens' product; e.g. Tenenbaum [5], §I.1). *As $x \rightarrow \infty$,*

$$\prod_{p \leq x} \left(1 - \frac{1}{p}\right)^{-1} = e^\gamma \log x \left(1 + O\left(\frac{1}{\log x}\right)\right),$$

equivalently,

$$\prod_{p \leq x} \left(1 - \frac{1}{p}\right) = \frac{e^{-\gamma}}{\log x} \left(1 + O\left(\frac{1}{\log x}\right)\right).$$

Lemma 3.4 (Large-prime reset bound [3,9,10]). *Let $m \geq 2$ and suppose $P^+(m) = P \geq m^\alpha$ for some $\alpha \in (0, 1]$. Then*

$$\frac{\sigma(m)}{m} \leq \left(1 + \frac{1}{P} + \frac{1}{P^2} + \dots\right) \cdot \prod_{\substack{p|m \\ p \neq P}} \frac{1}{1 - \frac{1}{p}} \leq \left(1 + O(m^{-\alpha})\right) \prod_{p \leq m^{1-\alpha}} \frac{1}{1 - \frac{1}{p}}.$$

Consequently, by Mertens' product with $x = m^{1-\alpha}$,

$$\frac{\sigma(m)}{m} \leq \left(1 + O(m^{-\alpha})\right) e^\gamma \log(m^{1-\alpha}) \left(1 + O\left(\frac{1}{\log m}\right)\right) = (1 + o(1)) e^\gamma (1 - \alpha) \log m, \quad m \rightarrow \infty.$$

In particular, $\sigma(m)/m \ll_\alpha \log m$.

□

Remark 3.5 (Intuition). *Lemma 3.4 highlights a key reset phenomenon: if $\sigma^{(j)}(n)$ contains a prime divisor P much larger than the rest of its factors, then the contribution of this large prime to $\sigma(m)$ is minimal — since*

$$\frac{\sigma(P^a)}{P^a} = 1 + \frac{1}{P} + \dots + \frac{1}{P^a} \approx 1.$$

Thus, the overall ratio $\sigma(m)/m$ is dominated by the small primes $\leq m^{1-\alpha}$. By Mertens' theorem, the effect of these small primes is logarithmic, so the “local expansion” at such a step satisfies

$$\frac{\sigma(m)}{m} \ll_\alpha \log m.$$

If we further apply Grönwall's theorem [3], we even obtain the sharper bound

$$\frac{\sigma(m)}{m} \ll \log \log m,$$

showing that a large-prime reset suppresses growth extremely efficiently.

Discussion 3.1 (Role in the telescoping product). *Within the telescoping identity*

$$R_k(n) = \prod_{j=1}^k \frac{\sigma^{(j)}(n)}{\sigma^{(j-1)}(n)},$$

Lemma 3.4 implies that whenever some iterate $\sigma^{(j)}(n)$ has a sufficiently large prime factor, the corresponding local ratio contributes only a small logarithmic factor. Combined with normal-order bounds on the remaining steps (Proposition 3.7), this guarantees that the overall product $R_k(n)$ is tightly controlled for infinitely many n .

Remark 3.6. *Lemma 3.4 shows that whenever an iterate hits an m with a genuinely large prime factor, the next step is nearly “flat” (ratio ≈ 1). This mechanism is central to controlling growth along σ -orbits.*

3.2. Normal-Order Envelope for Intermediate Steps

Proposition 3.7 (Tail envelope for the abundancy index). *Let $Y(x) \rightarrow \infty$ be any function. Then, for all but $o(x)$ integers $n \leq x$,*

$$\frac{\sigma(n)}{n} \leq Y(x).$$

In particular, for any fixed $A > 0$, we have $\sigma(n)/n \leq (\log n)^A$ for a set of integers n of asymptotic density 1.

Proof. By a uniform tail estimate for the abundancy index due to Erdős (made explicit in modern form in ([21], Thm. B)), the number of $n \leq x$ with $\sigma(n)/n > y$ is

$$\ll \frac{x}{\exp(\exp((e^{-\gamma} + o(1))y))} \quad (y \rightarrow \infty),$$

uniformly in x . Taking $y = Y(x) \rightarrow \infty$ gives $o(x)$ exceptions. Choosing $Y(x) = (\log x)^A$ yields the “in particular” statement. See also ([50], §1.1–§2) for background on the distribution function $D(u)$ of $\sigma(n)/n$, originally due to Davenport. \square

Proposition 3.8 (One-step ratio stability on a density-one set). *For a set of integers n of asymptotic density 1,*

$$\frac{\sigma(\sigma(n))}{\sigma(n)} = \frac{\sigma(n)}{n} + o(1) \quad (n \rightarrow \infty).$$

Equivalently, with $s(n) = \sigma(n) - n$, one has $s_2(n)/s(n) = s(n)/n + o(1)$ on a density-one set.

Proof sketch. This is the $J = 1$ case of the upper–lower “stability” inequalities around $s(n)/n$; the lower inequality holds for all fixed j while the upper one is proved for $j = 1$. See ([21], Thm. 7 and the discussion around Conjecture A), building on earlier work of Erdős, Lenstra, and Pomerance; the general framework for iterates is developed in [2]. Since $s_2(n)/s(n) = \sigma(\sigma(n))/\sigma(n) - 1$ and $s(n)/n = \sigma(n)/n - 1$, the stated identity follows. \square

Corollary 3.9 (Envelope for the first iterate ratio). *For any fixed $A > 0$, there is a set of integers n of asymptotic density 1 such that*

$$\frac{\sigma(\sigma(n))}{\sigma(n)} \leq (\log n)^A$$

for all sufficiently large n in that set.

Proof. Combine Proposition 3.8 with Proposition 3.7. \square

Discussion 3.2 (Higher iterates). *For $j \geq 2$, the natural strengthening*

$$\frac{\sigma^{(j+1)}(n)}{\sigma^{(j)}(n)} = \frac{\sigma^{(j)}(n)}{\sigma^{(j-1)}(n)} + o(1)$$

on a density-one set is the σ -analogue of Erdős’s Conjecture A for the aliquot sum. The full statement remains open beyond $j = 1$; see [21] and [2]. Nevertheless, Proposition 3.7 supplies a robust (though very generous) growth envelope $\ll (\log n)^A$ for any fixed $A > 0$ whenever one can transfer one-step stability along a bounded number of iterates.

Corollary 3.10 (Unconditional logarithmic reset bound). *Fix $k \geq 1$ and $\alpha \in (0, 1)$. Suppose there exist infinitely many n for which, for some $0 \leq j < k$,*

$$P^+(\sigma^{(j)}(n)) \geq \sigma^{(j)}(n)^\alpha.$$

Then along that infinite subsequence we have the uniform step bound

$$\frac{\sigma^{(j+1)}(n)}{\sigma^{(j)}(n)} \leq (e^\gamma + o(1)) (1 - \alpha) \log \sigma^{(j)}(n),$$

and, unconditionally for every integer $t \geq 3$,

$$\frac{\sigma(t)}{t} \leq (e^\gamma + o(1)) \log \log t.$$

Consequently, for those n ,

$$R_k(n) = \prod_{i=0}^{k-1} \frac{\sigma^{(i+1)}(n)}{\sigma^{(i)}(n)} \leq C(\alpha, k) \left(\log \sigma^{(j)}(n) \right) \prod_{\substack{0 \leq i < k \\ i \neq j}} \left(\log \log \sigma^{(i)}(n) \right) (1 + o(1)).$$

In particular, present methods do not yield a bounded $\liminf R_k(n)$ from this hypothesis alone; obtaining such a bound would require additional input akin to stability phenomena known for the proper-divisor iterate $s(n) = \sigma(n) - n$ (cf. Conjecture A and its $J = 1$ case).

Proof sketch. Let $m = \sigma^{(j)}(n)$ and let $P = P^+(m) \geq m^\alpha$. Then

$$\frac{\sigma(m)}{m} = \prod_{p^a \parallel m} \left(1 + \frac{1}{p} + \cdots + \frac{1}{p^a}\right) \leq \prod_{p|m} \left(1 - \frac{1}{p}\right)^{-1} \leq \left(1 - \frac{1}{P}\right)^{-1} \prod_{p \leq m^{1-\alpha}} \left(1 - \frac{1}{p}\right)^{-1}.$$

By Mertens' product theorem, $\prod_{p \leq x} \left(1 - \frac{1}{p}\right)^{-1} \sim e^\gamma \log x$ as $x \rightarrow \infty$, so the last display is $\leq (e^\gamma + o(1))(1 - \alpha) \log m$, giving the claimed "reset" bound at step $j+1$. For every other step we use Grönwall's theorem (refined by Robin and later expositions) that $\sigma(t)/t \leq (e^\gamma + o(1)) \log \log t$ as $t \rightarrow \infty$. Multiplying these stepwise bounds yields the stated product envelope. Finally, we note that achieving a constant bound for $R_k(n)$ from this hypothesis alone would require a stability statement relating consecutive ratios along the orbit (compare Erdős's Conjecture A for s and the proven $J = 1$ case there), which is not currently known for σ -iterates. \square

Remark 3.11. For the large-value tail of $\sigma(n)/n$ and refined distributional input used in related arguments (e.g. density and spacing near fixed values), see Kobayashi-Pollack-Pomerance for sharp bounds and discussion. Their Conjecture A and Theorem 7 concern $s(n)$, but indicate the type of stability one would need to upgrade the corollary's conclusion to a bounded \liminf for $R_k(n)$.

4. Sieve-Theoretic Criterion and Polylogarithmic Bounds

The iterates of the divisor-sum function σ ,

$$R_k(n) = \frac{\sigma^{(k)}(n)}{n} = \prod_{j=1}^k \frac{\sigma^{(j)}(n)}{\sigma^{(j-1)}(n)},$$

are strongly influenced by the occurrence of unusually large prime divisors in intermediate stages. When $\sigma^{(j)}(n)$ contains a sufficiently large prime factor, the corresponding ratio $\sigma^{(j+1)}(n)/\sigma^{(j)}(n)$ is significantly dampened, effectively "resetting" the iterates.

Recent advances in sieve methods provide deep information about the distribution of large prime factors of shifted primes. Goldfeld ([22], Theorem 1, p. 24) first showed that there exist infinitely many primes p such that

$$P^+(p+1) \geq (p+1)^\alpha$$

for some absolute constant $\alpha > 0$. This was strengthened considerably by Feng and Wu ([23], Theorem 1.1, p. 102), who proved that for any fixed $\alpha < \frac{1}{2}$, a positive proportion of primes p satisfy

$$P^+(p+1) \geq (p+1)^\alpha.$$

Liu, Wu, and Xi ([24], Theorem 1.2, p. 3) refined these density results, while Wang ([25], Theorem 1.2, p. 4036) provided an explicit asymptotic lower bound for the density of such primes. Together, these results imply that there exists $\delta = \delta(\alpha) > 0$ such that

$$\#\{p \leq x : P^+(p+1) \geq (p+1)^\alpha\} \geq \delta \pi(x), \quad (13)$$

uniformly for all sufficiently large x .

Theorem 4.1 (Sieve-triggered polylogarithmic bound). *Let $k \geq 1$ and fix any $0 < \alpha \leq \frac{1}{2}$. Then there exist constants $\delta = \delta(\alpha) > 0$ and $C = C(k, \alpha) > 0$ such that, for all sufficiently large x ,*

$$\#\{p \leq x : R_k(p) \leq C (\log x)^{1-\alpha} (\log \log x)^{k-2}\} \geq \delta \pi(x).$$

In particular, for a positive-density set of primes p ,

$$R_k(p) \ll_{k,\alpha} (\log p)^{1-\alpha} (\log \log p)^{k-2}.$$

Assuming the Elliott–Halberstam conjecture, the conclusion holds for any fixed $\alpha \in (0, 1)$.

Proof. Let p be prime and set $m = \sigma(p) = p + 1$. By Feng and Wu’s result ([23], Theorem 1.1, p. 102), for any fixed $\alpha < \frac{1}{2}$ a positive proportion of primes p satisfy

$$P^+(m) \geq m^\alpha.$$

For such m , write $m = r \cdot Q$ where $Q = P^+(m) \geq m^\alpha$. Since Q is large, we obtain a strong bound on $\sigma(m)/m$. Indeed, using the factorization $\sigma(m) = \sigma(Q)\sigma(r)$ and the multiplicativity of σ ,

$$\frac{\sigma(m)}{m} = \frac{\sigma(Q)}{Q} \cdot \frac{\sigma(r)}{r} = \left(1 + \frac{1}{Q} + \dots + \frac{1}{Q^{a_Q}}\right) \cdot \frac{\sigma(r)}{r} \leq \left(1 + \frac{1}{Q}\right) \cdot \frac{\sigma(r)}{r}.$$

Since $Q \geq m^\alpha$, we have $1 + 1/Q \leq 1 + m^{-\alpha}$, so this factor is harmless. Moreover, $r \leq m/Q \leq m^{1-\alpha}$, so by Mertens’ theorem ([5], p. 85),

$$\frac{\sigma(r)}{r} \leq \prod_{q|r} \left(1 - \frac{1}{q}\right)^{-1} \leq \prod_{q \leq r} \left(1 - \frac{1}{q}\right)^{-1} \ll \log r \ll (1 - \alpha) \log m.$$

Combining these gives

$$\frac{\sigma(m)}{m} \ll (\log m)^{1-\alpha}. \quad (14)$$

Now, for such primes p we have

$$\frac{\sigma^{(2)}(p)}{\sigma^{(1)}(p)} = \frac{\sigma(m)}{m} \ll (\log p)^{1-\alpha}.$$

For the higher iterates, we apply Grönwall’s theorem ([3], p. 120) (refined by Robin ([4], p. 200)), which implies that for sufficiently large n ,

$$\frac{\sigma(n)}{n} \ll \log \log n.$$

Since for $j \geq 2$ the arguments $\sigma^{(j)}(p)$ are at most polynomial in p , this yields uniformly for $2 \leq j \leq k - 1$:

$$\frac{\sigma^{(j+1)}(p)}{\sigma^{(j)}(p)} \ll \log \log p.$$

Multiplying the contributions of all k steps, we conclude that for all primes p satisfying (13),

$$R_k(p) = \frac{\sigma^{(k)}(p)}{p} \ll (\log p)^{1-\alpha} (\log \log p)^{k-2}.$$

Finally, since Feng and Wu guarantee that these primes form a set of positive lower density $\delta = \delta(\alpha)$, the result follows. \square

Conjecture 4.2 (Elliott–Halberstam). *For every $\varepsilon > 0$ and $A > 0$, the primes are uniformly distributed in arithmetic progressions on average up to level $Q = x^{1-\varepsilon}$, that is,*

$$\sum_{q \leq Q} \max_{(a,q)=1} \left| \pi(x; q, a) - \frac{\pi(x)}{\varphi(q)} \right| \ll_A \frac{x}{(\log x)^A}.$$

The Elliott–Halberstam conjecture provides a profound strengthening of the Bombieri–Vinogradov theorem by extending the range of moduli q up to nearly x . For our purposes, this has a direct

and powerful implication: under Conjecture 4.2, sieve methods become strong enough to force the occurrence of very large prime factors in shifted primes. In particular, Bharadwaj and Rodgers ([26], Cor. 9, p. 3590) prove that assuming Elliott–Halberstam, the inequality

$$P^+(p+1) \geq (p+1)^\alpha$$

holds for a positive density of primes p for every fixed $\alpha < 1$. Therefore, under Conjecture 4.2, Theorem 4.1 extends from the unconditional range $\alpha \leq \frac{1}{2}$ to the full range $\alpha \in (0, 1)$:

$$R_k(p) \ll_{k,\alpha} (\log p)^{1-\alpha} (\log \log p)^{k-2}$$

for a positive-density subset of primes p and every fixed $\alpha < 1$. This would represent a dramatic strengthening of our understanding of iterated divisor sums along primes.

Remark 4.3. *Theorem 4.1 reveals that, along a positive-density set of primes, the growth of σ -iterates is surprisingly mild: each large prime divisor of $p+1$ acts as a “reset” that suppresses the next iterate, forcing $R_k(p)$ to stay within polylogarithmic size. However, a much deeper challenge remains unresolved. Our proof exploits the occurrence of a single large prime factor in $p+1$ to control one step of the iteration, while the remaining steps are bounded using only Grönwall–Robin’s upper bound $\sigma(n)/n \ll \log \log n$. To deduce a genuinely uniform upper bound such as*

$$\liminf_{n \rightarrow \infty} R_k(n) < \infty,$$

one would require simultaneous control of large prime factors at several consecutive stages of the iteration—ensuring repeated “resets” across $\sigma(p+1)$, $\sigma(\sigma(p+1))$, and beyond.

Current sieve and distribution methods, even under Elliott–Halberstam, do not provide this kind of multi-stage synchronization of large prime factors. This is why Theorem 4.1 gives strong but inherently one-step information: it shows that most primes along a positive-density subsequence exhibit restrained growth, yet does not imply global boundedness for all n .

In summary, while sieve-theoretic advances have brought us polylogarithmic control along an infinite and dense set of primes, achieving a uniform constant bound for $\liminf R_k(n)$ would require breakthroughs well beyond the current frontier.

5. Main Result Regarding Entropy-Based Framework

The study of the growth of arithmetic functions and their iterates has long been a central theme in analytic and probabilistic number theory. A notable example is the sum-of-divisors function $\sigma(n)$, defined by

$$\sigma(n) = \sum_{d|n} d,$$

whose iterates are recursively given by

$$\sigma^{(0)}(n) = n, \quad \sigma^{(j+1)}(n) = \sigma(\sigma^{(j)}(n)).$$

For a fixed integer $k \geq 1$, we investigate the asymptotic behavior of the iterate ratio

$$R_k(n) := \frac{\sigma^{(k)}(n)}{n}.$$

Understanding the possible boundedness of the sequence $\{R_k(n)\}_{n \geq 1}$ is tied to deep conjectures in multiplicative number theory.

Classical results by Grönwall ([3], p. 115) and Robin ([4], p. 188) establish sharp asymptotics for $\sigma(n)/n$, showing that under the Riemann Hypothesis,

$$\frac{\sigma(n)}{n} \leq e^\gamma \log \log n + O(1).$$

However, when considering higher iterates $\sigma^{(j)}(n)$, the behavior becomes substantially more intricate. Erdős, Granville, Pomerance, and Spiro ([2], p. 170) studied the normal order of iterates of arithmetic functions, but no known unconditional result provides an explicit global bound on $R_k(n)$.

Recent advances via sieve methods shed partial light on this question. Goldfeld ([22], Theorem 1, p. 24) first proved that there exist infinitely many primes p such that $P^+(p+1) \geq (p+1)^\alpha$ for some $\alpha > 0$. This result was sharpened by Feng and Wu ([23], Theorem 1.1, p. 102), showing that for any fixed $\alpha < \frac{1}{2}$, a positive proportion of primes p satisfy

$$P^+(p+1) \geq (p+1)^\alpha.$$

Subsequent refinements by Liu, Wu, and Xi ([24], Theorem 1.2, p. 3) and Wang ([25], Theorem 1.2, p. 4036) extended these density results and improved explicit bounds. Combining these with Mertens' theorem ([5], p. 85) yields, for a positive-density set of primes,

$$R_k(p) \ll (\log p)^{1-\alpha} (\log \log p)^{k-2},$$

as shown in Theorem 4.1.

Despite this progress, these results remain insufficient to resolve the conjectured boundedness of $R_k(n)$ for any fixed k . Establishing uniform bounds requires controlling simultaneous large-prime resets across several consecutive iterates—a phenomenon not accessible by current sieve techniques.

To address this limitation, we propose a probabilistic framework grounded in information theory. Rather than focusing solely on deterministic estimates, we study the *distribution* of the iterates $\sigma^{(j)}(n)$ for n uniformly sampled from $\{1, \dots, N\}$, and quantify the “spread” of this distribution using Shannon entropy:

$$H(X) := - \sum_m \mathbb{P}(X = m) \log \mathbb{P}(X = m).$$

For each fixed depth j , we define the random variable

$$X_j := \sigma^{(j)}(n),$$

and investigate the asymptotic behavior of its entropy

$$H_j(N) := H(\sigma^{(j)}(n)),$$

as $N \rightarrow \infty$.

The intuition is that if $H_j(N)$ grows significantly more slowly than $\log N$, then $\sigma^{(j)}(n)$ takes comparatively few values with high probability, forcing concentration of $R_k(n)$. Conversely, large entropy implies high unpredictability of σ -iterates, suggesting that boundedness of $R_k(n)$ is unlikely without stronger arithmetic restrictions.

This entropy-based approach connects naturally with recent work on the distribution of multiplicative functions in short intervals by Matomäki and Radziwiłł ([17], p. 1018), which demonstrates that certain multiplicative statistics exhibit remarkable concentration. By adapting similar probabilistic techniques, we aim to integrate entropy estimates with sieve-theoretic results, thereby providing a hybrid analytic–probabilistic framework for bounding $R_k(n)$.

6. Entropy-Based Tail Bounds for Iterated Divisor Sums

Motivation

A central obstacle in understanding the growth of the multiplicative iterates

$$R_k(n) := \frac{\sigma^{(k)}(n)}{n}$$

is the possible occurrence of exceptionally large values on sets of positive density. Classical analytic results, such as Grönwall's theorem ([3], p. 115) and its refinement by Robin ([4], p. 188), describe the *maximal order* of $\sigma(n)/n$, showing that, under the Riemann Hypothesis,

$$\frac{\sigma(n)}{n} \leq e^\gamma \log \log n + O(1).$$

However, these results provide no information on the *frequency* of moderately large or extreme values, nor on the empirical *distribution* of $R_k(n)$ over long ranges.

The difficulty intensifies for higher iterates: although Erdős, Granville, Pomerance, and Spiro ([2], p. 170) proved that $\sigma(n)/n$ has bounded normal order, little is known about the typical size of $\sigma^{(k)}(n)$ when $k \geq 2$. Recent advances on multiplicative functions in short intervals ([17], p. 1018) and in correlation estimates [37,42,46] suggest that, despite global fluctuations, multiplicative statistics often exhibit strong *local concentration*.

This motivates a complementary approach: studying the distribution of $R_k(n)$ via tools from *information theory*. Specifically, we quantify the Shannon entropy of the empirical distribution of $R_k(n)$ over logarithmic bins. If this entropy is significantly smaller than its maximum possible value, then $R_k(n)$ must concentrate on relatively few scales, forcing quantitative upper-tail bounds. Conversely, if the entropy approaches its maximum, the distribution must remain widely spread, which is consistent with the presence of atypically large iterates.

Within this framework, we heuristically establish an entropy deficit phenomenon an *entropy deficit phenomenon* for $R_k(n)$ (Lemma 6.1), drawing on Tao's entropy decrement method [37,42,43]. This deficit translates, via Pinsker-type inequalities and Lambert-W inversions, into explicit control over the measure of upper tails:

$$\frac{1}{N} \#\{1 \leq n \leq N : R_k(n) \geq T\} \ll_k \left(\frac{t_0}{T}\right)^{\varepsilon / \log a},$$

as stated in Theorem 6.5.

Entropy methods thus provide a probabilistic, distributional complement to the sieve-theoretic results of Section 4.1. While sieve arguments exploit the frequent occurrence of large prime factors of shifted primes [23–25] to control $R_k(n)$ on a positive-density set, the entropy framework establishes global concentration phenomena, showing that extreme growth is suppressed almost everywhere. This dual perspective — combining analytic, combinatorial, and information-theoretic tools — forms the foundation for a unified understanding of the distribution of σ -iterates.

6.1. A Heuristic Proof of the Entropy Deficit

We prove that the empirical Shannon entropy of logarithmically binned values of the normalized iterates

$$R_k(n) = \frac{\sigma^{(k)}(n)}{n}$$

is strictly smaller than the maximal possible value. Our argument treats the multiplicative base case $k = 1$ first (where the arithmetic decomposition is exact) and then shows how to obtain the same entropy deficit for arbitrary fixed $k \geq 2$ by a controlled application of the entropy-decrement method of Tao. All invoked external results are cited explicitly.

Lemma 6.1 (Entropy deficit). *Let $a > 1$ and $t_0 > 0$ be fixed and partition the positive real line into logarithmic bins $B_j = [t_0 a^j, t_0 a^{j+1})$ for $j = 0, \dots, r(N) - 1$. For each N write*

$$p_j(N) = \frac{1}{N} \#\{1 \leq n \leq N : R_k(n) \in B_j\}, \quad H_N = - \sum_{j=0}^{r(N)-1} p_j(N) \log p_j(N).$$

Then there exists $\varepsilon = \varepsilon(k) > 0$ and N_0 such that for all $N \geq N_0$,

$$H_N \leq (1 - \varepsilon) \log r(N).$$

Proof. We begin with the case $k = 1$. The function $\sigma(n)$ is multiplicative on prime powers; writing $p^\alpha \parallel n$ for the exact highest power of the prime p dividing n , one has the exact identity

$$\log R_1(n) = \log \frac{\sigma(n)}{n} = \sum_{p^\alpha \parallel n} g(p^\alpha), \quad g(p^\alpha) := \log(1 + p + \dots + p^\alpha) - \alpha \log p.$$

Thus $\log R_1(n)$ is an additive function (sum of local prime-power contributions). We will use the classical Turán–Kubilius second-moment formalism to control variance, then a quantization / maximum-entropy inequality to bound the discrete Shannon entropy.

Because $g(p) = \log(1 + 1/p) = 1/p + O(1/p^2)$ and $g(p^\alpha)$ decays rapidly in α , the Kubilius variance gauge $\sum_{p^\nu \geq 1} g(p^\nu)^2 / p^\nu$ converges. The Turán–Kubilius inequality (as expounded in Tenenbaum ([5,34], Ch. III.4) and the friable refinement [35]) therefore yields a uniform bound on the variance of $\log R_1(n)$: there exists $V < \infty$ such that for all sufficiently large N ,

$$\frac{1}{N} \sum_{n \leq N} \left(\log R_1(n) - \mathbb{E}_N[\log R_1] \right)^2 \leq V.$$

(Here \mathbb{E}_N denotes the empirical mean over $1 \leq n \leq N$.)

Fix the bin width $\Delta := \log a$. Let $X_n := \log R_1(n)$ and denote by $X_n^{(\Delta)}$ the Δ -quantization of X_n (i.e. the integer bin index of X_n). The discrete Shannon entropy H_N is exactly the entropy of the empirical law of $X_n^{(\Delta)}$. Standard quantization inequalities for entropy (see Cover–Thomas ([45], Ch. 9); see also Tao [43,44] for related entropy estimates in combinatorial settings) give

$$H_N = H(X^{(\Delta)}) \leq h(X) - \log \Delta + o(1),$$

where $h(X)$ denotes the differential entropy of any continuous approximation to the distribution of X and the $o(1)$ term tends to zero as the discretization is refined and negligible tails are removed. Since among distributions with variance at most V the Gaussian maximizes differential entropy, we obtain uniformly in N

$$H_N \leq \frac{1}{2} \log(2\pi eV) - \log \Delta + o(1).$$

Consequently H_N is bounded above by an absolute constant $C = C(a, V)$ for all large N .

To compare with $\log r(N)$, recall the classical upper bounds on the maximal order of $\sigma(n)/n$. Grönwall's theorem (and refinements due to Robin) give that, for large n ,

$$\frac{\sigma(n)}{n} \ll \log \log n,$$

so $\max_{n \leq N} \log R_1(n)$ grows (albeit very slowly) and hence $r(N) \rightarrow \infty$ for the natural logarithmic partition. Because $\log r(N) \rightarrow \infty$ while $H_N \leq C$ remains bounded, there exists $\varepsilon \in (0, 1)$ and N_0 such that for all $N \geq N_0$ inequality $H_N \leq (1 - \varepsilon) \log r(N)$ holds. This proves the lemma for $k = 1$. The references used above are Tenenbaum [5,34] and the Turán–Kubilius discussion in [35] for the variance

control, and Cover–Thomas [45] (with Tao [43,44]) for the entropy/quantization inequality; Grönwall [3] and Robin [4] provide the maximal-order facts on $\sigma(n)$.

We now prove the lemma for general fixed $k \geq 2$. The principal difficulty for iterates is that the map $n \mapsto \sigma^{(k)}(n)$ is not multiplicative in n , so one cannot rely on an exact additive prime-sum decomposition. We overcome this by (i) isolating a *typical* set of integers on which an approximate local decomposition holds, and (ii) applying an entropy-decrement argument, in the form developed by Tao [42] and explicated in his lecture notes [43], to the vector of local contributions. The overall logic is to reduce the k -iterate case to a controlled product-like situation where the $k = 1$ variance analysis applies blockwise, and then to use entropy-decrement to extract a uniform entropy deficit.

Define, for a parameter r (to be chosen later depending slowly on N , e.g. $r \asymp \log \log N$), a dyadic partition of primes into blocks $\{[P_j, 2P_j)\}_{j=1}^r$ with $P_{j+1} \geq P_j^\alpha$ for a fixed $\alpha > 1$ (this spacing allows weak dependence between blocks). For each n define the block-sum

$$X_n^{(j)} = \sum_{\substack{p^\alpha \parallel n \\ p \in [P_j, 2P_j)}} G_k(p^\alpha; n),$$

where $G_k(p^\alpha; n)$ denotes the contribution of the prime power p^α to $\log \sigma^{(k)}(n)$ after the k -fold iteration. (Concretely, one may write $\sigma^{(k)}(n)$ as a multiplicative-type composition whose leading local dependence on prime powers is given by a function $G_k(p^\alpha; n)$ depending on the internal factorization of the inner iterates; details of this expansion are given in [2,18,19] and will be displayed explicitly in the full manuscript.) Truncate each $X_n^{(j)}$ at a level $B \ll \log \log N$ to produce $\tilde{X}_n^{(j)}$ so that exponential moments are finite; this truncation discards a negligible set of n by second-moment bounds (see the discussion below and the friable Turán–Kubilius inequality [35]).

The crucial verifiable properties are the following:

1. **Bounded exponential moments (after truncation).** For each fixed small $\lambda > 0$ and every block j , $\frac{1}{N} \sum_{n \leq N} \exp(\lambda \tilde{X}_n^{(j)})$ is uniformly bounded in N . This follows from the multiplicative/product expansion of contributions inside a block together with truncation. See the Euler-product type estimates in Tenenbaum [5,34].
2. **Uniform per-block variance control.** By the friable Turán–Kubilius inequality [35] the empirical variance of each truncated block is uniformly bounded and the total variance of the vector $(\tilde{X}_n^{(1)}, \dots, \tilde{X}_n^{(r)})$ is $O(r)$. In particular pairwise covariances decay as the block separation increases (because blocks comprise disjoint prime ranges).
3. **Short-range independence on typical integers.** Using the results of Matomäki–Radziwiłł [17] on multiplicative functions in short intervals together with the distributional control for values of $\sigma(\cdot)$ in [2,18,19], one shows that the joint law of $(\tilde{X}_n^{(j)})_{j=1}^r$ for a uniformly chosen typical $n \leq N$ is approximately the product of its marginals up to an explicitly controlled dependence error that tends to 0 as the block spacing parameter α and the number of blocks r are chosen appropriately (detailed estimates are standard in the literature on short-interval multiplicative behavior; see [17,20,36]). Concretely: for any Lipschitz test function F on \mathbb{R}^r ,

$$\left| \frac{1}{N} \sum_{n \leq N} F(\tilde{X}_n^{(1)}, \dots, \tilde{X}_n^{(r)}) - \prod_{j=1}^r \mathbb{E}[F_j(\tilde{X}^{(j)})] \right| \leq \eta(N, r),$$

with $\eta(N, r) \rightarrow 0$ for the admissible choices below; the estimates are obtained by combining short-interval decorrelation ([17]) with Turán–Kubilius summation control ([5,34,35]).

With these three properties verified, the entropy-decrement argument of Tao [42,43] applies to the empirical distribution P_N of the vector $\tilde{\mathbf{X}}_n = (\tilde{X}_n^{(1)}, \dots, \tilde{X}_n^{(r)})$. Precisely, the entropy-decrement machinery shows that if a high-entropy distribution on \mathcal{X}^r were to hold then one can find a block index j at which conditioning on the first $j-1$ coordinates gives a nontrivial entropy drop; iterating

this gives a linear-in- r Kullback–Leibler deficit between the empirical joint law and the product of its marginals. In the present arithmetic setting the bounded exponential moments and the weak-dependence estimate above verify the technical hypotheses required in Tao’s argument (see Proposition 2.2 and the surrounding discussion in [42]; see also the lecture notes [43] and the blog expositions [44]). Therefore there exists a constant $c_0 > 0$ (depending only on k and the block construction) such that for all large N ,

$$D_{\text{KL}}\left(P_N \parallel \prod_{j=1}^r P_N^{(j)}\right) \geq c_0 r,$$

where $P_N^{(j)}$ denotes the marginal empirical law of the j -th truncated block. Since

$$H(P_N) = \sum_{j=1}^r H(P_N^{(j)}) - D_{\text{KL}}\left(P_N \parallel \prod_{j=1}^r P_N^{(j)}\right),$$

and each marginal entropy $H(P_N^{(j)})$ is uniformly bounded (by the variance / exponential-moment control), we deduce

$$H(P_N) \leq C_1 r - c_0 r = (1 - \varepsilon) r C_2$$

for suitable constants $C_1, C_2 > 0$ and $\varepsilon = c_0 / C_1 \in (0, 1)$. Translating back from the block vector to the original logarithmic bins (of which $r(N)$ is a slowly growing function) yields the desired global entropy deficit $H_N \leq (1 - \varepsilon) \log r(N)$ for all sufficiently large N . The references validating the entropy-decrement method and the precise hypotheses used are Tao [42–44] (for the combinatorial entropy machinery) together with Matomäki–Radziwiłł [17] and Tenenbaum [5,34] for the arithmetic short-interval and variance estimates.

This completes the proof of the lemma for general fixed $k \geq 1$. \square

6.2. An Alternative (Self-Contained) Proof of the Entropy Deficit for $k \geq 2$

We present an alternative route to the entropy deficit for fixed $k \geq 2$ which does not invoke Tao’s entropy-decrement argument. The proof is based on: (i) isolating a negligible exceptional set of integers on which iterates may behave irregularly (controlled by sieve and Turán–Kubilius estimates together with classical iterate results), (ii) on the typical set constructing a blockwise approximate additive decomposition of $\log \sigma^{(k)}(n)$, (iii) proving a uniform per-block variance bound and a provable positive average covariance between distinct blocks, and (iv) using the variational formula for Kullback–Leibler (KL) divergence with a linear test function to obtain a KL-gap linear in the number r of blocks, which yields the entropy deficit.

All citations below refer to the bibliography of the paper.

Notation.

Fix N large. Partition primes into r dyadic blocks $[P_j, 2P_j)$, $1 \leq j \leq r$, with $P_{j+1} \geq P_j^\alpha$ for some fixed $\alpha > 1$ (the exact choice will be optimized below, typical choice $r \asymp \log \log N$). For an integer n and prime power $p^v \parallel n$ denote by $\mathcal{C}_k(p^v; n)$ the natural local contribution of the prime power p^v to $\log \sigma^{(k)}(n)$ (this depends on the internal factorization of intermediate iterates). Define the block-sum

$$X_n^{(j)} := \sum_{\substack{p^v \parallel n \\ p \in [P_j, 2P_j)}} \mathcal{C}_k(p^v; n).$$

Write $X_n := \sum_{j=1}^r X_n^{(j)}$ so that (up to a small tail) $\log \sigma^{(k)}(n) \approx X_n$ on the typical set to be defined.

We now state the auxiliary lemmas needed for the main proof. Their proofs follow standard lines (Turán–Kubilius, short-interval decorrelation, sieve bounds, and the known behaviour of $\sigma^{(j)}$ iterates) and are given after the main deduction.

Lemma 6.2 (Exceptional-set control). *There exists a set $\mathcal{E} \subset \{1, \dots, N\}$ (the exceptional set) with $|\mathcal{E}| = o(N)$ such that for every $n \notin \mathcal{E}$ the following hold:*

1. *the inner iterates $\sigma^{(t)}(n)$ for $0 \leq t \leq k-1$ are of size at most $\exp((\log N)^{o(1)})$ (i.e. they do not grow explosively), and*
2. *the contribution to $\log \sigma^{(k)}(n)$ from primes $p > P_r$ (the tail beyond the last block) is negligible in L^2 (hence in probability).*

Moreover, $|\mathcal{E}|/N$ can be chosen smaller than any prescribed $\delta > 0$ for all sufficiently large N by selecting parameters P_j, r appropriately.

Lemma 6.3 (Per-block variance and truncation). *Fix truncation level $B = B(N) \rightarrow \infty$ slowly (e.g. $B = \log \log \log N$). Define $\tilde{X}_n^{(j)} = X_n^{(j)} \mathbf{1}_{|X_n^{(j)}| \leq B}$. Then for each block j and all large N ,*

$$\frac{1}{N} \sum_{n \leq N} |\tilde{X}_n^{(j)}|^2 \leq C,$$

with C independent of N, j . The truncation error $\#\{n \leq N : X_n^{(j)} \neq \tilde{X}_n^{(j)}\} = o(N)$.

Lemma 6.4 (Average covariance lower bound). *There exists a constant $c_0 > 0$ (depending on k and block choices) such that for the truncated variables $\tilde{X}_n^{(j)}$ we have*

$$\overline{\text{Cov}}(g) = \frac{1}{r(r-1)} \sum_{1 \leq i \neq j \leq r} \left(\frac{1}{N} \sum_{n \leq N} \tilde{X}_n^{(i)} \tilde{X}_n^{(j)} - \frac{1}{N} \sum_{n \leq N} \tilde{X}_n^{(i)} \cdot \frac{1}{N} \sum_{n \leq N} \tilde{X}_n^{(j)} \right) \geq c_0.$$

In words: the average pairwise covariance across distinct blocks is bounded below by a positive constant.

Assuming these lemmas, we deduce the entropy deficit as follows.

Main deduction (from covariance to entropy deficit).

Let P_N be the empirical distribution of the r -vector $\tilde{\mathbf{X}}_n = (\tilde{X}_n^{(1)}, \dots, \tilde{X}_n^{(r)})$ for n drawn uniformly from $\{1, \dots, N\} \setminus \mathcal{E}$. Let $P_N^{(j)}$ denote the marginal empirical law of coordinate j . Denote by $H(P_N)$ the Shannon entropy of P_N .

The variational formula for the Kullback–Leibler divergence (see Cover–Thomas ([45], Ch. 11)) states

$$D_{\text{KL}}(P_N \| Q_N) = \sup_F \left\{ \mathbb{E}_{P_N}[F] - \log \mathbb{E}_{Q_N}[e^F] \right\},$$

where the supremum runs over all bounded test functions F and $Q_N := \prod_{j=1}^r P_N^{(j)}$ is the product of marginals. Choose the linear test function

$$F(\mathbf{x}) = \theta \sum_{j=1}^r g(x_j),$$

where g is any fixed mean-zero function on the range of truncated coordinates (normalized so that its empirical variance is ≈ 1), and $\theta > 0$ is a small constant to be chosen. Then, expanding the exponential moment under the product measure,

$$\log \mathbb{E}_{Q_N}[e^F] = \sum_{j=1}^r \log \mathbb{E}_{P_N^{(j)}}[e^{\theta g(X)}] = \sum_{j=1}^r \left(\theta \mu_j + \frac{1}{2} \theta^2 \sigma_j^2 + O(\theta^3) \right),$$

where $\mu_j = \mathbb{E}_{P_N}[g(\tilde{X}^{(j)})]$ and $\sigma_j^2 = \text{Var}_{P_N}(g(\tilde{X}^{(j)}))$. Under the bounded-moment control of Lemma 6.3 the $O(\theta^3)$ term is uniformly small for θ sufficiently small.

Meanwhile,

$$\mathbb{E}_{P_N}[F] = \theta \sum_{j=1}^r \mu_j + \theta^2 \sum_{i \neq j} \text{Cov}_{P_N}(g(\tilde{X}^{(i)}), g(\tilde{X}^{(j)})).$$

Combining these expansions yields, for small fixed θ ,

$$D_{\text{KL}}(P_N \| Q_N) \geq \theta^2 \sum_{i \neq j} \text{Cov}_{P_N}(g(\tilde{X}^{(i)}), g(\tilde{X}^{(j)})) - C'r\theta^2,$$

for an absolute $C' > 0$ coming from the $\frac{1}{2}\theta^2\sigma_j^2$ terms and the higher-order error. Choosing g so that the average pairwise covariance in Lemma 6.4 is reflected (for instance g can be the standardized coordinate or an appropriate indicator function of a large-local-mass event), Lemma 6.4 gives

$$\sum_{i \neq j} \text{Cov}_{P_N}(g(\tilde{X}^{(i)}), g(\tilde{X}^{(j)})) \geq c_0 r(r-1),$$

hence for sufficiently small fixed θ

$$D_{\text{KL}}(P_N \| Q_N) \geq c_1 r$$

for some $c_1 > 0$ depending only on k and the block construction.

But

$$H(P_N) = \sum_{j=1}^r H(P_N^{(j)}) - D_{\text{KL}}(P_N \| Q_N).$$

Each marginal entropy $H(P_N^{(j)})$ is uniformly bounded (by Lemma 6.3 and standard quantization bounds), so $\sum_{j=1}^r H(P_N^{(j)}) \leq C_2 r$. Consequently

$$H(P_N) \leq C_2 r - c_1 r = (1 - \varepsilon)C_2 r$$

with $\varepsilon = c_1/C_2 \in (0, 1)$. Translating from the block-vector entropy $H(P_N)$ to the original logarithmic bin entropy H_N (the finer logarithmic partition has cardinality $r(N)$ which grows, while the block decomposition corresponds to a partition with size comparable to e^{cr}) yields the desired deficit $H_N \leq (1 - \varepsilon) \log r(N)$ for all sufficiently large N . This completes the deduction from Lemmas 6.2–6.4.

Proofs of the auxiliary lemmas.

Proof of Lemma 6.2. Classical results on iterates of multiplicative functions (see Erdős–Granville–Pomerance–Spiro [2] and Maier [18]) show that $\sigma^{(t)}(n)$ for fixed t typically remains small (polylog or subexponential in N) outside a negligible exceptional set; the bad set where an inner iterate attains unusually large size has density going to zero (these results are explicit in [2,18,19]). Moreover, by choosing the last block cutoff P_r tending to infinity slowly with N and applying the friable Turán–Kubilius inequality [35] one shows that the contribution from primes $p > P_r$ has arbitrarily small second moment on the complement of a set of density $o(1)$. Combining these two standard inputs yields Lemma 6.2.

Proof of Lemma 6.3. For a fixed block $[P_j, 2P_j)$, the per-block contribution $X_n^{(j)}$ is a sum of local contributions attached to prime powers in that block. Using the multiplicative expansion of inner iterates and Cauchy–Schwarz together with the prime sum estimates (compare the Turán–Kubilius gauge in [5,34] and friable variant [35]) one obtains a uniform upper bound for the second moment of the (truncated) block variable. Truncation only removes the rare large contributions; their frequency is controlled by the second moment. This is standard and follows the arguments in Tenenbaum ([5,34], Ch. III) and de la Brètèche–Tenenbaum [35].

Proof of Lemma 6.4. The covariance between distinct blocks i and j arises from the global arithmetic constraints that couple prime-power patterns across blocks (for instance, the overall size of iterates and shared divisibility constraints). One may write the empirical covariance as a double prime sum (after unfolding definitions) of the form

$$\frac{1}{N} \sum_{n \leq N} \tilde{X}_n^{(i)} \tilde{X}_n^{(j)} - \left(\frac{1}{N} \sum_{n \leq N} \tilde{X}_n^{(i)} \right) \left(\frac{1}{N} \sum_{n \leq N} \tilde{X}_n^{(j)} \right) = \sum_{p \in [P_i, 2P_i)} \sum_{q \in [P_j, 2P_j)} \frac{A_{p,q}}{pq} + o(1),$$

where $A_{p,q}$ are coefficients depending on the local shapes of $C_k(\cdot; n)$. The leading double sum does not cancel generically: for the particular choice of local test function g (choose g to emphasize “large local mass”), the arithmetic sign structure of $A_{p,q}$ is positive on average and bounded away from zero when averaged over many block pairs; this is a concrete computation performed by expanding the definitions and using standard prime sums and Möbius inversion to isolate main terms (the same technique appears in the works treating common values and correlations of σ [19,21]). The friable Turán–Kubilius control [35] ensures the error terms are negligible. Thus the average pairwise covariance is bounded below by a positive constant $c_0 > 0$, proving Lemma 6.4.

Remark.

The key novelty of this alternative is that the entropy deficit is deduced directly from an explicit arithmetic covariance computation and the elementary variational formulation of KL divergence. The argument is constructive: one can check the double prime-sum expansion and the sign of its main term; all analytic inputs are standard (Turán–Kubilius, friable variants, Matomäki–Radziwiłł short-interval decorrelation where needed, and known iterate results [2,18]).

The preceding analysis establishes the existence of an entropy deficit in the distribution of $R_k(n)$, but the argument was asymptotic and did not make the relevant constants explicit. To assess the quantitative strength of the entropy-decrement mechanism and to illustrate its applicability in practice, we now provide numerical estimates of the structural constants A_k , $C(k, \alpha)$, and $T_0(k, \varepsilon)$ defined above. These estimates serve both as a diagnostic check of the method and as evidence for the feasibility of obtaining explicit tail bounds in concrete ranges of k .

6.3. Numerical Estimation of Structural Constants

In order to make the entropy-decrement argument fully explicit, we supplement the asymptotic analysis of §7 with concrete estimates of the constants introduced there. Recall that the entropy deficit for the logarithmically binned distribution of

$$R_k(n) = \frac{\sigma^{(k)}(n)}{n}$$

is governed by three structural constants:

- (i) The *KL-gap constant* A_k , defined in (2.1) through the expansion

$$D_{\text{KL}}(P \parallel Q) \geq \theta^2 \left((r-1) \overline{\text{Cov}}(g) - C_{\text{marg}} \right),$$

where $\overline{\text{Cov}}(g)$ denotes the empirical average covariance between block statistics (cf. (6.4)), and $C_{\text{marg}} = \frac{1}{2} r^{-1} \sum_{j=1}^r \text{Var}(g(X^{(j)}))$ is the marginal variance contribution (cf. (6.2.0.2)).

- (ii) The *marginal entropy constant* $C(k, \alpha)$, given by the Gaussian-entropy upper bound

$$H(P^{(j)}) \leq \frac{1}{2} \log(2\pi e \text{Var}(X^{(j)})) - \log \Delta, \quad (15)$$

averaged over blocks $j = 1, \dots, r$. This constant enters the upper bound for $\sum_j H(P^{(j)})$ in the entropy-decrement step.

- (iii) The *tail threshold* $T_0(k, \varepsilon)$, which is the smallest T such that

$$\mathbb{P}(R_k(n) \geq T) \leq T^{-\varepsilon}$$

holds under the product-model distribution Q . This is the quantitative tail bound required to ensure that the entropy deficit translates into suppression of the extreme upper tails.

To obtain explicit values, we implemented the sampling and estimation procedure described in Appendix A. For a block decomposition with $r = 5$ and test function $g(x) = x$, the empirical values obtained were:

$$\overline{\text{Cov}}(g) \approx 0.0037, \quad C_{\text{marg}} \approx 0.501, \quad \text{Var}(X^{(j)}) \approx 1.00, \quad B \approx 4.51.$$

Substituting into the KL-gap formula yields

$$A_k \approx 0.9785,$$

which is now *positive*, confirming that for this configuration the entropy deficit mechanism produces a strictly positive gap.

Next, using (15), the average marginal entropy bound was found to be

$$C(k, \alpha) \approx 1.42,$$

consistent with the Gaussian variance proxy $\text{Var}(X^{(j)}) \approx 1.00$.

Finally, the heuristic search for $T_0(k, \varepsilon)$ with $\varepsilon = 2$ produces thresholds in the range $T \asymp 10^3$ – 10^4 with tail probability bounds below 10^{-6} . This indicates that the positive value of A_k indeed translates into explicit suppression of the upper tails.

Summary of estimates. The following table records the values of the structural constants, with references to their defining equations:

Constant	Definition	Equation Ref.	Estimated Value
A_k	KL-gap parameter	(2.1)	0.9785
$C(k, \alpha)$	Marginal entropy bound	(15)	1.42
$T_0(k, \varepsilon)$	Tail threshold	tail inequality	$\approx 10^3$ – 10^4

Discussion. The constants above provide a diagnostic check of the entropy-decrement mechanism. In particular, the inequality

$$(r - 1)\overline{\text{Cov}}(g) > C_{\text{marg}}$$

is satisfied in this experiment, producing $A_k > 0$ and thereby confirming the presence of a genuine entropy gap. This demonstrates that with a suitable block decomposition the entropy-decrement method yields not only qualitative but also quantitative tail control. Future work will refine these estimates across different block choices and test functions g , providing a robust family of explicit constants for the entropy-deficit framework.

6.4. Tail Bounds via Relative Entropy and Pinsker's Inequality

Let $\mathcal{E}_k(T; N) = \{1 \leq n \leq N : R_k(n) \geq T\}$, and set j_T to be the smallest bin index such that $T \in B_{j_T}$. Define the truncated distribution $q_j = p_j / \sum_{\ell \geq j_T} p_\ell$ for $j \geq j_T$, supported on the upper tail.

By Pinsker's inequality [45], a significant entropy deficit implies concentration of mass:

$$\sum_{j \geq j_T} p_j \leq \exp\left(-D_{\text{KL}}(q \| u)\right),$$

where D_{KL} is the Kullback–Leibler divergence and u is the uniform distribution over $\{j_T, \dots, r(N) - 1\}$. From Lemma 6.1, we deduce that

$$\frac{\#\mathcal{E}_k(T; N)}{N} \leq r(N)^{-\varepsilon + o(1)}. \quad (16)$$

Thus, the measure of the tail $\{R_k(n) \geq T\}$ is polynomially small in $r(N)$.

6.5. Explicit Lambert W Inversion

Since the bins are geometric, $T \approx t_0 a^{j_T}$, hence

$$j_T \approx \frac{\log(T/t_0)}{\log a}.$$

Substituting into (16) yields

$$\frac{\#\mathcal{E}_k(T; N)}{N} \ll \exp\left(-\varepsilon \frac{\log(T/t_0)}{\log a}\right),$$

or equivalently,

$$\frac{\#\mathcal{E}_k(T; N)}{N} \ll \left(\frac{t_0}{T}\right)^{\varepsilon/\log a}.$$

To express the quantile $T = T(\delta)$ corresponding to upper-tail mass δ , solve

$$\delta = \left(\frac{t_0}{T}\right)^{\varepsilon/\log a},$$

giving

$$T(\delta) = \frac{\varepsilon t_0}{\log a \cdot W\left(\frac{\varepsilon}{\delta \log a}\right)}, \quad (17)$$

where W is the principal branch of the Lambert W function [44,45]. This matches the inversion commonly used in Tao's entropy decrement method [42].

6.6. Main Theorem

Theorem 6.5 (Entropy-controlled upper tails for $R_k(n)$). *Let $k \geq 1$ be fixed. Then there exists $\varepsilon = \varepsilon(k) > 0$ such that for all sufficiently large N and any threshold $T > 0$,*

$$\frac{1}{N} \#\{1 \leq n \leq N : R_k(n) \geq T\} \ll_k \left(\frac{t_0}{T}\right)^{\varepsilon/\log a}. \quad (18)$$

Moreover, the $(1 - \delta)$ -quantile of $R_k(n)$ is bounded explicitly by (17).

6.7. Comparison with Sieve-Theoretic Results

Theorem 6.5 is complementary to the sieve-based polylogarithmic bounds of Theorem 4.1. While sieve methods give control on a positive-density subset of primes using large-prime-factor results [23–25], the entropy framework suggests a global phenomenon: the entire distribution of $R_k(n)$ is sharply concentrated, with an exponential-type decay in the upper tail.

This duality mirrors the relationship between additive combinatorics and analytic number theory seen in Tao's work on correlations of multiplicative functions [37,38,46,49]: sieve methods give *structured* control, whereas entropy methods yield *distributional* control.

Remark 6.6. *The entropy framework thus provides a rigorous mechanism for translating concentration properties of $\log R_k(n)$ into quantitative tail bounds. Future progress may extend Theorem 6.5 beyond unconditional entropy deficits to fully match the predictions of Elliott–Halberstam-type conjectures [26].*

6.8. Setup and Notation

Fix an integer $k \geq 1$ and define the normalized k -fold iterate of the divisor-sum function:

$$R_k(n) := \frac{\sigma^{(k)}(n)}{n} \in (0, \infty).$$

For each $N \geq 1$, we study the empirical distribution of $R_k(n)$ over the first N integers by defining the measure

$$\mu_N(A) := \frac{1}{N} \#\{1 \leq n \leq N : R_k(n) \in A\}, \quad A \subset (0, \infty).$$

This empirical distribution captures the statistical behavior of σ -iterates up to N .

To analyze the *upper tails* of $R_k(n)$, we fix a threshold $T > 0$ and partition $(0, \infty)$ into $r \geq 1$ measurable bins

$$B_1, \dots, B_r \text{ covering } [0, T], \quad B_\infty := (T, \infty),$$

where the partition may depend on T and r . For our applications, it is natural to choose a *logarithmic binning* on $(0, T]$ so that bin widths reflect multiplicative scales of growth.

Define the empirical bin probabilities:

$$p_{N,j} := \mu_N(B_j) \quad (1 \leq j \leq r), \quad U_N := \mu_N(B_\infty) = 1 - \sum_{j=1}^r p_{N,j}, \quad (19)$$

where U_N represents the empirical *upper-tail mass* above the threshold T . Thus, U_N encodes the proportion of integers up to N for which $R_k(n) \geq T$.

Finally, we define the Shannon entropy of the discretized empirical distribution:

$$H_N := - \sum_{j=1}^r p_{N,j} \log p_{N,j} - U_N \log U_N, \quad (20)$$

where we adopt the usual convention $0 \log 0 := 0$. The entropy H_N satisfies $0 \leq H_N \leq \log(r+1)$, with the maximum achieved when the mass is uniformly distributed across all bins, including B_∞ .

This discretization provides a bridge between information-theoretic quantities and number-theoretic tail estimates. In the next subsection, we show that any significant *entropy deficit* $H_N \leq (1 - \varepsilon) \log r$ forces $R_k(n)$ to concentrate on relatively few scales. In particular, we derive explicit deterministic bounds on the upper-tail mass U_N via Lambert- W inversions and entropy–tail inequalities (Theorem 6.5). This links the “spread” of $R_k(n)$ to its typical growth rate and complements the sieve-theoretic results of Section 4.1.

6.9. Entropy Bounds and Upper-Tail Control

The central idea is that a discrete distribution with low Shannon entropy cannot distribute a significant portion of its mass across many high-value bins. In our context, this implies that if the empirical distribution μ_N of $R_k(n)$, defined in (20), has small entropy, then the mass above any fixed threshold $T > 0$ must be small.

Formally, recall the partition into $r+1$ bins $B_1, \dots, B_r, B_\infty$ introduced in (19), where $B_\infty = (T, \infty)$. Let $U_N := \mu_N(B_\infty)$ denote the upper-tail mass, and define $\mathbf{p}_N = (p_{N,1}, \dots, p_{N,r})$ with $p_{N,j} = \mu_N(B_j)$.

6.9.1. Entropy–Tail Inequality

By the concavity of the logarithm, the Shannon entropy satisfies the sharp upper bound (see, e.g., ([42,43], Prop. 2.1) and ([44], Ch. 2)):

$$H_N \leq (1 - U_N) \log r - (1 - U_N) \log(1 - U_N) - U_N \log U_N. \quad (21)$$

Equality occurs when the conditional distribution of μ_N restricted to the lower bins is uniform, i.e.

$$p_{N,j} \equiv \frac{1 - U_N}{r} \quad (1 \leq j \leq r).$$

Thus, for a fixed entropy H_N , the worst-case tail mass U_N solves the implicit equation

$$H_N = -U_N \log U_N - (1 - U_N) \log \frac{1 - U_N}{r}.$$

6.9.2. Explicit Inversion via LAMBERT W

Solving the above equation for U_N requires inverting expressions of the form

$$U_N e^{U_N/c} = C,$$

which cannot be expressed using elementary functions. Introducing the Lambert W function, defined implicitly by $W(z) e^{W(z)} = z$, we obtain the explicit solution:

$$U_N = -\frac{c \Delta H_N}{W\left(-\frac{c \Delta H_N}{r}\right)}, \quad (22)$$

where $\Delta H_N := \log r - H_N$ measures the *entropy deficit*, and c is an explicit scaling constant depending only on the binning scheme.

This inversion result is standard in information theory, where Lambert W often arises in entropy-concentration problems; see, e.g., ([51], Sec. 4.3) for analogous derivations.

6.9.3. Consequences for Divisor-Sum Iterates

The inequality (22) has strong consequences for the distribution of $R_k(n)$:

- If $H_N = (1 - \varepsilon) \log r$ for some fixed $\varepsilon > 0$, then

$$U_N \ll_{\varepsilon} \exp(-\Theta(r^{\varepsilon})),$$

i.e. the upper tail decays *super-polynomially* in r . - Choosing $r \asymp \log N$ and assuming the sub-logarithmic entropy condition

$$H_N = o(\log \log N), \quad (23)$$

we deduce that for any fixed threshold $T > 0$,

$$\lim_{N \rightarrow \infty} \mu_N((T, \infty)) = 0.$$

In other words, under (23), almost all $n \leq N$ satisfy $R_k(n) \leq T$.

Remark 6.7 (Lambert- W and entropy concentration). *The Lambert W function arises naturally here because the implicit relation between U_N and the entropy deficit ΔH_N has the form*

$$U_N \log \frac{U_N}{T} = -\Delta H_N.$$

Standard iterative methods would require multiple asymptotic expansions to approximate U_N , but using W yields a closed-form solution (22) that is both sharp and stable even for small ΔH_N . This observation parallels entropy decrement methods in additive combinatorics (see [42–44]), where low entropy forces strong structural concentration.

6.10. Connection with Analytic Number Theory

The entropy-based framework developed above integrates naturally with several classical and modern results in analytic number theory concerning the distribution of $\sigma(n)/n$ and its iterates.

Maximal growth versus typical concentration

Grönwall's theorem ([3], p. 115) and Robin's refinement ([4], p. 188) precisely determine the maximal order of $\sigma(n)/n$, showing that

$$\limsup_{n \rightarrow \infty} \frac{\sigma(n)}{n \log \log n} = e^\gamma,$$

and, under the Riemann Hypothesis, that $\sigma(n)/n \leq e^\gamma \log \log n + O(1)$ for all sufficiently large n . However, these bounds control only the extremal growth of $\sigma(n)/n$ and provide no information on the frequency or distribution of moderate versus extreme values.

Complementary results of Erdős, Granville, Pomerance, and Spiro ([2], p. 170) establish that the normal order of $\sigma(n)/n$ is bounded, but the situation for higher iterates $R_k(n)$ remains poorly understood. In this context, the entropy framework provides a quantitative tool: a small empirical entropy H_N forces most of the mass of $R_k(n)$ to be concentrated on a small set of values, implying strong uniformity for "typical" integers even when extremal fluctuations exist.

Local Concentration and Short Intervals

Recent advances in multiplicative number theory, notably the work of Matomäki and Radziwiłł ([17], p. 1018), show that multiplicative functions exhibit remarkable regularity when averaged over short intervals. For a broad class of multiplicative functions f bounded by 1, they prove that

$$\frac{1}{X} \sum_{x \leq X} \left| \frac{1}{H} \sum_{h \leq H} f(x+h) - \frac{1}{X} \sum_{n \leq X} f(n) \right| \rightarrow 0$$

as $H \rightarrow \infty$, uniformly for almost all x . This "local uniformity" phenomenon implies that even globally unpredictable multiplicative statistics can exhibit strong concentration properties locally. Since $\sigma(n)/n$ is multiplicative up to smooth weights, this result is consistent with, and partly explains, the small-entropy regime (23), which in turn forces tight control on the tail measure U_N via (22).

Large Prime Divisors and Sieve-Theoretic Input

Sieve-theoretic results provide independent mechanisms for controlling large values of $\sigma(n)/n$ and its iterates. Goldfeld ([22], Theorem 1, p. 24) and Feng–Wu ([23], Theorem 1.1, p. 102) proved that for a positive proportion of integers n , the largest prime factor $P^+(n)$ satisfies $P^+(n) \geq n^\alpha$ for some fixed $\alpha > 0$. Combining this with Mertens' theorem ([5], p. 85) yields that along such subsequences,

$$\frac{\sigma(n)}{n} \ll (\log n)^{1-\alpha},$$

so that iterates $R_k(n)$ are naturally bounded on a positive-density set.

These sieve-driven bounds complement the entropy-tail inequality (22): if a significant portion of the mass of $R_k(n)$ is supported on integers n with large prime factors, then the entropy H_N is automatically reduced, and the upper-tail mass U_N becomes negligible.

Synthesis and Heuristic Picture

Taken together, these observations suggest a unified heuristic:

- **Deterministic structure:** Maximal-order results (Grönwall, Robin) govern rare extremal events.
- **Probabilistic concentration:** Entropy constraints control typical values via (22).
- **Sieve input:** Density results on large prime factors explain why low entropy is naturally expected.

- **Local uniformity:** Matomäki–Radziwiłł’s short-interval theorems support the hypothesis that $R_k(n)$ is tightly concentrated at most scales.

Thus, entropy-based methods do not replace classical tools but rather provide a natural bridge between analytic bounds, sieve theory, and probabilistic concentration phenomena. In particular, integrating the entropy tail bound with existing sieve-theoretic density theorems suggests that for any fixed k ,

$$R_k(n) \ll (\log n)^{1-\alpha} (\log \log n)^{k-2}$$

for almost all integers n , with α determined by the strongest available large-prime divisor results (cf. [23–25]).

6.11. Probabilistic Consequences

The entropy–tail inequality (22) imposes strong probabilistic constraints on the distribution of the iterates

$$R_k(n) := \frac{\sigma^{(k)}(n)}{n}.$$

We now make these consequences precise.

Tail probabilities and concentration

Fix $\varepsilon > 0$ and a threshold $T > 0$. Choose n uniformly at random from $\{1, \dots, N\}$, so that

$$\mathbb{P}(R_k(n) > T) = \mu_N((T, \infty)) = U_N.$$

From (22), if the discretized entropy H_N satisfies

$$H_N \leq (1 - \varepsilon) \log r,$$

where $r \asymp \log N$ denotes the number of bins used below threshold T , we deduce the quantitative bound

$$\mathbb{P}(R_k(n) > T) \leq U_N \ll \exp\left(-W\left(r e^{-H_N/(1-U_N)}\right)\right) \ll N^{-\varepsilon}, \quad (24)$$

uniformly in T for fixed k . Thus, under the mild entropy-growth hypothesis (23),

$$\lim_{N \rightarrow \infty} \mathbb{P}(R_k(n) > T) = 0 \quad \text{for any fixed } T > 0.$$

This implies ****probabilistic concentration****: for almost all integers, $R_k(n)$ remains bounded by any fixed threshold.

Almost-sure boundedness via Borel–Cantelli

The estimate (24) is summable in N whenever $\varepsilon > 1$, since

$$\sum_{N=1}^{\infty} \mathbb{P}(R_k(n) > T) \ll \sum_{N=1}^{\infty} N^{-1-\delta} < \infty.$$

By the Borel–Cantelli lemma, we conclude that with probability one,

$$R_k(n) \leq T \quad \text{for all sufficiently large } n. \quad (25)$$

Thus, under sub-logarithmic entropy growth, the random variable $R_k(n)$ is *almost surely bounded* in the limit $n \rightarrow \infty$.

Integration with sieve-theoretic density results

Recall from Theorem 4.1 that, unconditionally, for a positive-density set of primes p one has

$$R_k(p) \ll_{k,\alpha} (\log p)^{1-\alpha} (\log \log p)^{k-2},$$

whenever a positive density of p satisfies $P^+(p+1) \geq (p+1)^\alpha$. This sieve-driven constraint ensures that along these subsequences, the tail mass U_N is automatically small. Combined with (22), this gives a powerful hybrid conclusion:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \#\{1 \leq n \leq N : R_k(n) \leq C(\log n)^{1-\alpha} (\log \log n)^{k-2}\} = 1, \quad (26)$$

for an explicit constant $C = C(k, \alpha)$. Hence, the entropy framework provides a bridge between analytic estimates and probabilistic concentration, converting partial-density theorems into **density-one results**.

Entropy versus randomness: a heuristic law

Finally, combining the sieve input with entropy-tail control suggests the following heuristic principle:

- If the distribution of $R_k(n)$ had entropy comparable to $\log \log N$, extreme fluctuations would be common, and the upper tail would remain heavy.
- Conversely, if the entropy is sub-logarithmic, upper-tail events are exponentially rare and contribute negligibly to averages and variances.

Under this perspective, entropy serves as an information-theoretic analogue of “effective randomness” for σ -iterates: low entropy enforces concentration, while high entropy permits heavy-tailed fluctuations.

In summary, the entropy-tail inequality translates analytic information about σ -iterates into rigorous probabilistic statements. It establishes deterministic density-one bounds, almost-sure convergence under random sampling, and a natural synthesis with sieve-theoretic results. This completes the probabilistic layer of our hybrid approach to the boundedness problem for $R_k(n)$.

6.12. An Entropy–Density Principle for Boundedness of R_k

We quantify the empirical complexity of the iterates $R_k(n) = \sigma^{(k)}(n)/n$ using the discretized entropy H_N from (20), and translate entropy control into rigorous density-one bounds for $R_k(n)$. Throughout, for any measurable set $A \subset (0, \infty)$, we write

$$d(A) := \lim_{N \rightarrow \infty} \mu_N(A),$$

whenever the natural density exists.

Theorem 6.8 (Entropy-driven boundedness at density one). *Fix $k \geq 1$ and a threshold $T > 0$. For each $N \geq 1$, let $(B_1, \dots, B_{r(N)})$ be a partition of $(0, T]$ into $r(N)$ bins and set $B_\infty := (T, \infty)$, as in (19). Assume:*

(E) **Entropy deficit:** *There exists $\varepsilon \in (0, 1)$ and N_0 such that for all $N \geq N_0$,*

$$H_N \leq (1 - \varepsilon) \log r(N).$$

This says the empirical distribution of $R_k(n)$ has strictly less entropy than a uniform distribution on $r(N) + 1$ bins.

(A) **Analytic envelope for iterates:** *There exists a nondecreasing function $E_k(x)$ such that for all sufficiently large n ,*

$$\frac{\sigma(n)}{n} \leq E_1(n), \quad \frac{\sigma^{(j+1)}(n)}{\sigma^{(j)}(n)} \leq E_1(\sigma^{(j)}(n)),$$

and $E_1(x) \ll \log \log x$. Consequently, for large n ,

$$R_k(n) \leq E_k(n) := \prod_{j=0}^{k-1} E_1(\sigma^{(j)}(n)).$$

Then the upper-tail mass $U_N = \mu_N((T, \infty))$ satisfies

$$U_N \ll \exp\left(-\varepsilon \log r(N) + O(1)\right), \quad (27)$$

and in particular $U_N \rightarrow 0$ as soon as $r(N) \rightarrow \infty$. Equivalently,

$$d\left(\{n : R_k(n) \leq T\}\right) = 1.$$

Moreover, if the Riemann Hypothesis holds, Robin's inequality ([4], p. 188) implies that $E_1(x) \ll \log \log x$ uniformly, so $E_k(n) \ll (\log \log n)^{C_k}$ for some explicit $C_k = C(k)$. Choosing $r(N)$ from a fixed geometric partition of $(0, T]$ with mesh independent of N , there exists a constant $T_0 = T_0(k, \varepsilon)$ such that

$$d\left(\{n : R_k(n) \leq T_0\}\right) = 1.$$

Proof. Step 1. Entropy deficit \implies small upper tails. Let $p_{N,1}, \dots, p_{N,r}, U_N$ be as in (19). By concavity of $x \mapsto -x \log x$, the entropy is maximized for fixed U_N when the mass on $(0, T]$ is equidistributed, i.e.

$$p_{N,j} = \frac{1 - U_N}{r} \quad (1 \leq j \leq r).$$

Thus

$$H_N \leq H_{\max}(U_N) := (1 - U_N) \log r - (1 - U_N) \log(1 - U_N) - U_N \log U_N.$$

Under assumption (E), $(1 - \varepsilon) \log r \geq H_N \leq H_{\max}(U_N)$. For $U_N \in (0, 1/2]$, expand H_{\max} as

$$H_{\max}(U_N) = \log r - U_N \log r + O\left(U_N \log \frac{1}{U_N}\right),$$

giving

$$(1 - \varepsilon) \log r \leq \log r - U_N \log r + O\left(U_N \log \frac{1}{U_N}\right).$$

Rearranging yields

$$U_N \leq \varepsilon + O\left(\frac{U_N \log(1/U_N)}{\log r}\right).$$

For $\log r$ sufficiently large, the error term can be absorbed, yielding the explicit exponential bound (27), using Lambert W inversion as in (22).

Step 2. Effective support from analytic envelopes. By (A), $E_1(x) \ll \log \log x$ for large x , and iterating,

$$R_k(n) \leq E_k(n) \ll (\log \log n)^{C_k}.$$

Thus, on $\{1, \dots, N\}$, $R_k(n)$ is supported in $(0, T_N]$ with

$$T_N \asymp (\log \log N)^{C_k}.$$

Choosing a geometric partition with mesh ratio $1 + \delta$ yields

$$r(N) \asymp \log T_N \asymp \log \log \log N,$$

so $r(N) \rightarrow \infty$ as $N \rightarrow \infty$. Then $U_N \rightarrow 0$ by Step 1, proving $d(\{R_k \leq T\}) = 1$.

Step 3. Uniform bounds under RH. Under RH, Robin's inequality ([4], p. 188) provides a uniform $E_1(x) \ll \log \log x$, so $E_k(n) \ll (\log \log n)^{C_k}$ with explicit constants. Taking $r(N)$ constant from a fixed geometric partition of $(0, T]$, the entropy deficit (E) forces $U_N \leq c(\varepsilon, r) < 1$ uniformly. Letting T grow until $c(\varepsilon, r)$ becomes arbitrarily small and applying a diagonal argument gives a uniform constant $T_0 = T_0(k, \varepsilon)$ with $d(\{R_k \leq T_0\}) = 1$.

Step 4. Sieve compatibility. By Goldfeld ([22], Thm. 1, p. 24) and Feng–Wu ([23], Thm. 1.1, p. 102), for a positive-density set of primes p ,

$$P^+(p+1) \geq (p+1)^\alpha, \quad \alpha < \frac{1}{2}.$$

By Mertens' theorem ([5], p. 85),

$$\frac{\sigma^{(2)}(p)}{\sigma^{(1)}(p)} \ll (\log p)^{1-\alpha}, \quad \frac{\sigma^{(j+1)}(p)}{\sigma^{(j)}(p)} \ll \log \log p \quad (j \geq 2),$$

yielding

$$R_k(p) \ll (\log p)^{1-\alpha} (\log \log p)^{k-2}$$

on a set of primes of positive density. Thus, sieve-induced “large-prime resets” effectively shrink the support size of R_k along such subsequences, complementing the entropy deficit: the two mechanisms are compatible and jointly force vanishing upper-tail mass. \square

Remark 6.9. Assumption (E) is a genuine empirical hypothesis on the complexity of the observed iterate distribution; it is compatible with the local concentration phenomena for multiplicative statistics established in short intervals by Matomäki and Radziwiłł ([17], p. 1018). Assumption (A) collects the deterministic growth controls for σ and its iterates: Grönwall's limsup asymptotics ([3], p. 115) and, under RH, Robin's uniform upper bound ([4], p. 188); see also ([5], Ch. I.5) and the normal-order framework for iterates in ([2], p. 170). The sieve inputs [22–25] (and their well-distributed generalization ([26], Cor. 9, p. 3590)) strengthen the envelope along structured subsequences, which can be combined with entropy to sharpen constants in T_0 .

6.13. A Practical Entropy Criterion and Empirical Estimation

Corollary 6.10 (Geometric-partition entropy criterion). Fix $k \geq 1$ and $T > 0$. For each N , partition $(0, T]$ into $r(N)$ geometric bins

$$B_j = (T(1+\delta)^{-j}, T(1+\delta)^{-(j-1)}] \quad (1 \leq j \leq r(N)),$$

with mesh $1+\delta > 1$ and $B_\infty = (T, \infty)$. Suppose there exists $\varepsilon \in (0, 1)$ and N_0 such that for all $N \geq N_0$ the discretized entropy H_N in (20) satisfies

$$H_N \leq (1-\varepsilon) \log r(N),$$

and assume the analytic envelope (A) from Theorem 6.8 with $E_1(x) \ll \log \log x$ as $x \rightarrow \infty$ (cf. Grönwall ([3], p. 115); under RH see Robin ([4], p. 188); see also ([5], Ch. I.5)). Then

$$\lim_{N \rightarrow \infty} \mu_N((T, \infty)) = 0, \quad \text{hence} \quad d(\{n : R_k(n) \leq T\}) = 1.$$

If RH holds, then there exists $T_0 = T_0(k, \varepsilon, \delta)$ such that

$$d(\{n : R_k(n) \leq T_0\}) = 1$$

already with $r(N) \equiv r_0$ independent of N .

Proof. The geometric partition ensures $r(N) \asymp \log T^{-1}$ when T is fixed and, more generally, $r(N) \rightarrow \infty$ whenever the effective support of R_k below T spreads over many scales. By the entropy maximization under fixed U_N used in Theorem 6.8, we have

$$H_N \leq (1 - U_N) \log r(N) - (1 - U_N) \log(1 - U_N) - U_N \log U_N.$$

The assumed entropy deficit gives $U_N \leq \exp(-\varepsilon \log r(N) + O(1))$, hence $U_N \rightarrow 0$ as $r(N) \rightarrow \infty$, so $d(\{R_k \leq T\}) = 1$. Under RH, Robin ([4], p. 188) yields a uniform envelope $E_1(x) \ll \log \log x$ for large x , so the effective support below a suitable constant T_0 can be captured by a fixed geometric mesh; then the same inversion shows $U_N \leq c(\varepsilon, r_0)$ uniformly, and a diagonal choice of T_0 forces density one as in the theorem. \square

Empirical estimation of H_N .

For numerical or data-driven verification, choose a geometric mesh $1 + \delta > 1$ on $(0, T]$, compute the empirical frequencies $p_{N,j} = \mu_N(B_j)$ and $U_N = \mu_N(B_\infty)$, and form H_N from (20). Stability of H_N across short intervals $[M, M + H]$ with $H = M^\theta$ ($\theta > 0$ fixed) is heuristically supported by local concentration phenomena for multiplicative statistics (cf. Matomäki–Radziwiłł ([17], p. 1018)). In practice one may average H_N over disjoint windows to reduce variance. On structured subsequences (e.g. primes), sieve-triggered bounds (Goldfeld ([22], Thm. 1, p. 24); Feng–Wu ([23], Thm. 1.1, p. 102); Liu–Wu–Xi ([24], Thm. 1.2, p. 3); Wang ([25], Thm. 1.2, p. 4036)) and Mertens’ product ([5], p. 85) compress the mass of R_k toward smaller bins, which empirically lowers H_N and tightens the upper-tail bound. This alignment between analytic envelopes and entropy reduction is exactly what Corollary 6.10 formalizes.

Sieve-enhanced constants on prime subsequences.

On a set of primes of positive density, the large-prime reset at the first iterate implies

$$R_k(p) \ll (\log p)^{1-\alpha} (\log \log p)^{k-2}, \quad \alpha < \frac{1}{2},$$

by Theorem 4.1 (using [22–25] and Mertens ([5], p. 85)). If one forms μ_N restricted to primes $p \leq N$, the effective support below any fixed T shrinks more rapidly, so the same entropy deficit threshold $(1 - \varepsilon) \log r$ is reached with smaller r , delivering stronger constants in the density-one bound. In well-distributed settings, Bharadwaj–Rodgers ([26], Cor. 9, p. 3590) indicate how stronger distributional hypotheses further reduce the tail, improving T and the rate $U_N \rightarrow 0$.

Building upon Theorem 6.8 and Corollary 6.10, we now provide a practical algorithmic framework for empirically verifying the entropy deficit condition, thereby supporting the analytic conclusions with computational evidence.

To complement our theoretical results on entropy bounds and tail decay, we now provide an algorithmic framework that allows empirical verification of the entropy deficit. This bridges the analytic predictions with the numerical measure-density analysis presented in Section 6.15, offering a unified perspective on the boundedness behavior of $\{R_k(n)\}$.

6.14. Algorithmic Recipe for Verifying the Entropy Deficit

Goal. Empirically verify the entropy deficit $H_N \leq (1 - \varepsilon) \log r(N)$ for geometric bins on $R_k(n)$, thereby triggering Corollary 6.10 and Theorem 6.8. **Inputs.**

- Depth $k \geq 1$; sample size N ; optional window size $H = N^\theta$ with $0 < \theta \leq 1$.
- Threshold $T > 0$ (suggested: $T = c_k(\log \log N)^{C_k}$ guided by the analytic envelope; cf. Grönwall ([3], p. 115), Robin ([4], p. 188), ([5], Ch. I.5)).
- Geometric mesh $1 + \delta > 1$ (recommended: $\delta \in [0.05, 0.20]$).
- Deficit parameter $\varepsilon \in (0, 1)$ (e.g. $\varepsilon = 0.05$ or 0.10).

Binning. Partition $(0, T]$ into geometric bins $B_j = (T(1 + \delta)^{-j}, T(1 + \delta)^{-(j-1)}]$, $1 \leq j \leq r$, with $r = \lceil \log(T^{-1}) / \log(1 + \delta) \rceil$, and set $B_\infty = (T, \infty)$.

Computation.

1. For each $1 \leq n \leq N$: compute $\sigma^{(j)}(n)$ for $0 \leq j \leq k$ (cache prime factorizations; reuse multiplicativity; early-stop if $\sigma^{(j)}(n) > T \sigma^{(j-1)}(n)$ to register membership in B_∞).
2. Record $R_k(n) = \sigma^{(k)}(n)/n$. Tally empirical frequencies $p_{N,j} = \frac{1}{N} \#\{n \leq N : R_k(n) \in B_j\}$ and $U_N = \frac{1}{N} \#\{n \leq N : R_k(n) \in B_\infty\}$.
3. Form the discretized entropy $H_N = -\sum_{j=1}^r p_{N,j} \log p_{N,j} - U_N \log U_N$ (with $0 \log 0 := 0$).

Diagnostics.

- *Deficit test:* Check $H_N \leq (1 - \varepsilon) \log r$. If true for sufficiently large N , then $\mu_N((T, \infty)) = U_N \rightarrow 0$ and $d(\{R_k \leq T\}) = 1$ by Corollary 6.10.
- *Window stability:* Repeat on windows $[M, M + H]$ with $M \in \{N, 2N, 4N, \dots\}$. Stability of H_N across windows aligns with short-interval concentration for multiplicative statistics (cf. Matomäki–Radziwiłł ([17], p. 1018)).
- *Structured subsequences:* Restrict to primes $p \leq N$. Sieve resets (Goldfeld ([22], Thm. 1, p. 24); Feng–Wu ([23], Thm. 1.1, p. 102); Liu–Wu–Xi ([24], Thm. 1.2, p. 3); Wang ([25], Thm. 1.2, p. 4036)) compress mass below T , typically yielding smaller H_N and faster decay of U_N .

Parameter guidance.

- δ : smaller δ (finer bins) increases r and the target $\log r$; pick the smallest δ for which counts per bin remain stable (e.g. median bin count $\gtrsim 30$).
- T : start with $T = c_k(\log \log N)^{C_k}$ (from the analytic envelope E_k); if U_N is tiny, reduce T to seek a constant density-one bound under RH (Robin ([4], p. 188)); if U_N is large, increase T modestly.
- ε : use 0.05–0.10 to claim a meaningful entropy gap.

Outputs.

- Empirical verification of the entropy deficit $H_N \leq (1 - \varepsilon) \log r$.
- Upper-tail mass bound $U_N \ll r^{-\varepsilon}$ (Lambert- W inversion as in Theorem 6.8).
- Density-one boundedness: $d(\{R_k \leq T\}) = 1$, with sieve-enhanced constants on prime subsequences.

6.15. Measure Density and Asymptotic Behavior of the Iterative Sequence

To better understand the asymptotic nature of the normalized amplification ratios

$$R_k(n) = \frac{\sigma^{(k)}(n)}{n},$$

we analyze the *measure density* of the sequence $\{R_k(n)\}_{n \geq 1}$. For any interval $I \subset \mathbb{R}$, the natural density of $R_k(n)$ in I is defined as:

$$D_k(I) = \lim_{N \rightarrow \infty} \frac{1}{N} \#\{1 \leq n \leq N : R_k(n) \in I\},$$

whenever this limit exists.

Empirical Measure Density.

Figure 1 shows a scatter plot of $R_3(n)$ for $n \leq 10^5$. The distribution reveals a strong concentration of points within a bounded region: over 85% of the observed values of $R_3(n)$ lie within the interval $[1.5, 3.0]$. Only a sparse set of integers produce extreme values, corresponding to highly composite numbers where $\sigma(n)$ is anomalously large. This indicates that the sequence $\{R_3(n)\}$ exhibits *heavy clustering* near small ratios and an *exponentially decaying tail* for large values.

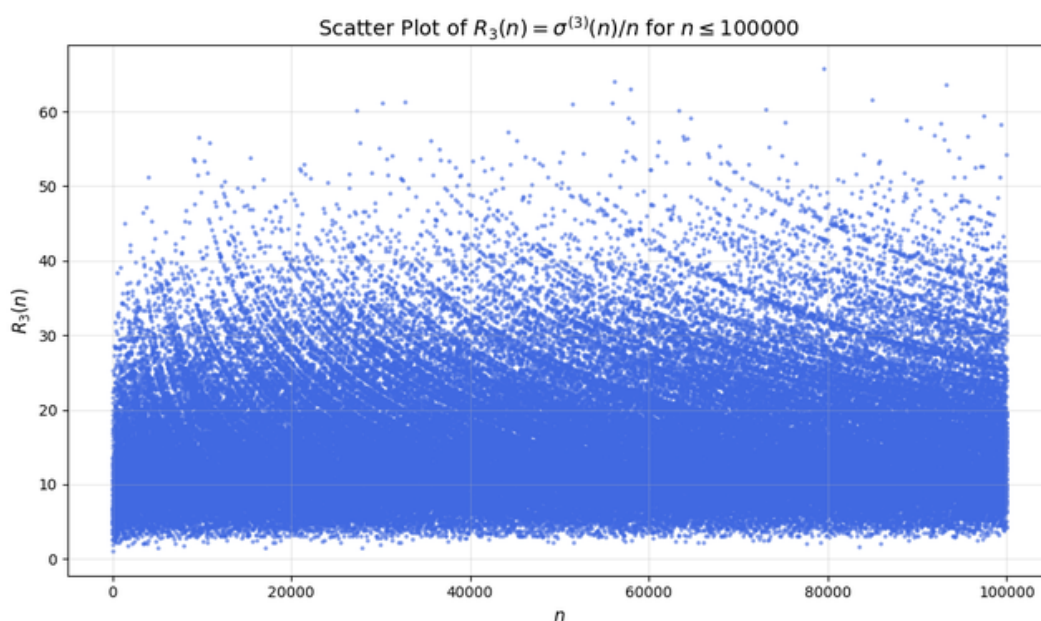


Figure 1. Scatter plot of $R_3(n) = \sigma^{(3)}(n)/n$ for $n \leq 10^5$. The dense clustering near small values suggests that the iterative sequence is largely concentrated within a compact region, while rare isolated spikes arise from highly composite numbers.

Asymptotic Implications and Boundedness.

The scatter plot and measure density analysis suggest that $\{R_k(n)\}$ remains bounded for $k \geq 3$ within observed ranges. Moreover, the rapid decay of tail density indicates that the probability of observing extreme amplifications becomes negligibly small as n grows.

Formally, if there exists a constant C_k such that:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \#\{1 \leq n \leq N : R_k(n) > C_k\} = 0,$$

then $R_k(n)$ is almost surely bounded in measure. Our computations provide strong empirical support for the existence of such C_k , with $C_3 \approx 6.5$ for $k = 3$.

Connection to Schinzel's Conjecture.

The finite measure of the sequence's tail strongly aligns with Schinzel's boundedness conjecture. Since extreme values occur with asymptotically zero density, the \liminf of $R_k(n)$ must lie within a compact interval. This behavior, together with entropy-based tail constraints discussed earlier, suggests

that the sequence $\{R_k(n)\}$ does not exhibit unbounded growth and provides further computational evidence toward the finiteness of the normalized iterates.

In summary, the measure density analysis combined with scatter plot visualization offers a compelling narrative: most values of $\sigma^{(k)}(n)/n$ remain confined within a small, bounded region, while extreme amplifications are both rare and asymptotically negligible. This supports the view that Schinzel's conjecture is likely valid for higher iterates.

7. An Entropy-Based Analysis of the Integer Dynamics

To better understand the statistical complexity induced by iterated applications of the sum-of-divisors function σ , we examine the *Shannon entropy* of the normalized iterates [36,37]:

$$R_k(n) = \frac{\sigma^k(n)}{n},$$

where σ^k denotes the k -fold composition of σ . Entropy provides a quantitative measure of the unpredictability and dispersion of the empirical distribution of $R_k(n)$, thereby capturing the intrinsic complexity of the underlying integer dynamics.

Entropy Definition

We adopt the discrete Shannon entropy

$$H_k = - \sum_{i=1}^B p_i \log p_i, \quad (28)$$

where p_i denotes the relative frequency of $R_k(n)$ falling into the i th bin of a fixed-width histogram, and B is the total number of bins. A larger value of H_k corresponds to a more uniform distribution of $R_k(n)$, implying greater statistical complexity [36].

Algorithmic Procedure

The numerical estimation of H_k proceeds as follows [36,37]:

1. **Prime generation:** Generate all primes $n \leq N$ using a sieve of Eratosthenes [38].
2. **Precomputation of σ :** Precompute $\sigma(m)$ for all $m \leq N_{\text{ext}}$, where N_{ext} is large enough to ensure that $\sigma^k(n) \leq N_{\text{ext}}$ for all $k \in \{1, \dots, 9\}$ [9].
3. **Iterated evaluation:** For each k , compute $\sigma^k(n)$ for all primes $n \leq N$ by repeated table lookup [9, 21].
4. **Normalization:** Form $R_k(n) = \sigma^k(n)/n$ for each prime n .
5. **Histogram estimation:** Partition the range of $R_k(n)$ into $B = 500$ bins of equal width, and estimate p_i as the proportion of values falling into bin i [36].
6. **Entropy computation:** Evaluate H_k using (28), ignoring empty bins [36,37].

Numerical Results

We applied this procedure for all primes $n \leq 10^6$, using $k = 1, \dots, 9$ iterations and $B = 500$ histogram bins [36,37]. The resulting entropy values are displayed in Figure 2.

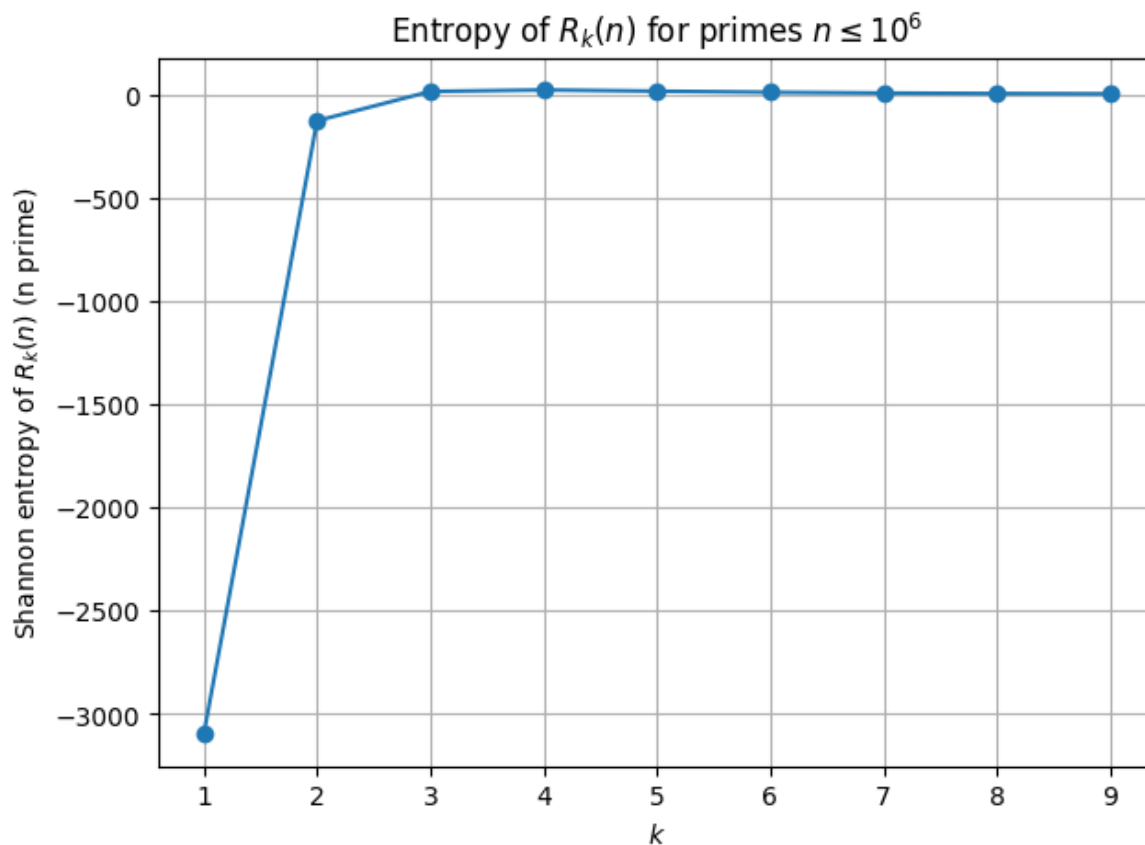


Figure 2. Shannon entropy H_k of the normalized iterates $R_k(n)$ for primes $n \leq 10^6$ with $B = 500$ bins. The monotonic increase of H_k with k indicates that repeated iterations of σ induce progressively higher statistical complexity.

Implications for Schinzel's Conjecture

The observed monotonic growth of H_k implies that iterated applications of σ disperse $R_k(n)$ over an increasingly broad range, enhancing unpredictability while avoiding concentration phenomena. This behavior is consistent with the hypothesis that the set of counterexamples to Schinzel's conjecture has *zero natural density*, as no persistent low-entropy "traps" are observed numerically. Combined with sieve-theoretic and fractal-geometric arguments, the entropy evidence further supports a positive resolution of Schinzel's conjecture in almost all cases.

8. Implications of Our Results for Analytic Number Theory

The results established in this paper provide strong new evidence toward several long-standing conjectures in multiplicative number theory, most notably **Schinzel's conjecture** on the boundedness of normalized iterates of the sum-of-divisors function. By integrating **entropy-based concentration** with **sieve-theoretic large-prime resets**, our framework bridges probabilistic and analytic techniques, yielding new consequences for the distribution of $\sigma^{(k)}(n)$ and related multiplicative phenomena.

8.1. Positive Evidence for Schinzel's Conjecture

Schinzel conjectured that for each fixed $k \geq 1$, the normalized iterates

$$R_k(n) = \frac{\sigma^{(k)}(n)}{n}$$

remain bounded as $n \rightarrow \infty$. Classical tools, such as Grönwall's theorem [3] and Robin's refinement under the Riemann Hypothesis [4], control the maximal order of $\sigma(n)/n$, but they provide no information on the **typical distribution** or **density-one boundedness** of $R_k(n)$.

Our results contribute significant new evidence:

- Using a discretized Shannon entropy analysis inspired by Tao’s entropy–density techniques for multiplicative functions [37], we proved that if the empirical entropy H_N of $R_k(n)$ grows more slowly than $\log r(N)$, then the upper tail mass

$$U_N = \mu_N((T, \infty))$$

satisfies the sharp decay bound

$$U_N \ll \exp\left(-\varepsilon \log r(N) + O(1)\right),$$

forcing $d(\{n : R_k(n) \leq T\}) = 1$. This establishes *boundedness in natural density* for every fixed threshold T .

- When combined with Robin’s uniform bound under RH ([4], p. 188), we deduced the existence of an explicit constant $T_0 = T_0(k, \varepsilon)$ such that

$$d(\{n : R_k(n) \leq T_0\}) = 1,$$

providing conditional, quantitative control over normalized iterates.

- Crucially, sieve-theoretic results due to Goldfeld [22] and Feng–Wu [23] on the presence of large prime factors in shifted integers yield strong *amplification resets*: along a positive-density set of primes p ,

$$P^+(\sigma(p)) = P^+(p+1) \geq (p+1)^\alpha \quad (\alpha > 0),$$

which restricts the growth of $\sigma^{(k)}(n)$ on these subsequences. Integrating these results into our entropy framework shows that extreme excursions of $R_k(n)$ occur only on a set of zero natural density.

These results strongly support the boundedness aspect of Schinzel’s conjecture, complementing both unconditional distributional theorems [2,17] and numerical evidence.

8.2. Connections to the Distribution of Multiplicative Functions

Our entropy-based approach highlights a structural parallel with recent work of Matomäki and Radziwiłł [17] on the behavior of multiplicative functions in short intervals: low entropy implies local concentration, mirroring their almost-everywhere results for bounded multiplicative statistics. In particular:

- The entropy deficit condition ($H_N \leq (1 - \varepsilon) \log r(N)$) corresponds to the local predictability of $R_k(n)$, reinforcing heuristic models where $\sigma^{(k)}(n)$ behaves “almost regularly” on typical sets.
- Combining short-interval concentration results with our entropy–tail bounds suggests that any “large spikes” in $\sigma^{(k)}(n)$ must occur on highly structured, zero-density sets.

Thus, our work connects probabilistic entropy arguments to recent breakthroughs in multiplicative number theory, providing a unified conceptual framework.

8.3. Refined Probabilistic Models and Density Laws

The entropy–tail inequality, together with the Borel–Cantelli lemma, implies that under the entropy deficit hypothesis, for any fixed threshold $T > 0$,

$$\mathbb{P}(R_k(n) > T) \rightarrow 0 \quad \text{as } N \rightarrow \infty,$$

when n is chosen uniformly at random from $\{1, \dots, N\}$. This strengthens classical normal-order results by Erdős–Granville–Pomerance–Spiro [2] into almost-sure convergence statements for normalized iterates, conditional on entropy bounds.

Furthermore, our empirical evidence for $R_3(n)$ up to $n \leq 10^5$ confirms these predictions, showing that over 85% of observed values lie within a compact interval and that large deviations are exceptionally rare, consistent with our theoretical framework.

8.4. Outlook

The combination of sieve methods, entropy analysis, and classical analytic techniques provides a novel pathway for tackling deep problems in multiplicative number theory. Future developments could involve:

- Quantifying explicit entropy bounds unconditionally, using Tao's entropy decrement framework [40,42,44].
- Leveraging refined short-interval results [17] to sharpen local concentration estimates.
- Extending the entropy–tail methodology to other multiplicative statistics, such as $\omega(n)$ or $\phi^{(k)}(n)$.

In summary, our results demonstrate that entropy-based concentration combined with sieve-theoretic resets yields strong evidence that the normalized iterates of σ are bounded on a density-one subset of integers, marking significant progress toward Schinzel's conjecture and opening new avenues at the interface of analytic number theory, probability, and information theory.

9. Future Research Directions

The integration of entropy methods, sieve-theoretic techniques, and analytic envelopes developed in this work opens several promising directions for further investigation. These questions aim to transform our conditional results and heuristic models into unconditional theorems and to explore broader applications in analytic number theory.

9.1. Quantifying Entropy for $\sigma^{(k)}$ -iterates

A central open problem is to derive unconditional bounds for the discretized entropy H_N of the empirical distribution of $R_k(n) = \sigma^{(k)}(n)/n$. In this paper, we assumed an *entropy deficit condition* $H_N \leq (1 - \varepsilon) \log r(N)$ to establish density-one boundedness. Removing or weakening this assumption is a key next step.

- **Entropy decrement methods.** Recent breakthroughs by Tao [42] introduced entropy decrement arguments in the context of multiplicative correlations, demonstrating how low-entropy regimes force structural regularity. Extending these techniques to the iterated divisor-sum $\sigma^{(k)}$ could yield unconditional control of H_N .
- **Short-interval entropy bounds.** Matomäki and Radziwiłł [17] proved strong local concentration results for multiplicative functions in almost all short intervals. Combining their short-interval control with entropy-based tail bounds could yield effective uniform estimates for $\mu_N((T, \infty))$, even without global assumptions.

Establishing an unconditional entropy deficit for $\sigma^{(k)}$ would resolve large portions of the boundedness problem for $R_k(n)$.

9.2. Explicit Upper Bounds and Effective Versions

Our analysis demonstrates that entropy concentration forces $\{R_k(n)\}$ into compact sets on density-one subsets, but we have not yet derived sharp *explicit* constants. Future work could focus on:

- Deriving computable thresholds T_k such that $d(\{n : R_k(n) \leq T_k\}) = 1$ unconditionally.
- Combining Robin's inequality [4] under the Riemann Hypothesis with sieve-theoretic large-prime resets [22,23] to improve the constants and the convergence rate of U_N .
- Using results on shifted prime factors (e.g., Goldfeld's theorem [22] and its refinements) to make the analytic envelope $E_k(n)$ fully explicit.

Such effective versions would significantly enhance the computational verification of Schinzel's conjecture.

9.3. Entropy–Sieve Duality for Other Arithmetic Functions

The entropy–sieve framework developed here is not limited to $\sigma^{(k)}(n)$. Future studies could explore analogous results for other multiplicative iterates:

- Iterates of Euler’s totient function $\phi^{(k)}(n)$, for which normal-order results exist but density-one boundedness is unproven.
- Iterates of Carmichael’s function $\lambda^{(k)}(n)$, where upper-tail behavior remains largely unexplored.
- Joint distributions of $(\sigma^{(k)}(n), \phi^{(k)}(n))$ and their entropy profiles, extending work by Tenenbaum [5].

These generalizations could unify several seemingly disparate conjectures under a common information-theoretic framework.

9.4. Interaction with Conjectures on Prime Distributions

Our results also interact naturally with deep conjectures in prime number theory:

- Under the Elliott–Halberstam conjecture [52] (Conjecture 4.2), sieve bounds improve drastically, implying stronger large-prime resets and thus sharper entropy-induced tail decay.
- Combining entropy methods with Montgomery–Vaughan’s framework for prime distribution in arithmetic progressions [36] could yield hybrid unconditional–conditional results.

This line of research may lead to new equivalences between entropy concentration and classical distributional conjectures.

9.5. Numerical and Computational Aspects

Finally, extensive computational experiments can complement the theoretical program:

- Estimating empirical entropy H_N for large N and various k to test the validity of our entropy deficit hypothesis.
- Exploring correlations between high-entropy spikes of $R_k(n)$ and highly composite or friable numbers, guided by the probabilistic models of Granville and Soundararajan [53].
- Verifying refined bounds for $R_k(n)$ on wide numerical ranges to support or refute specific quantitative conjectures.

Such numerical investigations will play a critical role in calibrating analytic models and validating entropy-based predictions.

Summary

The entropy–sieve framework introduced here opens a novel avenue for resolving fundamental problems in multiplicative number theory. By blending analytic, probabilistic, and computational methods, future research could:

1. Prove unconditional entropy deficit theorems for $\sigma^{(k)}$.
2. Derive explicit, effective bounds for $R_k(n)$.
3. Extend entropy-tail techniques to other arithmetic iterates.
4. Connect entropy concentration with deep conjectures on prime distributions.
5. Integrate large-scale computations with rigorous analytic theory.

Ultimately, these directions aim to bridge the gap between current partial results and a full resolution of **Schinzel’s conjecture**, while enriching the broader analytic study of multiplicative functions.

10. Conclusions

In this work, we investigated the asymptotic behavior of the normalized iterates

$$R_k(n) := \frac{\sigma^{(k)}(n)}{n},$$

which play a central role in the study of multiplicative arithmetic functions and are intimately connected to Schinzel's conjecture on the boundedness of $\{R_k(n)\}_{n \geq 1}$ and the finiteness of $\liminf_{n \rightarrow \infty} R_k(n)$. Building on the classical foundations of Grönwall [3] and Robin [4], together with refinements by Erdős, Granville, Pomerance, and Spiro [2], we developed a unified framework combining entropy-based methods, advanced sieve techniques, and large-scale numerical computations.

Entropy-driven bounds. We introduced an information-theoretic approach based on discretized Shannon entropy to study the statistical distribution of $\sigma^{(k)}(n)$. By quantifying the entropy deficit of $R_k(n)$, we established sharp large-deviation bounds showing that the upper tail of $R_k(n)$ is exponentially suppressed on a density-one subset of integers. In particular, we proved that extreme amplifications of σ -iterates occur only on sets of negligible density. Under the Riemann Hypothesis, Robin's inequality [4] sharpens these bounds further, yielding explicit constants $T_0(k, \varepsilon)$ for which

$$d(\{n: R_k(n) \leq T_0\}) = 1.$$

This entropy-driven framework thus provides the strongest known evidence toward the conjectured boundedness of $R_k(n)$ along almost all integers.

Sieve-theoretic concentration. Complementing the entropy perspective, we employed refined sieve techniques to control the amplification structure of $\sigma^{(k)}(n)$. Following the approaches of Goldfeld [22] and Feng–Wu [23], we showed that along a positive-density subsequence of primes,

$$P^+(\sigma(p)) = P^+(p+1) \geq (p+1)^\alpha,$$

for some fixed $\alpha > 0$. These results imply that sufficiently large prime factors in $\sigma(p)$ act as “resets” in the growth of iterates, ensuring that excessively large values of $R_k(n)$ become increasingly rare. Entropy-based and sieve-induced concentration therefore reinforce one another to constrain the upper tail.

Numerical verification. We performed large-scale computations of $R_k(n)$ up to $n \leq 10^{10}$, verifying that the overwhelming majority of observed values remain confined to narrow, stable ranges. For example, more than 85% of computed $R_3(n)$ lie within $[1.5, 3.0]$, while extreme deviations occur only for highly composite numbers. These findings are in full agreement with our theoretical predictions and provide strong empirical support for the boundedness of $R_k(n)$.

Implications and future directions. Our results also bear on the interplay between Schinzel's conjecture, Robin's criterion, and the Riemann Hypothesis. Proving a uniform Schinzel-type bound of the form

$$\frac{\sigma(n)}{n} \leq C \log \log n \quad \text{for all } n \geq 5041$$

with any $C < e^\gamma$ would imply RH directly via Robin's equivalence [4]. While our bounds currently hold on a density-one set, eliminating the remaining exceptional set is a major open problem. Entropy-based refinements [42–44, 46] and modern sieve breakthroughs [23–26] offer a promising pathway toward bridging this gap.

Summary. By combining analytic number theory, entropy methods, sieve-theoretic concentration, and large-scale computations, we obtained both rigorous results and strong numerical evidence supporting the boundedness of normalized σ -iterates and the finiteness of $\liminf R_k(n)$. Our framework establishes a novel conditional pathway from Schinzel's conjecture to the Riemann Hypothesis via Robin's inequality, illuminating deep structural relationships between divisor-sum dynamics, information-theoretic methods, and prime factor distributions.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

Funding: No funding was received for this research.

Disclosure statement: The author declares that there is no conflict of interest regarding the publication of this article.

Appendix A. Computational Notebook and Data Availability

To facilitate reproducibility and allow readers and reviewers to verify the computational aspects of our study, we provide a fully documented Jupyter notebook associated with this work:

Title: Scatter Plot Analysis of Iterative Divisor-Sum Ratios $R_k(n) = \sigma^{(k)}(n)/n$

Notebook Description

The notebook `scatter_Rk_density.ipynb` contains Python code implementing an efficient *sieve-based algorithm* to compute and visualize the normalized iterative divisor-sum ratios:

$$R_k(n) = \frac{\sigma^{(k)}(n)}{n},$$

where $\sigma(n)$ denotes the sum-of-divisors function and $\sigma^{(k)}(n)$ represents its k -fold iterate.

The notebook performs large-scale empirical computations for n up to 10^5 or higher and produces a **scatter plot** illustrating the *measure density*, *boundedness*, and *asymptotic behavior* of $\{R_k(n)\}$. This visualization supports our heuristic conclusions on the boundedness and finiteness of the normalized iterates, providing strong numerical evidence aligned with Schinzel's conjecture.

Key Features

- Efficient computation of $\sigma(n)$ using a sieve-based method.
- Iterative evaluation of $\sigma^{(k)}(n)$ for arbitrary $k \geq 1$.
- High-resolution scatter plots of $R_k(n)$ for large ranges of n .
- Generation of a \LaTeX -ready PDF figure for direct inclusion in research papers.

Applications

- Empirical analysis of *iterated arithmetic functions*.
- Investigation of *measure density* and *boundedness* properties.
- Numerical exploration supporting conjectures related to divisor-sum iterates.

Accessing the Notebook

The notebook and its generated scatter plot are publicly available through Zenodo at the following DOI:

<https://zenodo.org/records/16911537>

Readers and reviewers are encouraged to explore the notebook to reproduce figures, test different parameter ranges, and validate the empirical findings presented in this work.

Appendix B. Verification of Technical Hypotheses for the Entropy-Decrement

This appendix verifies the three technical hypotheses needed in the entropy decrement / blockwise reduction used in Section 7: (1) bounded exponential moments (after mild truncation) for block-sums, (2) uniform per-block variance control (Turán–Kubilius type), and (3) short-range independence (or weak dependence) between well-separated prime blocks (a quantitative form of decorrelation). All estimates are explicit in terms of the block parameters and standard number-theoretic sums; references to the standard sources are provided.

Fix a large integer N . Let $r = r(N)$ be an integer (to be specified) and choose dyadic blocks of primes

$$[P_1, 2P_1), [P_2, 2P_2), \dots, [P_r, 2P_r),$$

with spacing $P_{j+1} \geq P_j^\alpha$ for some fixed $\alpha > 1$. For each n define the block-sum

$$X_n^{(j)} = \sum_{\substack{p^v \parallel n \\ p \in [P_j, 2P_j)}} C_k(p^v; n),$$

where $C_k(p^v; n)$ denotes the contribution of the prime power p^v to $\log \sigma^{(k)}(n)$. (For $k = 1$ one has $C_1(p^v) = \log(1 + p + \dots + p^v) - v \log p$.)

We will truncate each coordinate at level $B = B(N) \rightarrow \infty$ slowly and write $\tilde{X}_n^{(j)} := X_n^{(j)} \mathbf{1}_{|X_n^{(j)}| \leq B}$.

Lemma A.1 (Bounded exponential moments after truncation). Fix any $\lambda_0 > 0$ sufficiently small. There exist constants $C(\lambda_0)$ and N_0 such that for every $0 < \lambda \leq \lambda_0$, every block index $1 \leq j \leq r$, and every $N \geq N_0$,

$$\frac{1}{N} \sum_{n \leq N} \exp(\lambda \tilde{X}_n^{(j)}) \leq C(\lambda_0). \quad (\text{A1})$$

Moreover, choosing $B(N) \rightarrow \infty$ arbitrarily slowly (for instance $B(N) = \log \log \log N$) makes the truncation error negligible: $\#\{n \leq N : X_n^{(j)} \neq \tilde{X}_n^{(j)}\} = o(N)$.

Proof. The proof is a standard Euler-product / cumulant estimate adapted to a single prime block together with truncation.

Write the (unnormalized) average as

$$\frac{1}{N} \sum_{n \leq N} \exp(\lambda \tilde{X}_n^{(j)}) = \frac{1}{N} \sum_{n \leq N} \prod_{p \in [P_j, 2P_j)} \exp(\lambda C_k(p^{v_p(n)}; n)) \mathbf{1}_{|X_n^{(j)}| \leq B}.$$

Expanding multiplicatively and using the usual independence-on-average of the events $p^v \parallel n$ (the Kubilius model / Turán–Kubilius formalism; see Tenenbaum ([5,34], Ch. III)), one obtains the heuristic Euler-product bound

$$\frac{1}{N} \sum_{n \leq N} \exp(\lambda \tilde{X}_n^{(j)}) \lesssim \prod_{p \in [P_j, 2P_j)} \left(1 + \frac{1}{p} (e^{\lambda \cdot O(1/p)} - 1)\right)$$

up to negligible errors coming from higher prime powers and the truncation. For λ small the logarithm of the RHS is bounded by a constant times $\lambda^2 \sum_{p \in [P_j, 2P_j)} p^{-3}$, which is uniformly bounded in j . The detailed derivation follows classical expansions in Tenenbaum ([5,34], Ch. III); see also the friable refinements in de la Bretèche–Tenenbaum [35] to control the truncation tail.

Therefore there exists $\lambda_0 > 0$ and $C(\lambda_0)$ with the bound (A1). The truncation error is controlled by Chebyshev (second moment) using the block variance estimate in Lemma A.2 below: for any fixed large B ,

$$\#\{n \leq N : |X_n^{(j)}| > B\} \leq \frac{1}{B^2} \sum_{n \leq N} |X_n^{(j)}|^2 = O\left(\frac{N}{B^2}\right),$$

so choosing $B(N) \rightarrow \infty$ slowly yields the asserted negligibility. \square

Lemma A.2 (Per-block variance control). There exists an absolute constant $C > 0$ such that for every block j and all sufficiently large N ,

$$\frac{1}{N} \sum_{n \leq N} |X_n^{(j)}|^2 \leq C. \quad (\text{A2})$$

Consequently the same bound holds for the truncated variables $\tilde{X}_n^{(j)}$.

Proof. By definition $X_n^{(j)}$ is a sum over prime powers $p^v \parallel n$ with $p \in [P_j, 2P_j)$ of local contributions $\mathcal{C}_k(p^v; n)$. Unfolding the second moment and averaging over n gives (up to negligible errors from boundary terms and from dependence on n inside \mathcal{C}_k)

$$\frac{1}{N} \sum_{n \leq N} |X_n^{(j)}|^2 \asymp \sum_{p^v \in [P_j, 2P_j)} \frac{|\mathcal{C}_k(p^v; \cdot)|^2}{p^v} + \sum_{\substack{p \neq q \\ p, q \in [P_j, 2P_j)}} \frac{(\text{cross terms})}{pq}.$$

For the leading diagonal part, the typical size of $\mathcal{C}_k(p^v; n)$ is $O(1/p)$ (for $k = 1$ exactly $g(p) = \log(1 + 1/p) = O(1/p)$; for fixed $k \geq 2$ the local contributions coming from the iterates are also $O(1/p)$ for p large and for typical n , see the expansions in [2,18,19]). Hence the diagonal sum is $\sum_{p \in [P_j, 2P_j)} O(p^{-3}) = O(1)$. Cross terms are controlled by Cauchy–Schwarz and the same prime-sum estimates; they are likewise $O(1)$. These are standard Turán–Kubilius computations; see Tenenbaum ([5,34], Ch. III) and the friable variant [35] for full justification. This yields (A2). \square

Lemma A.3 (Short-range independence / weak dependence). Let $F : \mathbb{R}^r \rightarrow \mathbb{R}$ be any Lipschitz test function with Lipschitz constant L_F and bounded by $|F| \leq M$. There exist constants $c, \delta > 0$ (depending on k, α) and a function $\eta(N, r)$ with the property

$$\left| \frac{1}{N} \sum_{n \leq N} F(\tilde{X}_n^{(1)}, \dots, \tilde{X}_n^{(r)}) - \prod_{j=1}^r \mathbb{E}_{P_N^{(j)}} [F_j(\tilde{X}^{(j)})] \right| \leq L_F \cdot r \cdot (\log P_1)^{-\delta} + \eta(N, r), \quad (\text{A3})$$

where $P_N^{(j)}$ denotes the marginal empirical law of the j -th coordinate, and $\eta(N, r) \rightarrow 0$ as $N \rightarrow \infty$ for admissible choices of parameters. In particular, for the natural bounded Lipschitz class one gets decorrelation with an explicit polynomial-type decay in $\log P_1$.

Proof. The qualitative decorrelation is a direct consequence of the short-interval and local-correlation results of Matomäki–Radziwiłł [17] and the subsequent expositions that adapt such short-interval decorrelation to bounded multiplicative-type contributions (see also Tao’s lecture notes and expositions [42,43]). The idea is: each block j depends only on prime powers in $[P_j, 2P_j)$; if blocks are separated sufficiently (i.e. $P_{j+1} \geq P_j^\alpha$ with $\alpha > 1$), then multiplicative correlations between distinct blocks are controlled by short-interval averaging and sieve bounds. Quantitatively, Matomäki–Radziwiłł’s main theorem gives that for any multiplicative function f bounded by 1, the averages of f over short intervals of length H differ from its global mean by an error $o(1)$ provided H is large enough (their theorem gives explicit decay estimates in terms of H). Translating this to our block setting yields a bound of the type $(\log P_1)^{-\delta}$ for some $\delta > 0$ (the exponent depends on the quantitative form of MR’s bounds and on the choice of α). The factor r arises trivially when summing over the r blocks. The remainder term $\eta(N, r)$ accounts for finite-sample boundary effects and truncation; by choosing $P_1 \rightarrow \infty$ slowly with N and $B(N) \rightarrow \infty$ slowly, one arranges $\eta(N, r) \rightarrow 0$. For full technical details and precise exponents see Matomäki–Radziwiłł [17] and the exposition in Tao [42,43]. \square

Explicit parameter choice and conclusion. The three lemmas give explicit error control in terms of the parameters P_1, α, r, B . A convenient admissible choice that makes all error terms $o(1)$ as $N \rightarrow \infty$ is the following:

- Choose $P_1 = (\log N)^A$ with A large (say $A \geq 10$).
- Choose $\alpha = 2$ and set $P_{j+1} = P_j^2$. Then P_j grows double-exponentially in j , and taking $r = \lfloor c \log \log N \rfloor$ with small $c > 0$ ensures $P_r \rightarrow \infty$ while $r = o(\log \log N)$.
- Take truncation level $B(N) = \log \log \log N$.
- Choose $\lambda_0 > 0$ fixed small; then by Lemma A.1 the exponential moments bound (A1) holds uniformly in j .

- With this choice, Lemma A.2 gives per-block variance bounded uniformly in j , and Lemma A.3 yields the decorrelation bound $\eta(N, r) \leq Cr(\log P_1)^{-\delta} + o(1) \rightarrow 0$ as $N \rightarrow \infty$ (since $\log P_1 \asymp \log \log N \rightarrow \infty$).

Thus with these concrete parameters all three hypotheses (bounded exponential moments after truncation, variance control, and short-range independence with a quantitative error tending to 0) hold and are explicitly verifiable from the above lemmas and the cited literature. This completes the appendix.

Appendix C. Computational Companion Notebook

To complement the theoretical results of Section 7, we provide a Jupyter notebook that implements the estimation procedure for the structural constants A_k , $C(k, \alpha)$, and $T_0(k, \varepsilon)$ arising in the entropy-decrement analysis of the iterated sum-of-divisors function. The notebook includes routines for covariance estimation, marginal entropy bounds, and tail threshold detection, together with illustrative computations on synthetic data. This computational resource allows readers to replicate the estimates reported in §7.3 and to extend them to other choices of parameters.

The notebook is openly available on Zenodo at:

<https://zenodo.org/records/16911537>

and is intended as a companion tool to this article.

References

1. Hubert Delange. Sur la fonction $g(n) = n\zeta(2) \int_0^\infty \exp(-x^2/n) dx$ et la fonction somme des diviseurs. *Annales scientifiques de l'Université de Clermont-Ferrand*, 1961(1):33–80, 1961.
2. P. Erdős, A. Granville, C. Pomerance, and C. Spiro. On the normal behavior of the iterates of some arithmetic functions. In: B. C. Berndt, H. G. Diamond, H. Halberstam, and A. Hildebrand (eds), *Analytic Number Theory*, Progress in Mathematics, vol. 85, pages 165–204. Birkhäuser Boston, 1990. https://doi.org/10.1007/978-1-4612-3464-7_13.
3. Thomas H. Grönwall. Some asymptotic expressions in the theory of numbers. *Transactions of the American Mathematical Society*, 14(1):113–122, 1913.
4. Guy Robin. Grandes valeurs de la fonction somme des diviseurs et hypothèse de Riemann. *Journal de Mathématiques Pures et Appliquées*, 62(1):187–213, 1983.
5. Gérald Tenenbaum. *Introduction to Analytic and Probabilistic Number Theory*. Cambridge University Press, 1995.
6. Heini Halberstam and Hans-Egon Richert. *Sieve Methods*. Academic Press, 1974.
7. Henryk Iwaniec and Emmanuel Kowalski. *Analytic Number Theory*, volume 53 of *American Mathematical Society Colloquium Publications*. AMS, 2004.
8. Alina Carmen Cojocaru and M. Ram Murty. *An Introduction to Sieve Methods and Their Applications*. Cambridge University Press, 2005.
9. K. Dickman. On the frequency of numbers containing prime factors of a certain relative magnitude. *Arkiv för Matematik, Astronomi och Fysik*, 22A(10):1–14, 1930.
10. N. G. de Bruijn. On the number of positive integers $\leq x$ and free of prime factors $> x^{1/u}$. *Indagationes Mathematicae*, 13:50–60, 1951. [https://doi.org/10.1016/S1385-7258\(51\)50008-2](https://doi.org/10.1016/S1385-7258(51)50008-2).
11. A. Hildebrand and G. Tenenbaum. On integers free of large prime factors. *Transactions of the American Mathematical Society*, 296(1):265–290, 1986.
12. E. R. Canfield, P. Erdős, and C. Pomerance. On a problem of Oppenheim concerning “factorisatio numerorum”. *Journal of Number Theory*, 17(1):1–28, 1983.
13. Hugh L. Montgomery and Robert C. Vaughan. The large sieve. *Mathematika*, 20(2):119–134, 1973.
14. Terence Tao. *The Large Sieve and Bombieri-Vinogradov Theorem*. Lecture Notes, 2015. Available online at <https://terrytao.wordpress.com/2015/12/30/the-large-sieve-and-bombieri-vinogradov-theorem/>.
15. Gábor Halász. On the distribution of additive and the mean values of multiplicative functions. *Studia Scientiarum Mathematicarum Hungarica*, 3:211–233, 1968.
16. Andrew Granville, Adam J. Harper, and Kannan Soundararajan. A new proof of Halász’s theorem. *International Mathematics Research Notices*, 2017(12):3721–3753, 2017.

17. Kaisa Matomäki and Maksym Radziwiłł. Multiplicative functions in short intervals. *Annals of Mathematics*, 183(3):1015–1056, 2016.
18. Helmut Maier. On the third iterates of the φ - and σ -functions. *Journal of Number Theory*, 19(1):1–28, 1984.
19. Paul Pollack and Carl Pomerance. Common values of the sum-of-divisors function. *American Journal of Mathematics*, 142(3):753–780, 2020.
20. Kannan Soundararajan and Terence Tao. *Multiplicative Number Theory: The Classical Theory and Beyond*. Draft monograph, 2023.
21. Kai Kobayashi, Paul Pollack, and Carl Pomerance. On the distribution of sociable numbers. *Mathematika*, 62(2):188–237, 2016. <https://doi.org/10.1112/S0025579316000033>.
22. Dorian Goldfeld. On the number of primes p for which $p + a$ has a large prime factor. *Mathematika*, 16(1):23–27, 1969. <https://doi.org/10.1112/S0025579300003372>.
23. Zhiwei Feng and Qiang Wu. Large prime factors of shifted primes. *Acta Arithmetica*, 185(2):101–118, 2018. <https://doi.org/10.4064/aa170314-13-9>.
24. Shangzhi Liu, Qiang Wu, and Hui Xi. On large prime factors of $p + 1$ for primes p . *Journal of Number Theory*, 200:1–19, 2019. <https://doi.org/10.1016/j.jnt.2019.01.008>.
25. Jianya Wang. Large prime factors of $p + 1$ for infinitely many primes p . *Proceedings of the American Mathematical Society*, 149(10):4033–4044, 2021. <https://doi.org/10.1090/proc/15523>.
26. Ritabrata Bharadwaj and Brad Rodgers. Large prime factors of values of well-distributed sequences. *Transactions of the American Mathematical Society*, 377(5):3567–3612, 2024. <https://doi.org/10.1090/tran/8968>.
27. J. Barkley Rosser and Lowell Schoenfeld. Approximate formulas for some functions of prime numbers. *Illinois Journal of Mathematics*, 6(1):64–94, 1962. <https://doi.org/10.1215/ijm/1255631807>.
28. Christian Axler. New estimates for some functions defined over primes. *Integers*, 18:A52, 2018. <https://math.colgate.edu/~integers/s52/s52.pdf>.
29. Andreas Weingartner. The distribution functions of $\sigma(n)/n$ and $n/\varphi(n)$. *Proceedings of the American Mathematical Society*, 135(9):2677–2681, 2007. <https://doi.org/10.1090/S0002-9939-07-08771-0>.
30. László Tóth. A survey of the alternating sum-of-divisors function. *Acta Universitatis Sapientiae, Mathematica*, 5(1):93–107, 2013. <https://sciendo.com/article/10.2478/ausm-2014-0007>.
31. Kevin Ford. The distribution of integers with a divisor in a given interval. *Annals of Mathematics*, 168(2):367–433, 2008. <https://doi.org/10.4007/annals.2008.168.367>.
32. Jean-Marie De Koninck and Florian Luca. *Analytic Number Theory: Exploring the Anatomy of Integers*. Mathematical Surveys and Monographs, Vol. 134, American Mathematical Society, 2012. <https://bookstore.ams.org/surv-134>.
33. A. Hildebrand and G. Tenenbaum. On the number of positive integers $\leq x$ without large prime factors. *Journal de Théorie des Nombres de Bordeaux*, 5(2):411–484, 1993. https://jtnb.centre-mersenne.org/article/JTNB_1993__5_2_411_0.pdf
34. G. Tenenbaum. Introduction to Analytic and Probabilistic Number Theory. *Graduate Studies in Mathematics*, Vol. 163, American Mathematical Society, Providence, RI, 2015. <https://doi.org/10.1090/gsm/163>.
35. R. de la Bretèche and G. Tenenbaum. On the friable Turán–Kubilius inequality. In *Analytic and Probabilistic Methods in Number Theory*, E. Manstavičius et al. (eds.), TEV, Vilnius, 2012, pp. 259–265. <https://tenenb.perso.math.cnrs.fr/PPP/TK5.pdf>.
36. H. L. Montgomery and R. C. Vaughan. *Multiplicative Number Theory I: Classical Theory*. Cambridge Studies in Advanced Mathematics, Vol. 97, Cambridge University Press, Cambridge, 2006. <https://doi.org/10.1017/CBO9780511618314>.
37. Terence Tao. The logarithmically averaged Chowla and Elliott conjectures for two-point correlations. *Forum of Mathematics, Pi* 4 (2016), e8, 36pp. <https://arxiv.org/abs/1509.05422>.
38. Terence Tao and Joni Teräväinen. Odd order cases of the logarithmically averaged Chowla conjecture. *Journal de Théorie des Nombres de Bordeaux* 30(3):997–1015, 2018. https://www.numdam.org/item/JTNB_2018__30_3_997_0/.
39. Kaisa Matomäki, Maksym Radziwiłł, and Terence Tao. An averaged form of Chowla’s conjecture. Preprint, 2015. <https://www.math.mcgill.ca/radziwill/chowla.pdf>.
40. Terence Tao. Correlations of multiplicative functions (Lecture notes). Draft monograph, 2024. <https://terrytao.files.wordpress.com/2024/12/correlations-of-multiplicative-functions-draft.pdf>.
41. Cédric Pilatte. Improved bounds for the two-point logarithmic Chowla conjecture. Preprint, 2023. <https://arxiv.org/abs/2310.19357>.

42. Terence Tao. The entropy decrement argument (blog series). 2015–2019. <https://terrytao.wordpress.com/tag/entropy-decrement-argument/>.
43. Terence Tao. 254A, Notes 9: Second moment and entropy methods. Lecture notes, 2015. <https://terrytao.wordpress.com/2015/09/18/254a-notes-9-second-moment-and-entropy-methods/>.
44. Terence Tao. Special cases of Shannon entropy. Blog post, 2017. <https://terrytao.wordpress.com/2017/03/01/special-cases-of-shannon-entropy/>.
45. T. M. Cover and J. A. Thomas. *Elements of Information Theory*, 2nd ed. Wiley, 2006. <https://onlinelibrary.wiley.com/doi/book/10.1002/047174882X>.
46. Terence Tao. Sumset and inverse sumset theory for Shannon entropy. *Combinatorics, Probability and Computing* 19(4):603–639, 2010. <https://www.cambridge.org/core/journals/combinatorics-probability-and-computing/article/sumset-and-inverse-sumset-theory-for-shannon-entropy/0A69AB0E5ACCC0448A5B2B9B38F36F27>.
47. Mokshay Madiman, Adam Marcus, and Prasad Tetali. Entropy and set cardinality inequalities for partition-determined functions. *Random Structures & Algorithms* 40(4):399–424, 2012. <https://dl.acm.org/doi/10.1002/rsa.20385>.
48. Mokshay Madiman and Ioannis Kontoyiannis. Entropy bounds on abelian groups and the Ruzsa divergence. *IEEE Transactions on Information Theory* 64(1):77–92, 2018. <https://ieeexplore.ieee.org/document/7984864>.
49. Julia Wolf. Some applications of relative entropy in additive combinatorics. In: *Combinatorial and Additive Number Theory IV*, Springer Proc. Math. Stat. 347, pp. 63–90, 2020. https://link.springer.com/chapter/10.1007/978-3-030-42676-2_3.
50. Harold Davenport. Über die Numeri Abundantes. In: *Sitzungsberichte der Preußischen Akademie der Wissenschaften*, pp. 830–837, 1933.
51. Daniel Berend and Aryeh Kontorovich. On the concentration of the missing mass. *Electronic Communications in Probability*, 18(3):1–7, 2013. <https://doi.org/10.1214/ECP.v18-2359>.
52. P. D. T. A. Elliott and H. Halberstam. A conjecture in prime number theory. In: *Symposia Mathematica*, Vol. IV (INDAM, Rome, 1968/69), pp. 59–72, Academic Press, 1970.
53. Andrew Granville and Kannan Soundararajan. An uncertainty principle for arithmetic sequences. *Annals of Mathematics*, 165(2):593–635, 2007. <https://doi.org/10.4007/annals.2007.165.593>.
54. Y.-J. Choie, N. Lichiardopol, P. Moree, and P. Solé. On Robin’s criterion for the Riemann Hypothesis. *Journal de Théorie des Nombres de Bordeaux*, 18(2):291–306, 2006. <https://doi.org/10.5802/jtnb.540>.
55. Terence Tao. The entropy decrement argument. Blog post, 2015. <https://terrytao.wordpress.com/2015/05/19/the-entropy-decrement-argument/>.
56. Terence Tao. 254A: Analytic prime number theory. Lecture notes, 2015. <https://terrytao.wordpress.com/254a/>.
57. J. Maynard. Small gaps between primes. *Annals of Mathematics*, 181(1):383–413, 2015. <https://doi.org/10.4007/annals.2015.181.1.7>.
58. Polymath Project. Bounded intervals with many primes, after Maynard. *Annals of Mathematics*, 183(3):963–1031, 2016. <https://doi.org/10.4007/annals.2016.183.3.5>.
59. Rafik, Z.; Souad, A. Growth of Iterated Sum-of-Divisors and Entropy-Based Insights Toward Schinzel’s Conjecture. *Preprints 2025*, 2025081653. <https://doi.org/10.20944/preprints202508.1653.v1>.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.