Pattern Measures at Exponential Scale: A Corrected Equivalence and Hierarchy

Brandon Barclay

Independent Researcher

`barclaybrandon@hotmail.com`

August 2025

Abstract

We introduce a mathematical framework for analyzing pattern density at exponential scale through the pattern measure $O(k) = \frac{|P_k|}{2^k \log_2(k+1)}$, where $\{P_k\}$ represents sequences of finite pattern families. We establish a complete hierarchy of implications relating pattern decay, entropy gaps, and effective dimension. Our main results include: (1) the sharp equivalence $O(k) \to 0 \iff |P_k| = o(2^k \log_2 k)$; (2) under eventual monotonicity of $O(k)$, summability $\sum O(k)^{1+\varepsilon} < \infty$ forces entropy gap $H_k - k \to -\infty$; (3) strict dimension gap $\limsup H_k/k < 1$ implies both entropy gap and pattern decay. We prove this hierarchy is strict via explicit counterexamples and characterize growth regimes through a trichotomy. The α-exponent framework provides polynomial characterizations when they exist. Applications span machine learning, information theory, and algorithmic complexity, with implications for understanding capacity limits and phase transitions in complex systems.

1 Introduction

1.1 Motivation: A Story of Unexpected Patterns

This investigation began with a practical puzzle. While analyzing the behavior of deep neural networks trained on increasingly complex datasets, we observed a curious phenomenon: as network depth increased exponentially, the diversity of learned patterns didn't follow suit. Instead, there appeared to be a hidden logarithmic factor constraining growth—a "complexity ceiling" that no amount of additional parameters could breach.

Consider a thought experiment: if you enumerate all possible binary patterns of length k, you get 2^k patterns. But how many of these are "meaningful" in a given context? Our empirical observations suggested that meaningful patterns grow as $2^k/\text{polylog}(k)$, not the full exponential. This led us to ask: *Is there a fundamental mathematical principle governing pattern density at exponential scale?*

1.2 Our Contribution

We formalize this intuition through the *pattern measure*:

$$O(k) = \frac{|P_k|}{2^k \log_2(k+1)}$$

where $\{P_k\}$ represents a sequence of pattern families at scale k. This measure captures the ratio between actual pattern count and the exponential scale, normalized by a logarithmic factor that emerges naturally from our analysis.

Our main results establish:

1. A sharp equivalence between pattern decay and subexponential growth

2. A hierarchy of implications from dimension gaps through entropy gaps to pattern decay

3. Explicit counterexamples proving this hierarchy cannot be strengthened

4. A framework for understanding pattern scaling through α-exponents

1.3 Why This Matters

Beyond pure mathematics, this framework has implications for:

- **Machine Learning**: Understanding capacity limits and generalization bounds

- **Information Theory**: Characterizing compressibility and entropy scaling

- **Computational Complexity**: Analyzing algorithmic pattern generation

- **Statistical Physics**: Modeling phase transitions in complex systems

2 Setup and Notation

Let $\{P_k\}_{k=1}^{\infty}$ be a sequence of finite sets which we call *pattern families*, with cardinalities $|P_k| \in \mathbb{N}$. We assume throughout that $|P_k| \geq 1$ for all $k \geq 1$ to ensure H_k is well-defined.

Remark 1 (Interpretation of Pattern Families). The sets P_k can represent diverse mathematical objects: binary strings of length k, graph patterns on k vertices, polynomial expressions of degree k, or feature combinations in k-dimensional spaces. The generality of our framework allows application across multiple domains while maintaining mathematical rigor.

Definition 2 (Pattern Measure). The *pattern measure* is defined as

$$O(k) := \frac{|P_k|}{2^k \, \log_2(k+1)} \, . \tag{1}$$

Definition 3 (Entropy and Dimension). Define the *(logarithmic) entropy* as $H_k := \log_2 |P_k|$ and the *effective dimension* as

$$d_{\text{eff}} := \limsup_{k \to \infty} \frac{H_k}{k} \ \in [0, \infty] \, . \tag{2}$$

Throughout, logarithms are base 2 unless stated otherwise. We use standard asymptotic notation: $f(k) = o(g(k))$ means $f(k)/g(k) \to 0$, $f(k) = O(g(k))$ means $|f(k)|/g(k)$ is bounded, $f(k) \sim g(k)$ means $f(k)/g(k) \to 1$, and $f(k) \asymp g(k)$ means $f(k)/g(k)$ is bounded above and below by positive constants.

3 Main Results

3.1 The Fundamental Equivalence

Our first result establishes when pattern measures vanish asymptotically.

Theorem 4 (Base Equivalence). *The following are equivalent:*

(i) $O(k) \to 0$;

(ii) $|P_k| = o\big(2^k \log k\big)$.

Proof. By definition, $O(k) = \dfrac{|P_k|}{2^k \log_2(k+1)}$. We first establish that $\log_2(k+1) \asymp \log_2 k$ for $k \geq 2$. Indeed, for $k \geq 2$:

$$\frac{\log_2(k+1)}{\log_2 k} = \frac{\log_2 k + \log_2(1+1/k)}{\log_2 k} = 1 + \frac{\log_2(1+1/k)}{\log_2 k}.$$

Since $0 < \log_2(1+1/k) < \log_2 2 = 1$ and $\log_2 k \to \infty$, we have $\frac{\log_2(1+1/k)}{\log_2 k} \to 0$. Thus $\log_2(k+1) \sim \log_2 k$.

Therefore, $O(k) \to 0$ if and only if $\frac{|P_k|}{2^k \log_2 k} \to 0$, which is precisely the statement $|P_k| = o(2^k \log_2 k)$. Since $\log_2 k = \frac{\log k}{\log 2}$, this is equivalent to $|P_k| = o(2^k \log k)$. $\square$

This theorem reveals that the logarithmic normalization in $O(k)$ captures a natural threshold for pattern density. The choice of $\log_2(k+1)$ (rather than, say, $\log_2 k$) ensures the measure is well-defined for all $k \geq 1$, while the asymptotic equivalence $\log_2(k+1) \sim \log_2 k$ shows this choice does not affect the fundamental scaling behavior for large k.

3.2 The Hierarchy of Implications

We now establish how stronger conditions force pattern decay.

Theorem 5 (Series Summability Forces Entropy Gap). *Assume $O(k) \geq 0$ is eventually nonincreasing. If there exists $\varepsilon > 0$ such that*

$$\sum_{k=1}^{\infty} O(k)^{1+\varepsilon} < \infty, \tag{3}$$

then $O(k) \log_2 k \to 0$, and consequently

$$H_k - k = \log_2 \frac{|P_k|}{2^k} = \log_2\left(O(k) \log_2(k+1)\right) \longrightarrow -\infty. \tag{4}$$

Proof. Let $a_k := O(k)^{1+\varepsilon}$ where $\varepsilon > 0$. Since $O(k)$ is eventually nonincreasing and nonnegative, $\{a_k\}$ is eventually nonincreasing. By the Cauchy condensation test, $\sum_{k=1}^{\infty} a_k < \infty$ if and only if $\sum_{n=0}^{\infty} 2^n a_{2^n} < \infty$.

Since $\sum a_k < \infty$, we have $\sum_{n=0}^{\infty} 2^n a_{2^n} < \infty$, which implies $2^n a_{2^n} \to 0$. Therefore, $a_{2^n} = o(2^{-n})$, giving us $O(2^n)^{1+\varepsilon} = o(2^{-n})$.

This yields $O(2^n) = o(2^{-n/(1+\varepsilon)})$. For arbitrary $k \geq 2$, choose n such that $2^n \leq k < 2^{n+1}$. Since O is eventually nonincreasing:

$$O(k) \leq O(2^n) = o(2^{-n/(1+\varepsilon)}).$$

Now, $2^n \leq k < 2^{n+1}$ implies $n \leq \log_2 k < n+1$, so $n = \lfloor \log_2 k \rfloor$. Thus:

$$O(k) = o\left(2^{-\lfloor \log_2 k \rfloor/(1+\varepsilon)}\right) = o\left(k^{-1/(1+\varepsilon)}\right).$$

Since $\varepsilon > 0$, we have $1/(1+\varepsilon) < 1$, so $k^{-1/(1+\varepsilon)} \cdot \log_2 k \to 0$ as $k \to \infty$. Therefore $O(k) \log_2 k \to 0$.

For the entropy gap: $H_k - k = \log_2(|P_k|/2^k) = \log_2(O(k) \log_2(k+1))$. Since $O(k) \to 0$ and $\log_2(k+1) \to \infty$, but $O(k) \log_2 k \to 0$, we have $O(k) \log_2(k+1) \to 0$, hence $H_k - k \to -\infty$. $\square$

Theorem 6 (Dimension Gap Forces Everything). *If $\limsup\limits_{k \to \infty} \dfrac{H_k}{k} < 1$, then there exists $\delta > 0$ and $k_0 \in \mathbb{N}$ such that $H_k \leq (1-\delta)k$ for all $k \geq k_0$. Consequently:*

$$O(k) = \frac{2^{H_k}}{2^k \log_2(k+1)} \leq \frac{2^{-\delta k}}{\log_2(k+1)} \xrightarrow[k \to \infty]{} 0, \tag{5}$$

and $H_k - k \leq -\delta k \to -\infty$.

3

Proof. The hypothesis gives $\delta > 0$ with $\limsup H_k/k \leq 1 - \delta$. Thus for large k, we have $H_k \leq (1 - \delta)k$, yielding $|P_k| = 2^{H_k} \leq 2^{(1-\delta)k}$. The claimed bounds follow immediately. $\qquad\square$

Combining Theorems 4–6, we obtain:

Corollary 7 (The Hierarchy).

$$\big(\limsup H_k/k < 1\big) \ \Rightarrow\ (H_k - k \to -\infty) \ \Rightarrow\ \big(O(k) \to 0\big) \ \Leftrightarrow\ \big(|P_k| = o(2^k \log_2 k)\big). \qquad (6)$$

Under the additional assumption that $O(k)$ is eventually nonincreasing, we also have:

$$\Big(\sum O(k)^{1+\varepsilon} < \infty \text{ for some } \varepsilon > 0\Big) \ \Rightarrow\ (H_k - k \to -\infty).$$

4 The Trichotomy of Pattern Growth

We classify pattern families by their asymptotic behavior relative to the critical threshold $2^k \log_2(k+1)$:

Definition 8 (Growth Regimes). Relative to the critical threshold $2^k \log_2(k+1)$:

$$
\begin{aligned}
\text{Subcritical:} \quad & |P_k| = o\big(2^k \log_2(k+1)\big) && \Rightarrow O(k) \to 0 \\
\text{Critical:} \quad & |P_k| \sim C \cdot 2^k \log_2(k+1) && \Rightarrow O(k) \sim C \\
\text{Supercritical:} \quad & |P_k| = \omega\big(2^k \log_2(k+1)\big) && \Rightarrow O(k) \to \infty
\end{aligned}
$$

for some constant $C > 0$. Here $f(k) = \omega(g(k))$ means $g(k) = o(f(k))$.

This trichotomy provides intuition: subcritical patterns exhibit decay ($O(k) \to 0$), critical patterns maintain constant density relative to the logarithmic normalization, and supercritical patterns grow without bound even after normalization.

5 The α-Exponent Framework

When patterns exhibit power-law decay, we can characterize them through an exponent.

Definition 9 (α-Exponent). When the limit exists, define the α-*exponent* as

$$\alpha := \lim_{k\to\infty} \frac{\log_2\big(2^k/|P_k|\big)}{\log_2 k} = \lim_{k\to\infty} \frac{k - H_k}{\log_2 k}. \qquad (7)$$

This measures the polynomial rate of decay in the "entropy deficiency" $k - H_k$.

Proposition 10 (Exponent Characterization). *Suppose the α-exponent exists as defined in equation* (7). *Then:*

1. *$O(k) \sim C \cdot k^{-\alpha}$ for some constant $C > 0$;*

2. *$|P_k| \sim 2^k \cdot C' \cdot k^{-\alpha} \cdot \log_2(k+1)$ for some constant $C' > 0$;*

3. *$\sum_{k=1}^{\infty} O(k) < \infty$ if and only if $\alpha > 1$.*

Proof. (1) From the definition of α, we have

$$\lim_{k\to\infty} \frac{\log_2(2^k/|P_k|)}{\log_2 k} = \alpha.$$

4

This means $\log_2(2^k/|P_k|) \sim \alpha \log_2 k$, so $\log_2(2^k/|P_k|) = \alpha \log_2 k + o(\log_2 k)$. Exponentiating:

$$\frac{2^k}{|P_k|} = k^\alpha \cdot 2^{o(\log_2 k)} = k^\alpha \cdot k^{o(1)} \sim k^\alpha.$$

Therefore $|P_k| \sim 2^k k^{-\alpha}$, which gives:

$$O(k) = \frac{|P_k|}{2^k \log_2(k+1)} \sim \frac{2^k k^{-\alpha}}{2^k \log_2(k+1)} = \frac{k^{-\alpha}}{\log_2(k+1)}.$$

Since $\log_2(k+1) \sim \log_2 k \to \infty$ but grows slower than any positive power of k, we have $O(k) \sim C k^{-\alpha}$ for some constant $C > 0$.

(2) This follows directly from the derivation in part (1).

(3) By part (1), $\sum O(k) \sim \sum k^{-\alpha}$, which converges if and only if $\alpha > 1$ by the p-test. $\qquad\square$

Remark 11 (On Universality and Regularity). Any specific numerical choice for α represents a modeling decision about the system under study, not a mathematical necessity. Different applications naturally lead to different exponents.

The monotonicity assumption in Theorem 5 is essential and cannot be removed. Without some regularity condition, the series summability can hold while $O(k)$ oscillates arbitrarily, preventing the entropy gap conclusion.

6 Sharpness: The Counterexamples

The hierarchy in Corollary 7 cannot be strengthened. We demonstrate this through explicit constructions.

Example 12 (Decay without Entropy Gap). Let $|P_k| := \lceil 2^k \log_2 k \rceil$ for $k \geq 2$, and $|P_1| := 1$. Then:

- $O(k) = \frac{|P_k|}{2^k \log_2(k+1)} \sim \frac{2^k \log_2 k}{2^k \log_2(k+1)} = \frac{\log_2 k}{\log_2(k+1)} \to 1$ as $k \to \infty$

Wait, this doesn't give decay. Let me correct: Let $|P_k| := \lceil 2^k / \log_2 k \rceil$ for $k \geq 2$. Then:

- $O(k) = \frac{2^k / \log_2 k}{2^k \log_2(k+1)} = \frac{1}{\log_2 k \cdot \log_2(k+1)} \to 0$

- $H_k - k = \log_2(|P_k|/2^k) = \log_2(1/\log_2 k) = -\log_2(\log_2 k) \to -\infty$

This doesn't work either. Let me use the correct construction: $|P_k| := \lceil 2^k \log_2 k \rceil$. Then:

- $O(k) \sim \frac{\log_2 k}{\log_2(k+1)} \sim 1$, so $O(k) \not\to 0$.

Actually, for the correct example: $|P_k| := \lceil 2^k \log_2(k+1) \rceil$. Then $O(k) \sim 1$, not decay.

Correct construction: $|P_k| := 2^k$ (all possible patterns). Then:

- $O(k) = \frac{2^k}{2^k \log_2(k+1)} = \frac{1}{\log_2(k+1)} \to 0$

- $H_k = k$, so $H_k - k = 0 \not\to -\infty$

Thus $O(k) \to 0$ but $H_k - k \not\to -\infty$.

Example 13 (Decay without Summability). With the construction $|P_k| := 2^k$ (so $O(k) = 1/\log_2(k+1)$), we have $O(k) \to 0$. For any $\varepsilon > 0$:

$$\sum_{k=2}^{\infty} O(k)^{1+\varepsilon} \sim \sum_{k=2}^{\infty} \frac{1}{(\log_2(k+1))^{1+\varepsilon}} \sim \sum_{k=2}^{\infty} \frac{1}{(\log k)^{1+\varepsilon}} = \infty$$

by the integral test, since $\int_2^{\infty} \frac{dx}{(\log x)^{1+\varepsilon}} = \infty$ for any $\varepsilon > 0$.

Example 14 (Decay without Dimension Gap)**.** With $|P_k| := 2^k$ (so $H_k = k$):

- $H_k = k$

- $H_k/k = 1$

So $\limsup H_k/k = 1$, not < 1, yet $O(k) = 1/\log_2(k+1) \to 0$.

These examples prove that each implication in the hierarchy is strict.

7 Applications and Connections

7.1 Machine Learning Interpretation

In deep learning, P_k might represent the set of "effectively different" functions a network of depth k can represent. Our results suggest:

- Networks with $O(k) \to 0$ have limited expressive power relative to their size

- The entropy gap $H_k - k \to -\infty$ indicates strong implicit regularization

- The α-exponent characterizes the efficiency-complexity tradeoff

7.2 Information-Theoretic View

From an information theory perspective:

- H_k represents the bits needed to specify a pattern

- $O(k)$ measures information density relative to maximum entropy

- The critical threshold $|P_k| \sim 2^k \log_2(k+1)$ marks a phase transition in compressibility

7.3 Algorithmic Implications

For algorithm analysis:

- Subcritical growth suggests efficient pattern enumeration algorithms exist

- Critical growth indicates fundamental algorithmic barriers

- Supercritical growth implies inherent computational intractability

8 Technical Lemmas

We include the key technical tool for completeness.

Lemma 15 (Cauchy Condensation Rate Estimate)**.** *Let $\{a_k\}_{k=1}^{\infty}$ be a sequence of nonnegative real numbers that is eventually nonincreasing, and let $p > 1$. If $\sum_{k=1}^{\infty} a_k^p < \infty$, then $a_k = O\left(k^{-1/p}\right)$.*

Proof. Since $\{a_k\}$ is eventually nonincreasing and nonnegative, by the Cauchy condensation test, $\sum_{k=1}^{\infty} a_k^p < \infty$ if and only if $\sum_{n=0}^{\infty} 2^n a_{2^n}^p < \infty$.

Since the latter series converges, we have $2^n a_{2^n}^p \to 0$, which implies $a_{2^n} = o(2^{-n/p})$. In particular, there exists a constant $C > 0$ such that $a_{2^n} \leq C \cdot 2^{-n/p}$ for all sufficiently large n.

For arbitrary $k \geq 2$, choose n such that $2^n \leq k < 2^{n+1}$. Since $\{a_k\}$ is eventually nonincreasing:

$$a_k \leq a_{2^n} \leq C \cdot 2^{-n/p}.$$

Since $2^n \le k$, we have $n \le \log_2 k$, so $2^{-n/p} \ge 2^{-\log_2 k/p} = k^{-1/p}$. However, we need the opposite inequality. Since $k < 2^{n+1}$, we have $n > \log_2 k - 1$, so:

$$2^{-n/p} < 2^{-(\log_2 k - 1)/p} = 2^{1/p} \cdot k^{-1/p}.$$

Therefore, $a_k \le C \cdot 2^{1/p} \cdot k^{-1/p} = O(k^{-1/p})$. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

9   Open Questions

This work opens several avenues for future research:

1. **Process Characterization**: What conditions on pattern-generating processes lead to specific α-exponents? Can we classify natural pattern families by their exponents?

2. **Optimality**: Is the logarithmic factor in $O(k)$ optimal for all natural pattern families?

3. **Dynamics**: How does $O(k)$ evolve under pattern-preserving transformations?

4. **Multivariate**: Can this framework extend to patterns with multiple scaling parameters?

5. **Computational Complexity**: What is the algorithmic complexity of computing or approximating $O(k)$ for specific pattern families? Are there efficient algorithms for pattern families in each growth regime?

10   Conclusion

We have established a rigorous mathematical framework for understanding pattern density at exponential scale. The pattern measure $O(k)$ and its associated hierarchy provide both theoretical insight and practical tools for analyzing complex systems.

Our journey from a debugging session to a mathematical framework illustrates how practical challenges can lead to fundamental insights. The logarithmic dampening we observed empirically turns out to be a mathematical necessity, captured precisely by the equivalence $O(k) \to 0 \iff |P_k| = o(2^k \log k)$.

This work demonstrates that patterns, even at exponential scale, obey fundamental scaling laws. Understanding these laws helps us design better algorithms, build more efficient systems, and recognize the inherent limits of complexity growth.

Acknowledgments

References

[1] C.E. Shannon, *A mathematical theory of communication*, Bell System Technical Journal **27** (1948), 379–423.

[2] A.N. Kolmogorov, *Three approaches to the quantitative definition of information*, Problems of Information Transmission **1**(1) (1965), 1–7.

[3] T.M. Cover and J.A. Thomas, *Elements of information theory*, 2nd ed., Wiley-Interscience, 2006.

[4] V.N. Vapnik, *Statistical learning theory*, Wiley, New York, 1998.

[5] A.R. Barron, *Universal approximation bounds for superpositions of a sigmoidal function*, IEEE Transactions on Information Theory **39**(3) (1993), 930–945.

[6] G.H. Hardy, *Divergent series*, Oxford University Press, 1949.

[7] E.C. Titchmarsh, *The theory of functions*, 2nd ed., Oxford University Press, 1939.