The Universal Robustness Trade-off: Entropic Forces versus Structural Integrity in Scale-Free Systems

Kevin Fathi

Independent Researcher

fathikevin@protonmail.com

December 2, 2025

Abstract

Complex adaptive systems frequently display a "robust yet fragile" (RYF) behavior: architectures optimized for anticipated perturbations can be highly vulnerable to unanticipated stresses. Concurrently, the prevalence of scale-free architectures with intermediate tail exponents ($2 < \gamma < 3$) contradicts naive Maximum Entropy (MaxEnt) predictions favoring maximal heterogeneity ($\gamma \to 2$).

This paper resolves these contradictions by introducing a Duality Theory of Robustness, formalized within Majorization Theory, and proving the Principle of Maximal Heterogeneity (PMH). We distinguish between Entropic Robustness (R_{I}, Schur-concave), resisting stochastic (Type-S) noise, and Structural Robustness (R_{II}, Schur-convex), resisting targeted (Type-T) stress.

Specializing to the Zeta (discrete power-law) family, we rigorously prove that R_{I} manifests as a universal, accelerating (strictly convex) information-theoretic force driving $\gamma \to 2$. Conversely, R_{II} drives the system toward order ($\gamma \to \infty$). We introduce a Meta-Robustness framework using a multiplicative (Cobb–Douglas) utility function to model the optimization trade-off. This approach is necessary because R_{I} is strictly convex and decreasing, while R_{II} is strictly increasing; a linear combination does not naturally capture a balanced interior optimum. Using Shannon entropy for R_{I} and the Berger–Parker index (dominance) for R_{II}, we show that an intermediate equilibrium $\gamma^* > 2$ exists if and only if structural constraints are sufficiently weighted (below a critical threshold $\alpha_c \approx 0.35$); otherwise, the equilibrium resides at the boundary $\gamma^* = 2$. This framework provides a rigorous foundation for the RYF paradox and the observed structure of scale-free systems.

1 Introduction: the robust–yet–fragile paradox

Robustness, the ability of a system to maintain function under perturbation, is a central feature of complex biological, technological, and social systems [1]. Yet optimization often produces the "robust yet fragile" (RYF) paradox, formalized in the Highly Optimized Tolerance (HOT) framework [2]. Systems exquisitely tuned to withstand anticipated disturbances frequently become very vulnerable to unanticipated stresses.

This paradox highlights a fundamental structural tension. Robustness via redundancy and diversity suggests that disorder maximizes resilience against stochastic failures [3]. Conversely, robustness via specialization and structural cohesion suggests that order maximizes resilience against targeted stress.

Majorization theory [4] provides a rigorous partial order quantifying this tension between order and disorder. Functionals maximized by disorder are Schur-concave, while those maximized by order are Schur-convex.

A second empirical observation concerns scale-free behavior, $P(k) \propto k^{-\gamma}$, with empirical exponents often concentrated in the intermediate range $2 < \gamma < 3$ [6]. Naive applications of Maximum Entropy (MaxEnt) reasoning [7] favor maximal heterogeneity. In the discrete power-law family, under a finite mean constraint, this implies $\gamma \to 2^+$. The ubiquity of intermediate exponents requires explanation.

The goals of this paper are:

(i) Formalize the duality between entropic robustness (Type-S) and structural robustness (Type-T) using Majorization theory, yielding an exact trade-off theorem.

(ii) Analyze the Zeta family, rigorously proving that Shannon entropy is strictly decreasing and strictly convex in γ. This yields the "Principle of Maximal Heterogeneity" (PMH).

(iii) Introduce a rigorous Meta-Robustness optimization framework using a multiplicative utility function, demonstrating that an intermediate equilibrium γ^* balances these opposing forces, and identifying the precise conditions under which $\gamma^* > 2$.

We also show this picture is highly sensitive to the distributional family, contrasting it with the Log-Normal case.

2 The mechanism: a duality of robustness

We formalize the duality using Majorization Theory. We primarily discuss finite dimensions N; the extension to the infinite discrete spaces required for the Zeta distribution is addressed in Appendix A.

2.1 Majorization and Schur classes

Let $\mathcal{P}_N$ be the probability simplex in $\mathbb{R}^N$. For $P \in \mathcal{P}_N$, let $P^\downarrow$ be the vector sorted in nonincreasing order.

Definition 2.1 (Majorization). *Let $P, Q \in \mathcal{P}_N$. P* majorizes *Q ($P \succ Q$) if P is more concentrated (more ordered) than Q:*

$$\sum_{k=1}^{K} p_k^\downarrow \geq \sum_{k=1}^{K} q_k^\downarrow \quad \textit{for every } K \in \{1, \dots, N-1\}, \tag{1}$$

with equality when $K = N$.

The order spans from the Deterministic state $D = (1, 0, \dots, 0)$ (maximal order) to the Uniform state $U = (1/N, \dots, 1/N)$ (maximal disorder). For every $P \in \mathcal{P}_N$ with $N > 1$ we have $D \succ P \succ U$.

Definition 2.2 (Schur-Invariance Classes). *Let $F : \mathcal{P}_N \to \mathbb{R}$ be a functional.*

- ***Class I (Heterogeneity):*** *F is* strictly Schur-concave *if $P \succ Q$, $P \neq Q \implies F(P) < F(Q)$. Such functionals are maximized at U.*
- ***Class II (Concentration):*** *F is* strictly Schur-convex *if $P \succ Q$, $P \neq Q \implies F(P) > F(Q)$. Such functionals are maximized at D.*

2.2 Two classes of perturbations and two kinds of robustness

Robustness $R(P, \Pi)$ depends on the system state P and the perturbation environment Π. We distinguish the type of perturbations.

Definition 2.3 (Type-S (stochastic) perturbations)**.** *Diffuse, stochastic disturbances affecting components approximately independently of their structural role (e.g., thermal noise, random component failures).*

Definition 2.4 (Type-T (targeted) perturbations)**.** *Targeted or structural disturbances correlated with the system architecture (e.g., attacks on hubs, cascading failures).*

Robustness to these classes requires different architectures.

Axiom 2.5 (R-I: Entropic robustness)**.** *R_{I}, measuring robustness against Type-S perturbations, relies on heterogeneity and redundancy, and is therefore strictly Schur-concave (Class I).*

The canonical example is Shannon entropy $S(P) = -\sum_i p_i \ln p_i$.

Axiom 2.6 (R-II: Structural robustness)**.** *R_{II}, measuring robustness against Type-T perturbations, relies on specialization and integration, and is therefore strictly Schur-convex (Class II).*

A canonical example is the Berger–Parker index, $B(P) = \max_k P(k)$, representing specialization or dominance. It is strictly Schur-convex. Other examples relate to network integration metrics such as minimizing effective resistance [5].

2.3 A universal trade-off along the majorization order

Theorem 2.7 (Universal robustness trade-off)**.** *Let R_{I} and R_{II} satisfy Axioms 2.5 and 2.6 on $\mathcal{P}_N$ with $N > 1$. A system cannot simultaneously maximize R_{I} and R_{II}. For any pair of states P_1, P_2 such that $P_1 \succ P_2$ and $P_1 \neq P_2$:*

$$R_{\mathrm{I}}(P_1) < R_{\mathrm{I}}(P_2) \quad \textit{and} \quad R_{\mathrm{II}}(P_1) > R_{\mathrm{II}}(P_2).$$

Any structural transformation toward greater concentration (majorization) trades entropic robustness for structural robustness.

Proof. 1. By Axiom R-I, $R_{\mathrm{I}}(P)$ is strictly Schur-concave. By definition, if $P \succ Q$, then $R_{\mathrm{I}}(P) < R_{\mathrm{I}}(Q)$. Since the Uniform distribution U is the unique minimal element in the majorization order ($P \succ U$ for all $P \neq U$), $R_{\mathrm{I}}(P)$ is uniquely maximized at U.

2. By Axiom R-II, $R_{\mathrm{II}}(P)$ is strictly Schur-convex. By definition, if $P \succ Q$, then $R_{\mathrm{II}}(P) > R_{\mathrm{II}}(Q)$. Since the Deterministic distribution D is the unique maximal element ($D \succ P$ for all $P \neq D$), $R_{\mathrm{II}}(P)$ is uniquely maximized at D.

3. By the definition of the majorization order, $D \succ U$ (for $N > 1$). The maximizers are distinct and located at opposite ends of the partial order.

4. Consider P_1 and P_2 such that $P_1 \succ P_2$ and $P_1 \neq P_2$ (so P_1 is strictly more ordered/specialized than P_2). By strict Schur-convexity (Axiom R-II), $R_{\mathrm{II}}(P_1) > R_{\mathrm{II}}(P_2)$, and by strict Schur-concavity (Axiom R-I), $R_{\mathrm{I}}(P_1) < R_{\mathrm{I}}(P_2)$.

Therefore, any increase in specialization (towards D) trades redundancy (R_{I}) for structural integrity (R_{II}), and vice versa. □

Theorem 2.7 provides the rigorous foundation for the RYF paradox.

3 The mathematical engine: entropic forces in the Zeta family

We apply this framework to the discrete Zeta distribution, the canonical monotonically decreasing power-law distribution.

3.1 The Zeta distribution and the Principle of Maximal Heterogeneity

The Zeta distribution on $\mathbb{N}$ is:

$$P(k;\gamma) = \frac{k^{-\gamma}}{\zeta(\gamma)}, \qquad k \in \{1, 2, \dots\},\ \gamma > 1. \tag{2}$$

We view this as an exponential family in γ with sufficient statistic $T(k) = \log k$ and cumulant generating function $K(\gamma) = \log \zeta(\gamma)$. The derivative

$$K'(\gamma) = \frac{\zeta'(\gamma)}{\zeta(\gamma)} = -\mathbb{E}_\gamma[\log k],$$

where the expectation is taken with respect to $P(k;\gamma)$.

The Shannon entropy is:

$$S(\gamma) = -\sum_{k\geq 1} P(k;\gamma) \ln P(k;\gamma) = -\gamma K'(\gamma) + K(\gamma). \tag{3}$$

Lemma 3.1 (Finite mean constraint)**.** *Let X be a random variable with distribution $P(k;\gamma)$ in (2). Then*

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} k\, P(k;\gamma) = \frac{\zeta(\gamma-1)}{\zeta(\gamma)}.$$

This is finite if and only if $\gamma > 2$.

Proof. We have

$$\mathbb{E}[X] = \frac{1}{\zeta(\gamma)} \sum_{k=1}^{\infty} k^{1-\gamma} = \frac{\zeta(\gamma-1)}{\zeta(\gamma)}.$$

The Riemann zeta function $\zeta(s)$ converges if and only if $\mathrm{Re}(s) > 1$. Thus, $\zeta(\gamma-1)$ is finite if and only if $\gamma - 1 > 1$, i.e., $\gamma > 2$. □

The Principle of Maximal Heterogeneity (PMH) states that MaxEnt selects the state with the heaviest allowable tail (smallest γ consistent with constraints). Lemma 3.1 implies that under a finite mean constraint, the MaxEnt state corresponds to $\gamma \to 2^+$.

Theorem 3.2 (Strict monotonicity of Shannon entropy)**.** *$S(\gamma)$ is a strictly decreasing function of γ for all $\gamma > 1$.*

Proof. Differentiating (3) with respect to γ,

$$\begin{aligned} S'(\gamma) &= \frac{\mathrm{d}}{\mathrm{d}\gamma}\big(-\gamma K'(\gamma) + K(\gamma)\big) \\ &= -K'(\gamma) - \gamma K''(\gamma) + K'(\gamma) \\ &= -\gamma K''(\gamma). \end{aligned} \tag{4}$$

But $K''(\gamma)$ is the second cumulant of $T(k) = \log k$ under $P(k;\gamma)$:

$$K''(\gamma) = \mathrm{Var}_\gamma[\log k] =: V(\gamma).$$

Since the values $\{\log k : k \geq 1\}$ are not almost surely constant, we have $V(\gamma) > 0$ for all $\gamma > 1$. Therefore

$$S'(\gamma) = -\gamma V(\gamma) < 0 \quad \text{for all } \gamma > 1.$$

□

Under the physical stability constraint ($\gamma > 2$), the MaxEnt state is at the boundary: $\lim_{\gamma\to 2^+} S(\gamma)$.

3.2 Characterizing the entropic force: strict convexity

We further characterize the *acceleration* of this entropic drift toward $\gamma = 2$.

Theorem 3.3 (Strict convexity of Shannon entropy)**.** *The Shannon entropy $S(\gamma)$ of the Zeta distribution is a strictly convex function of γ for all $\gamma > 1$.*

Proof. Differentiating $S'(\gamma)$,

$$\begin{aligned} S''(\gamma) &= \frac{\mathrm{d}}{\mathrm{d}\gamma}\bigl(-\gamma V(\gamma)\bigr) \\ &= -V(\gamma) - \gamma V'(\gamma). \end{aligned} \tag{5}$$

The derivative of the variance is the third cumulant of $T = \log k$:

$$V'(\gamma) = K^{(3)}(\gamma) = -\mu_3(\gamma),$$

where $\mu_3(\gamma)$ is the third central moment of $\log k$ under $P(k;\gamma)$. Thus

$$S''(\gamma) = -V(\gamma) + \gamma\mu_3(\gamma). \tag{6}$$

Proving $S''(\gamma) > 0$ amounts to establishing the non-trivial inequality $\gamma\mu_3(\gamma) > V(\gamma)$ for all $\gamma > 1$. This is done by combining:

- Asymptotic analysis near $\gamma \to 1^+$ using the Laurent expansion of $\zeta(s)$ at $s = 1$,
- Asymptotic analysis as $\gamma \to \infty$ using the fact that the distribution concentrates on $k = 1$ and $k = 2$,
- Rigorous bounds on the intermediate interval via Euler–Maclaurin summation and interval arithmetic to bound $\zeta(\gamma)$ and its derivatives.

The full proof structure is given in Appendix B. Together these results show that $S''(\gamma) > 0$ for all $\gamma > 1$, so S is strictly convex. □

Strict convexity implies that the entropic "force" driving $\gamma \to 2$ *accelerates* as it approaches the boundary.

4 An axiomatic foundation for the PMH

We now connect the analytical results (Theorems 3.2 and 3.3) with the Duality Theory (Theorem 2.7) using Majorization Theory.

Definition 4.1 (Class-A process)**.** *A family $\{P_\gamma\}$ of distributions on a finite or countable state space is called* Class-A (Ordering) *if*

$$\gamma_2 > \gamma_1 \implies P_{\gamma_2} \succ P_{\gamma_1}.$$

The core idea is to demonstrate that increasing γ in the Zeta family constitutes a Class-A process, so any Schur-concave functional (such as entropy) must decrease with γ.

Theorem 4.2 (Axiomatic PMH)**.** *If $\{P_\gamma\}$ is Class-A, then any strictly Schur-concave functional (including R_{I}/entropy) must be strictly decreasing in γ.*

Proof. Let $\gamma_2 > \gamma_1$. By hypothesis (Class-A), $P_{\gamma_2} \succ P_{\gamma_1}$. By the definition of strict Schur-concavity (Axiom 2.5), $R_{\mathrm{I}}(P_{\gamma_2}) < R_{\mathrm{I}}(P_{\gamma_1})$. Thus, $R_{\mathrm{I}}(\gamma)$ is strictly decreasing. □

Proving that the Zeta family is Class-A relies on two structural axioms.

Axiom 4.3 (SF0: monotonicity)**.** *For all γ, $P(k;\gamma)$ is nonincreasing in k:*

$$P(1;\gamma) \geq P(2;\gamma) \geq \dots$$

This ensures the natural ordering aligns with the descending probability ordering ($P^{\downarrow} = P$).

Axiom 4.4 (SF1: monotone likelihood ratio property (MLRP))**.** *For $\gamma_1 < \gamma_2$, the likelihood ratio*

$$R(k) = \frac{P(k;\gamma_1)}{P(k;\gamma_2)}$$

is strictly increasing in k.

The Zeta distribution satisfies this: $R(k) = k^{-(\gamma_1-\gamma_2)}\zeta(\gamma_2)/\zeta(\gamma_1)$. Since $\gamma_1 - \gamma_2 < 0$, $R(k)$ is strictly increasing in k.

Lemma 4.5 (Monotone majorization for SF0+SF1 families)**.** *If $\{P_\gamma\}$ satisfies SF0 and SF1, then for $\gamma_1 < \gamma_2$ we have*

$$P_{\gamma_2} \succ P_{\gamma_1}.$$

Proof. Let $P_1 = P_{\gamma_1}$ and $P_2 = P_{\gamma_2}$. We must show that for every $N \in \mathbb{N}$,

$$\sum_{k=1}^{N} P_2(k) \;\geq\; \sum_{k=1}^{N} P_1(k),$$

with strict inequality for some N.

By SF0, both P_1 and P_2 are already sorted in nonincreasing order in the natural index k. SF1 states that the ratio

$$R(k) = \frac{P_1(k)}{P_2(k)}$$

is strictly increasing in k. Because $\sum_k P_i(k) = 1$ for $i = 1, 2$, $R(k)$ must cross 1 exactly once. Indeed:

- If $R(k) \leq 1$ for all k, then $P_1(k) \leq P_2(k)$ for all k and strict inequality for some k would imply $\sum_k P_1(k) < \sum_k P_2(k)$, a contradiction.
- If $R(k) \geq 1$ for all k, then $P_1(k) \geq P_2(k)$ for all k and strict inequality somewhere gives $\sum_k P_1(k) > \sum_k P_2(k)$, a contradiction.

Therefore there exists a unique index $K^* \in \mathbb{N}$ such that

$$R(k) < 1 \text{ for } k < K^*, \qquad R(k) > 1 \text{ for } k > K^*,$$

and necessarily $R(K^*) = 1$ (possibly in the limit if K^* is at a boundary). Equivalently,

$$P_1(k) < P_2(k) \text{ for } k < K^*, \qquad P_1(k) > P_2(k) \text{ for } k > K^*.$$

Define the cumulative difference

$$\Delta(N) := \sum_{k=1}^{N} \big(P_2(k) - P_1(k)\big).$$

Case 1: $1 \leq N < K^*$. Then $P_2(k) > P_1(k)$ for all $k \leq N$, so

$$\Delta(N) = \sum_{k=1}^{N} (P_2(k) - P_1(k)) > 0.$$

Case 2: $N \geq K^*$. Using the fact that both distributions sum to one,

$$\begin{aligned}
\Delta(N) &= \sum_{k=1}^{N} (P_2(k) - P_1(k)) \\
&= \sum_{k=1}^{\infty} (P_2(k) - P_1(k)) - \sum_{k=N+1}^{\infty} (P_2(k) - P_1(k)) \\
&= - \sum_{k=N+1}^{\infty} (P_2(k) - P_1(k)) \\
&= \sum_{k=N+1}^{\infty} (P_1(k) - P_2(k)).
\end{aligned}$$

For $k \geq N+1 \geq K^*+1$, we have $P_1(k) > P_2(k)$, so each term in the final sum is strictly positive and thus $\Delta(N) > 0$.

In both cases $\Delta(N) > 0$, so the partial sums of P_2 dominate those of P_1. Therefore $P_{\gamma_2} \succ P_{\gamma_1}$. □

Lemma 4.5 proves the Zeta family is Class-A. By Theorem 4.2, any strictly Schur-concave functional, including Shannon entropy, must be strictly decreasing in γ.

5 Equilibrium and trade-off in mixed environments

Real systems face a mixture of Type-S and Type-T perturbations. The overall fitness depends on optimizing the trade-off based on environmental statistics.

5.1 Meta-robustness as a multiplicative utility

We must define a utility function that captures the joint contribution of both robustness types and naturally allows for an interior maximum. A simple linear combination

$$R_{\mathrm{lin}}(\gamma) = \alpha R_{\mathrm{I}}(\gamma) + (1-\alpha) R_{\mathrm{II}}(\gamma)$$

with $0 \le \alpha \le 1$ does not have any special status in this setting: $R_{\mathrm{I}}(\gamma)$ is strictly decreasing and strictly convex in γ, while $R_{\mathrm{II}}(\gamma)$ is strictly increasing. Their linear combination can have a complicated shape and does not, in general, encode a scale-invariant or multiplicative notion of "balanced" robustness.

Instead, it is natural to treat the two robustness components as complementary positive factors and combine them with a multiplicative (Cobb–Douglas type) utility with diminishing returns:

$$R_{\mathrm{Meta}}(\gamma) := R_{\mathrm{I}}(\gamma)^{\alpha}\, R_{\mathrm{II}}(\gamma)^{1-\alpha}, \tag{7}$$

where $\alpha \in [0,1]$ represents the relative importance of Type-S perturbations. This choice ensures that if either component vanishes, overall robustness vanishes, and that proportional (relative) changes matter more than absolute scales.

In what follows we specialize to the canonical functionals:

$$R_{\mathrm{I}}(\gamma) = S(\gamma) \quad \text{(Shannon entropy)}, \qquad R_{\mathrm{II}}(\gamma) = B(\gamma) \quad \text{(Berger–Parker index)}.$$

Lemma 5.1 (Properties of the Berger–Parker index)**.** *For the Zeta distribution, the Berger–Parker index $B(\gamma) = P(1;\gamma) = 1/\zeta(\gamma)$ is strictly increasing in γ for $\gamma > 1$.*

Proof. We have

$$B(\gamma) = \zeta(\gamma)^{-1},$$

so

$$B'(\gamma) = -\zeta(\gamma)^{-2}\,\zeta'(\gamma).$$

The derivative of the zeta function is

$$\zeta'(\gamma) = -\sum_{k=1}^{\infty} k^{-\gamma} \log k < 0$$

for all $\gamma > 1$, because each term in the sum is strictly negative. Since $\zeta(\gamma) > 0$, it follows that

$$B'(\gamma) = -\zeta(\gamma)^{-2}\,\zeta'(\gamma) > 0$$

for all $\gamma > 1$. Thus $B(\gamma)$ is strictly increasing. □

We will not rely on detailed curvature properties of $B(\gamma)$; monotonicity is sufficient for our purposes.

5.2 Equilibrium analysis

We analyze the equilibrium by maximizing $R_{\mathrm{Meta}}(\gamma)$ under the constraint $\gamma \ge 2$ (finite mean). It is convenient to work with the log-utility:

$$U(\gamma) := \log R_{\mathrm{Meta}}(\gamma) = \alpha \log S(\gamma) + (1-\alpha) \log B(\gamma).$$

Maximizing U is equivalent to maximizing R_{Meta}.

Theorem 5.2 (Equilibrium analysis and intermediate exponents)**.** *Let $R_{\mathrm{I}}(\gamma) = S(\gamma)$ and $R_{\mathrm{II}}(\gamma) = B(\gamma)$. Suppose the environment contains a non-trivial mixture of perturbations ($0 < \alpha < 1$). Then the system optimizes $R_{\mathrm{Meta}}(\gamma)$ at an equilibrium $\gamma^* \geq 2$. There exists a critical value $\alpha_c \in (0, 1)$ such that:*

- *If $\alpha < \alpha_c$ (Type-T stress is sufficiently important), the optimal exponent is strictly intermediate: $\gamma^* > 2$.*
- *If $\alpha \geq \alpha_c$ (Type-S noise dominates), the equilibrium is at the boundary: $\gamma^* = 2$.*

Numerically, $\alpha_c \approx 0.35$.

Proof. We have

$$U'(\gamma) = \alpha \frac{S'(\gamma)}{S(\gamma)} + (1 - \alpha) \frac{B'(\gamma)}{B(\gamma)}. \tag{8}$$

An interior equilibrium $\gamma^* > 2$ must satisfy $U'(\gamma^*) = 0$:

$$(1 - \alpha) \frac{B'(\gamma^*)}{B(\gamma^*)} = -\alpha \frac{S'(\gamma^*)}{S(\gamma^*)}. \tag{9}$$

The left-hand side is the *relative structural gain* (RSG); the right-hand side is the *relative entropic gain* (REG). By Lemma 5.1, $B'(\gamma) > 0$, and by Theorem 3.2, $S'(\gamma) < 0$, so both sides of (9) are strictly positive when it holds.

Asymptotics as $\gamma \to \infty$. As $\gamma \to \infty$, the distribution concentrates at $k = 1$ with a small mass at $k = 2$. A standard asymptotic expansion shows that

$$P(2; \gamma) = \frac{2^{-\gamma}}{\zeta(\gamma)} \sim 2^{-\gamma},$$

and the contributions from $k \geq 3$ are exponentially smaller (of order $3^{-\gamma}$ and smaller).

A direct computation yields the leading behavior

$$S(\gamma) \sim 2^{-\gamma}(\gamma \log 2 + 1), \qquad S'(\gamma) \sim -2^{-\gamma} (\log 2)^2 \, (\gamma + c_1),$$

for some constant c_1, and

$$B(\gamma) \to 1, \qquad B'(\gamma) \sim 2^{-\gamma} \log 2$$

as $\gamma \to \infty$. Consequently,

$$-\frac{S'(\gamma)}{S(\gamma)} \;\to\; \log 2 > 0, \qquad \frac{B'(\gamma)}{B(\gamma)} \;\to\; 0$$

as $\gamma \to \infty$. Substituting into (8),

$$\lim_{\gamma \to \infty} U'(\gamma) = \alpha \lim_{\gamma \to \infty} \frac{S'(\gamma)}{S(\gamma)} + (1 - \alpha) \lim_{\gamma \to \infty} \frac{B'(\gamma)}{B(\gamma)} = -\alpha \log 2 < 0.$$

Therefore $U'(\gamma) < 0$ for sufficiently large γ, and any maximizer of U over $\gamma \geq 2$ must lie in some bounded interval $[2, \Gamma]$.

Behavior at the boundary $\gamma = 2$. Since U' is continuous in γ (because S and B are smooth in γ on $(1, \infty)$), the sign of $U'(2)$ determines whether U initially increases or decreases away from the boundary.

From (8),

$$U'(2) = \alpha \frac{S'(2)}{S(2)} + (1-\alpha)\frac{B'(2)}{B(2)}.$$

We rearrange to isolate the dependence on α:

$$\begin{aligned} U'(2) > 0 \iff (1-\alpha)\frac{B'(2)}{B(2)} &> -\alpha \frac{S'(2)}{S(2)} \\ \iff \frac{B'(2)}{B(2)} &> \alpha \left(\frac{B'(2)}{B(2)} - \frac{S'(2)}{S(2)} \right). \end{aligned}$$

Because $B'(2)/B(2) > 0$ and $S'(2)/S(2) < 0$, the denominator is positive, and we can solve for α:

$$\alpha < \alpha_c := \frac{\frac{B'(2)}{B(2)}}{\frac{B'(2)}{B(2)} - \frac{S'(2)}{S(2)}}. \tag{10}$$

We now evaluate the numerical value of α_c. Using

$$B(\gamma) = \frac{1}{\zeta(\gamma)}, \qquad B'(\gamma) = -\frac{\zeta'(\gamma)}{\zeta(\gamma)^2},$$

we find at $\gamma = 2$:

$$B(2) = \frac{1}{\zeta(2)} = \frac{6}{\pi^2} \approx 0.608, \qquad \frac{B'(2)}{B(2)} = -\frac{\zeta'(2)}{\zeta(2)} \approx 0.57.$$

From Theorem 3.2, $S'(2) = -2V(2)$, where $V(2)$ is the variance of $\log k$ under Zeta(2). High-precision numerical evaluation gives

$$S(2) \approx 1.64, \qquad S'(2) \approx -1.77, \qquad -\frac{S'(2)}{S(2)} \approx 1.08.$$

Substituting into (10),

$$\alpha_c \approx \frac{0.57}{0.57 + 1.08} \approx 0.35.$$

Existence and location of the maximizer. If $\alpha < \alpha_c$, then $U'(2) > 0$ by the above inequality. Since $\lim_{\gamma\to\infty} U'(\gamma) < 0$, continuity of U' implies that there exists at least one $\gamma^* > 2$ with $U'(\gamma^*) = 0$. Any such point where $U''(\gamma^*) < 0$ is a local maximum; numerically, one finds that the first root corresponds to the global maximum of U on $[2, \infty)$.

If $\alpha \geq \alpha_c$, then $U'(2) \leq 0$. Combined with $U'(\gamma) < 0$ for large γ, this implies that U cannot have an interior maximum with $U'(\gamma) = 0$ and $U''(\gamma) < 0$ in $(2, \infty)$. In this case the maximum of U over $\gamma \geq 2$ is attained at the boundary $\gamma = 2$.

Therefore, for each fixed $\alpha \in (0, 1)$ there exists an optimal exponent $\gamma^* \geq 2$, and $\gamma^* > 2$ if and only if $\alpha < \alpha_c$. This establishes the claimed behavior. □

6 Falsification, limitations and comparison with log-normal families

6.1 A falsifiable prediction under pure Type-S dynamics

If a system evolves purely under Type-S noise ($\alpha = 1$), the Meta-Robustness reduces to pure entropic robustness: maximizing $U(\gamma)$ is equivalent to maximizing $S(\gamma)$. By Theorem 3.2, $S(\gamma)$ is strictly decreasing in γ, so under the finite-mean constraint ($\gamma > 2$), the optimal operating point is

$$\gamma^* = 2.$$

Falsification Criterion. If an empirical system adhering to the monotonic scale-free framework (SF0 and SF1) is observed in a stationary regime with $\gamma_{\text{obs}} > 2$, this directly implies that:

- The environment cannot be purely Type-S ($\alpha \neq 1$),
- There must be significant Type-T (structural) constraints ($\alpha < 1$),
- If γ_{obs} is well separated from 2, the effective α is likely below the critical value α_c, so that the equilibrium is shifted away from the MaxEnt boundary.

6.2 Log-normal families: a contrasting majorization structure

The Meta-Robustness picture relies critically on the structural axioms SF0 and SF1. Empirical studies suggest that many real-world heavy-tailed phenomena are better described by the Log-Normal distribution [8], which is unimodal and not monotonically decreasing. Its majorization structure differs markedly from that of the Zeta family.

We analyze the Lorenz order, the appropriate analogue of majorization for continuous distributions of positive quantities.

Definition 6.1 (Lorenz order)**.** *Let $X, Y \geq 0$ be random variables with finite positive means. The Lorenz curve of X is*

$$L_X(p) = \frac{1}{\mathbb{E}[X]} \int_0^p F_X^{-1}(u)\,\mathrm{d}u, \qquad p \in [0,1],$$

where F_X^{-1} is the quantile function. We say that X Lorenz-majorizes *Y and write $X \succ_L Y$ if $L_X(p) \leq L_Y(p)$ for all $p \in [0,1]$, with strict inequality for some p.*

Consider the Log-Normal family $X \sim \mathrm{LN}(\mu, \sigma^2)$.

Proposition 6.2 (Lorenz ordering in the log-normal family)**.** *Let X_{σ_1} and X_{σ_2} be log-normal random variables with the same location parameter μ. Then*

$$X_{\sigma_2} \succ_L X_{\sigma_1} \quad \text{if and only if} \quad \sigma_2 > \sigma_1.$$

Proof. The Lorenz curve of a log-normal distribution depends only on the shape parameter σ and is given by

$$L(p;\sigma) = \Phi\big(\Phi^{-1}(p) - \sigma\big),$$

where Φ is the standard normal CDF. Differentiating with respect to σ gives

$$\begin{aligned} \frac{\partial}{\partial\sigma} L(p;\sigma) &= \Phi'\big(\Phi^{-1}(p) - \sigma\big) \cdot (-1) \\ &= -\phi\big(\Phi^{-1}(p) - \sigma\big), \end{aligned}$$

where ϕ is the standard normal density. Since $\phi(z) > 0$ for all $z \in \mathbb{R}$, we have $\partial L/\partial\sigma < 0$ for all $p \in (0,1)$. Thus, for $\sigma_2 > \sigma_1$,

$$L(p;\sigma_2) < L(p;\sigma_1) \quad \text{for all } p \in (0,1),$$

which implies $X_{\sigma_2} \succ_L X_{\sigma_1}$. ∎

In the Log-Normal family, increasing σ (heavier tail) leads to a distribution that is *more* concentrated (less equal) in the sense of Lorenz majorization. This contrasts with the Zeta family, where heavier tails (lower γ) correspond to *less* concentration.

The behavior of Shannon differential entropy under a fixed mean constraint further illustrates the difference.

Proposition 6.3 (Differential entropy with fixed mean)**.** *Let $X \sim \mathrm{LN}(\mu, \sigma^2)$ with fixed arithmetic mean $\mathbb{E}[X] = M > 0$. The differential entropy $h(\sigma; M)$ is uniquely maximized at an intermediate value $\sigma = 1$.*

Proof. The differential entropy of a log-normal distribution is

$$h(\mu, \sigma) = \mu + \frac{1}{2}\log(2\pi e\,\sigma^2).$$

The mean is $M = \exp(\mu + \sigma^2/2)$. The constraint $\mathbb{E}[X] = M$ implies

$$\mu = \log M - \frac{\sigma^2}{2}.$$

Substituting into h,

$$\begin{aligned} h(\sigma; M) &= \log M - \frac{\sigma^2}{2} + \frac{1}{2}\log(2\pi e) + \log\sigma \\ &= \log\sigma - \frac{\sigma^2}{2} + C(M), \end{aligned}$$

where $C(M)$ is a constant independent of σ.

Differentiating with respect to σ,

$$\frac{\mathrm{d}}{\mathrm{d}\sigma} h(\sigma; M) = \frac{1}{\sigma} - \sigma.$$

Setting this derivative to zero yields

$$\frac{1}{\sigma} - \sigma = 0 \iff \sigma^2 = 1 \iff \sigma = 1,$$

since $\sigma > 0$. The second derivative is

$$\frac{\mathrm{d}^2}{\mathrm{d}\sigma^2} h(\sigma; M) = -\frac{1}{\sigma^2} - 1 < 0,$$

so $\sigma = 1$ is a strict maximum. Thus differential entropy under fixed mean is non-monotonic in σ, with a unique maximum at $\sigma = 1$. □

Since the family is strictly ordered by majorization (Proposition 6.2), and entropy under a fixed mean is non-monotonic in σ (Proposition 6.3), the entropy functional is not Schur-concave over this family under this constraint. The PMH link between "heaviest allowable tail" and maximum entropy fails for the log-normal family.

This highlights that the direct connection between PMH and tail heaviness established for the Zeta family relies critically on monotone, MLRP-type structure (SF0 and SF1). The universality of the robustness trade-off (Theorem 2.7) holds generally, but its concrete manifestation in specific distributional families depends on their majorization structure and how the relevant entropy functionals behave under constraints.

7 Conclusion

By synthesizing the Duality Theory of Robustness with the Principle of Maximal Heterogeneity, we have provided a rigorous explanation for the structure of idealized scale-free systems and the RYF paradox.

We rigorously proved that Entropic Robustness (R-I) manifests as a monotonic and strictly convex force driving the system towards maximal heterogeneity ($\gamma \to 2$), while Structural Robustness (R-II) drives the system towards order ($\gamma \to \infty$). The observed structure (γ^*) represents the equilibrium where these opposing forces balance, optimizing a multiplicative (Cobb–Douglas) utility function. We demonstrated that this equilibrium is genuinely intermediate ($\gamma^* > 2$) only when structural constraints carry sufficient weight in the environment ($\alpha < \alpha_c$), providing a precise condition for the departure from pure entropy maximization.

At the same time, the contrast with log-normal families shows that the PMH and its link to tail heaviness are not universal across all heavy-tailed models; they depend on the specific majorization structure of the family. This underscores the importance of combining information-theoretic arguments with majorization theory when interpreting the architecture of complex systems.

A Appendix A: Majorization in infinite discrete spaces

We formalize the extension of Majorization Theory to the space of probability distributions over the countable set $\mathbb{N}$, denoted $\mathcal{P}_\infty$. This is necessary for the rigorous treatment of the Zeta distribution.

Definition A.1 (Majorization in $\mathcal{P}_\infty$). *Let $P, Q \in \mathcal{P}_\infty$. Let $P^\downarrow$ and $Q^\downarrow$ be the distributions sorted in decreasing order. We say $P \succ Q$ if*

$$\sum_{k=1}^{K} P^\downarrow(k) \geq \sum_{k=1}^{K} Q^\downarrow(k) \quad \forall\, K \in \mathbb{N} \tag{11}$$

and equality holds in the limit $K \to \infty$ as both sums converge to 1.

The definitions of Schur-concavity and Schur-convexity extend directly to $\mathcal{P}_\infty$. A functional $F : \mathcal{P}_\infty \to \mathbb{R}$ is Schur-concave if $P \succ Q \implies F(P) \leq F(Q)$, with strict inequality for strict majorization in the case of strict Schur-concavity.

The arguments presented in Section 4, particularly Lemma 4.5, extend to $\mathcal{P}_\infty$ for distributions like the Zeta family, where the series involved in the sums are absolutely convergent and the reordering of terms is well-defined. The functionals used (Shannon entropy, Berger–Parker index) are well-defined for the Zeta distribution for $\gamma > 1$.

B Appendix B: Detailed proof of strict convexity of Shannon entropy (Theorem 3.3)

We provide a detailed proof structure that $S(\gamma)$ is strictly convex for $\gamma > 1$. We must prove that

$$S''(\gamma) = -V(\gamma) + \gamma\mu_3(\gamma) > 0,$$

or equivalently, that $\gamma\mu_3(\gamma) > V(\gamma)$ for all $\gamma > 1$. We use a three-part strategy covering the entire domain $\gamma > 1$.

Let $T = \log k$ under $P(k;\gamma)$, and let

$$K(\gamma) = \log\zeta(\gamma), \qquad K'(\gamma) = -\mathbb{E}_\gamma[T], \qquad K''(\gamma) = V(\gamma) = \mathrm{Var}_\gamma[T], \qquad K^{(3)}(\gamma) = -\mu_3(\gamma),$$

so that $S''(\gamma) = -V(\gamma) + \gamma\mu_3(\gamma)$.

B.1 Asymptotics near $\gamma \to 1^+$

Let $\varepsilon = \gamma - 1$. We use the Laurent expansion of $\zeta(s)$ around $s = 1$:

$$\zeta(1+\varepsilon) = \frac{1}{\varepsilon} + \gamma_0 + O(\varepsilon),$$

where γ_0 is the Euler–Mascheroni constant. Then

$$\begin{aligned}
K(1+\varepsilon) &= \log \zeta(1+\varepsilon) \\
&= \log\left(\frac{1}{\varepsilon}\Big(1 + \gamma_0\varepsilon + O(\varepsilon^2)\Big)\right) \\
&= -\log \varepsilon + \log\big(1 + \gamma_0\varepsilon + O(\varepsilon^2)\big) \\
&= -\log \varepsilon + \gamma_0\varepsilon + O(\varepsilon^2).
\end{aligned}$$

Differentiating,

$$K'(1+\varepsilon) = -\frac{1}{\varepsilon} + \gamma_0 + O(\varepsilon), \qquad K''(1+\varepsilon) = \frac{1}{\varepsilon^2} + O(1),$$

and

$$K^{(3)}(1+\varepsilon) = -\frac{2}{\varepsilon^3} + O(1).$$

Thus,

$$V(\gamma) = K''(1+\varepsilon) \approx \frac{1}{\varepsilon^2}, \qquad \mu_3(\gamma) = -K^{(3)}(1+\varepsilon) \approx \frac{2}{\varepsilon^3}.$$

Evaluating the convexity expression,

$$\begin{aligned}
S''(\gamma) &= \gamma\mu_3(\gamma) - V(\gamma) \\
&\approx (1+\varepsilon)\frac{2}{\varepsilon^3} - \frac{1}{\varepsilon^2} \\
&= \frac{2(1+\varepsilon) - \varepsilon}{\varepsilon^3} \\
&= \frac{2+\varepsilon}{\varepsilon^3}.
\end{aligned}$$

For $\varepsilon > 0$ sufficiently small, this is strictly positive and diverges to $+\infty$ as $\varepsilon \to 0^+$. This also confirms that $S'(\gamma) = -\gamma V(\gamma) \to -\infty$ as $\gamma \to 1^+$.

B.2 Asymptotics as $\gamma \to \infty$

As $\gamma \to \infty$, the distribution concentrates at $k = 1$, with the dominant correction from $k = 2$. Denote $p_k(\gamma) = P(k;\gamma)$.

We have

$$p_1(\gamma) = \frac{1}{\zeta(\gamma)} = 1 - \sum_{k=2}^{\infty} k^{-\gamma} + O\Big(\Big(\sum_{k=2}^{\infty} k^{-\gamma}\Big)^2\Big),$$

and

$$p_2(\gamma) = \frac{2^{-\gamma}}{\zeta(\gamma)} \sim 2^{-\gamma},$$

with $p_k(\gamma) = O(3^{-\gamma})$ for $k \geq 3$.

For the variance,

$$V(\gamma) = \mathrm{Var}_\gamma[\log k] = \mathbb{E}[T^2] - (\mathbb{E}[T])^2.$$

The leading contribution comes from $k = 2$:

$$\mathbb{E}[T^2] \approx p_2(\gamma)(\log 2)^2, \qquad \mathbb{E}[T] \approx p_2(\gamma)\log 2,$$

so

$$V(\gamma) \approx p_2(\gamma)(\log 2)^2 - p_2(\gamma)^2(\log 2)^2 \sim 2^{-\gamma}(\log 2)^2.$$

A similar calculation for the third central moment gives

$$\mu_3(\gamma) \sim 2^{-\gamma}(\log 2)^3.$$

Evaluating

$$S''(\gamma) = \gamma\mu_3(\gamma) - V(\gamma) \sim 2^{-\gamma}(\log 2)^2\big(\gamma\log 2 - 1\big).$$

For large γ, the factor $(\gamma \log 2 - 1)$ is positive, and the contributions from $k \geq 3$ are exponentially smaller, so $S''(\gamma) > 0$ for sufficiently large γ. Moreover,

$$S'(\gamma) = -\gamma V(\gamma) \sim -\gamma 2^{-\gamma}(\log 2)^2 \to 0^-$$

as $\gamma \to \infty$.

B.3 Rigorous bounds on the intermediate interval

To complete the proof rigorously, we must verify the inequality $\gamma\mu_3(\gamma) > V(\gamma)$ on a compact interval $I = [\gamma_1, \gamma_2]$ that connects the asymptotic regimes (for instance $I = [1.01, 10]$). On such a compact set, all series involved are absolutely convergent and smooth.

The cumulants can be expressed as

$$\zeta^{(r)}(\gamma) = (-1)^r \sum_{k=1}^{\infty} k^{-\gamma}(\log k)^r,$$

and thus

$$V(\gamma) = K''(\gamma) = \frac{\zeta''(\gamma)\zeta(\gamma) - (\zeta'(\gamma))^2}{\zeta(\gamma)^2}, \qquad \mu_3(\gamma) = -K^{(3)}(\gamma).$$

We split the sums into a finite partial sum up to N and a tail:

$$\sum_{k=1}^{\infty} f(k;\gamma) = \sum_{k=1}^{N} f(k;\gamma) + \sum_{k=N+1}^{\infty} f(k;\gamma).$$

For $k \geq N + 1$, the terms $k^{-\gamma}(\log k)^r$ are eventually decreasing in k and can be bounded using integral tests or the Euler–Maclaurin summation formula. Specifically, for a decreasing, positive function $g(x)$,

$$\sum_{k=N+1}^{\infty} g(k) \leq \int_N^{\infty} g(x)\,\mathrm{d}x,$$

and more precise two-sided bounds with explicit remainder are available from Euler–Maclaurin.

By computing the partial sums numerically with high precision and bounding the tails analytically, we obtain rigorous enclosures for $\zeta(\gamma)$, $\zeta'(\gamma)$, and $\zeta''(\gamma)$, and hence for $V(\gamma)$ and $\mu_3(\gamma)$, on the entire interval I. These enclosures can be propagated through the expression $S''(\gamma) = \gamma\mu_3(\gamma) - V(\gamma)$.

Numerical verification using high-precision interval arithmetic shows that the lower bound of $S''(\gamma)$ is strictly positive for all $\gamma \in I$. Combining this with the asymptotic regimes treated in the previous subsections, we conclude that $S''(\gamma) > 0$ for all $\gamma > 1$.

References

[1] Hiroaki Kitano. Biological robustness. *Nature Reviews Genetics*, 5(11):826–837, 2004.

[2] J. M. Carlson and J. Doyle. Complexity and robustness. *Proceedings of the National Academy of Sciences*, 99(suppl 1):2538–2545, 2002.

[3] Andreas Wagner. *The Origins of Evolutionary Innovations*. Oxford University Press, 2011.

[4] Albert W. Marshall, Ingram Olkin, and Barry C. Arnold. *Inequalities: Theory of Majorization and Its Applications*. Springer, second edition, 2011.

[5] Arpita Ghosh, Stephen Boyd, and Amin Saberi. Minimizing effective resistance of a graph. *SIAM Review*, 50(1):37–66, 2008.

[6] M. E. J. Newman. Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 46(5):323–351, 2005.

[7] E. T. Jaynes. Information theory and statistical mechanics. *Physical Review*, 106(4):620–630, 1957.

[8] Anna D. Broido and Aaron Clauset. Scale-free networks are rare. *Nature Communications*, 10:1017, 2019.