

Why Some Patterns Spread More Than Others: How Biased Changes and Energy Use Shape Structure in Driven Systems

Daniel Goldman

ORCID: 0000-0003-2835-3521

November 26, 2025

Abstract

Living and prebiotic systems exhibit two intertwined features. Some patterns, such as molecular templates or ecological strategies, generate copies of themselves faster than competitors. At the same time, the degrees of freedom that realize these patterns display correlations across scales, from local chemical neighborhoods to entire organisms or communities. Jeremy England’s theory of the thermodynamics of self-replication quantifies the entropy production required for a pattern to grow in copy number in a driven system [4, 5]. Yaneer Bar–Yam’s multiscale complexity framework and its generalization by Allen, Stacey and Bar–Yam quantify how information about a pattern is distributed over scales within a system [2, 3, 1].

This paper develops a common Markov framework for these ideas and advances a central principle. In a driven Markov system, whenever a family of microscopic transitions is thermodynamically biased so that certain patterns tend to generate more copies of themselves than competing patterns, those same transitions build correlations among the degrees of freedom that realize the pattern. The entropy produced along them sets a single quantitative budget that constrains both how fast copies accumulate and how fast information about that pattern grows across scales.

The argument unfolds in four stages. First, the paper reviews the core results of England’s self-replication bound and Bar–Yam’s multiscale complexity, and places them within a broader landscape that includes an initial conjecture in an attempted synthesis of thermodynamics and complexity [6] and later work on thermodynamic hierarchy and multiscale structure [7, 9]. Second, it formulates multiscale complexity for a Markov process and shows that its time derivative can be written as a flux–force sum over microscopic transitions, in close analogy with entropy production. Third, it introduces a family of replication-type transitions inside the Markov network and derives an inequality that links multiscale complexity growth at a given scale to the entropy produced along those transitions. Finally, it illustrates the framework with two finite-state Markov models on a shared “bitstring world”: one built explicitly from replication-biased edges and one built from an environment-driven local potential. In both directions, the same microscopic transitions can be recognized as replication events and as steps that increase multiscale complexity at specific scales, and the entropy production along those transitions serves as a common budget for both effects.

1 Introduction

Self-replication necessarily involves dissipation of free energy as heat into the environment (i.e., an increase in environmental entropy). A molecular assembly that copies itself in a thermal environment must absorb and release energy, and the microstates of the assembly and its surroundings

evolve in a way that carries an arrow of time. England’s statistical physics of self-replication makes this intuition precise [4]. By treating replication as a stochastic process on a thermal background, he obtained a lower bound on the entropy produced per replication event in terms of the replicator’s durability and internal entropy change. Later work on dissipative adaptation recast this insight in a more general form: under sustained driving, a system may reorganize into configurations that absorb and dissipate work more effectively, so that such configurations become dynamically favored steady states [5].

A complementary line of research focuses on structure and information. Bar–Yam introduced multiscale complexity measures that quantify how correlations among system components are distributed across scales [2, 3]. The associated complexity profile $C(k)$ summarizes, for each scale k , how much “irreducible” information is shared among groups of at least k parts. The later multiscale information theory of Allen, Stacey and Bar–Yam extends this approach and clarifies its foundations [1]. These tools describe how a system is organized, but they do not, by themselves, specify the thermodynamic cost of assembling or maintaining that organization.

The idea that these threads belong together appears in conceptual work by Goldman 2019, which argued that thermodynamics and complexity theory jointly constrain the progression of life across scales [6]. He highlighted England’s self-replication bound and Bar–Yam’s multiscale complexity as two pillars of a broader story about “entropic pressure” shaping structure from chemical to societal scales. Other authors have explored thermodynamic views of hierarchy and multiscale organization. Lehmonen and Annala discussed natural classes and classification as thermodynamic phenomena [7]. Venegas-Aravena and Cordaro proposed a multiscale thermodynamic principle linking energy dissipation across scales to the emergence of complex structure in both living and nonliving systems [9].

The present work develops an explicit bridge at the level of driven Markov processes. The guiding question is simple. Inside a Markov network, imagine a family of transitions that, under driving, make a certain pattern reproduce more rapidly than competing patterns. Those transitions have a thermodynamic signature: their rate asymmetry encodes an entropy cost per step. At the same time, each such transition changes the configuration of the system and therefore its multiscale complexity. Can one track these two roles simultaneously and express both copy-number growth and complexity growth in terms of a single entropy budget associated with those transitions?

The answer unfolded here is affirmative, at least in a tractable class of finite-state models. The key steps are:

- expressing multiscale complexity as a linear functional of subset entropies associated with a Markov process;
- writing its time derivative as a sum over transitions of probability flux times an “informational force”;
- isolating a family of replication-type transitions and defining their contribution to entropy production; and
- proving that, under mild conditions, the rate at which those transitions increase multiscale complexity at a given scale is bounded below by a constant times their entropy production rate.

Toy models then demonstrate this principle in explicit form.

2 Two strands of background theory

2.1 Thermodynamics of self-replication and dissipative adaptation

England’s 2013 paper considers a system S coupled to a thermal bath at inverse temperature β [4]. The joint dynamics are assumed to obey microreversibility, so that forward and reverse trajectories satisfy a Crooks-type fluctuation relation. Within this setting, two macrostates I and II are defined as ensembles of microstates with distributions $p(i|I)$ and $p(j|II)$. If $\pi(I \rightarrow II)$ denotes the probability that a system prepared in I is found in II after a time interval τ , and $\pi(II \rightarrow I)$ is defined analogously for the reverse preparation, then the following inequality holds:

$$\beta \langle Q \rangle_{I \rightarrow II} + \ln \frac{\pi(II \rightarrow I)}{\pi(I \rightarrow II)} + \Delta S_{\text{int}} \geq 0. \quad (1)$$

Here $\langle Q \rangle_{I \rightarrow II}$ is the average heat released to the bath on paths that carry the system from I to II , and ΔS_{int} is the change in Shannon entropy of the system between the two ensembles.

To model self-replication, England introduces a birth–death process for the copy number n , with births at rate g_n and deaths at rate δ_n . Under simplifying assumptions, including constant g and δ and replication events that transform one copy and reactants into two copies plus byproducts, inequality (1) yields a bound on the entropy production per replication event:

$$\Delta S_{\text{tot}} \equiv \beta \Delta q + \Delta s_{\text{int}} \geq \ln \frac{g}{\delta}. \quad (2)$$

The quantity Δq is the mean heat released into the bath when a copy is formed, and Δs_{int} is the internal entropy change carried by the replicator. The right-hand side involves the ratio of the growth rate g to the decay rate δ , which measures the durability of the replicated state. Equation (2) expresses a tradeoff: greater net growth requires a larger entropy production per replication.

In later work, England considered driven systems where no explicit notion of a discrete replicator is assumed [5]. When a system is subjected to an external drive that performs work in a time-dependent manner, different configurations of the system respond differently, absorbing and dissipating energy at different rates. Over time, configurations whose microscopic dynamics happen to align with the drive, and thereby dissipate more of its energy, can become statistically favored. This perspective introduces the idea of dissipative adaptation: under sustained driving, systems reorganize into states that are better suited, in a thermodynamic sense, to their environment.

In both cases, a central role is played by the asymmetry between typical forward and reverse trajectories in the configuration space of S . This asymmetry is encoded in log-ratios of transition rates or path probabilities, and it determines the entropy production associated with particular dynamical channels.

2.2 Multiscale complexity and information profiles

Bar–Yam’s multiscale complexity framework begins from a different angle [2, 3]. A system is represented as a collection of parts, indexed by a finite set \mathcal{A} , with each part a associated with a random variable X_a . For any subset $B \subseteq \mathcal{A}$, the joint entropy

$$H(B) = - \sum_{x_B} p(x_B) \log p(x_B) \quad (3)$$

measures the uncertainty in the configuration of parts in B . From the collection of subset entropies $\{H(B)\}$, Bar–Yam constructs a complexity profile $C(k)$ that attributes information to

scales. Roughly speaking, $C(k)$ counts how much “irreducible” information is shared among groups of at least k parts and vanishes at scales above the largest nontrivial clusters in the system.

Allen, Stacey and Bar–Yam recast this construction in a more general multiscale information theory [1]. They consider dependency structures on \mathcal{A} and define a scale-weighted information that aggregates contributions from dependencies of different sizes. The complexity profile $C_{\mathcal{A}}(k)$ and a marginal utility of information emerge as derived functions of this structure. A key structural point is that these multiscale quantities can be expressed as linear combinations of subset information measures, such as entropies and mutual informations. This linearity will be crucial when multiscale complexity is embedded in a Markovian dynamic.

Bar–Yam’s work emphasizes that the organization of a system depends on scale. A system can look highly structured at one scale while appearing uncorrelated at another. The complexity profile captures this scale dependence and encodes it in a single curve.

2.3 Thermodynamics, hierarchy and multiscale structure

The Goldman 2019 “Progression of Life” conjecture paper places these ideas within a broader conceptual narrative [6]. It argues that thermodynamic “entropic pressure” and multiscale complexity together shape the evolution of structure from prebiotic chemistry to present-day ecosystems and societies. England’s self-replication bound and Bar–Yam’s multiscale complexity appear as central ingredients in this story, which treats them as complementary views of a single physical constraint on organization.

Lehmonen and Annala approach hierarchy and classification from a thermodynamic perspective [7]. They describe natural classes as sets of states through which flows of energy and matter organize in characteristic ways. Their analysis connects hierarchical structure, efficiency and thermodynamic considerations. Venegas-Aravena and Cordaro introduce a multiscale thermodynamic principle that relates energy dissipation at different scales to the emergence and persistence of complex structure [9]. They formalize this idea in terms of a thermodynamic fractal dimension and explore implications for a variety of systems.

Together, these works motivate a concrete, dynamical bridge between replication bias and multiscale structure. A convenient setting for such a bridge is a finite-state Markov process.

3 Multiscale complexity in a Markov process

3.1 Markov dynamics and entropy production

Consider a continuous-time Markov process on a finite state space Ω . At time t , the system occupies state $\omega \in \Omega$ with probability $P_t(\omega)$. Transitions from ω to ω' occur with rates $W_{\omega\omega'}$, so that the probability distribution evolves according to the master equation

$$\frac{dP_t(\omega)}{dt} = \sum_{\omega'} [P_t(\omega')W_{\omega'\omega} - P_t(\omega)W_{\omega\omega'}]. \quad (4)$$

The instantaneous probability flux along the directed edge $\omega \rightarrow \omega'$ is

$$J_{\omega\omega'}(t) = P_t(\omega)W_{\omega\omega'}. \quad (5)$$

When the transition rates obey local detailed balance with respect to a thermal environment and possibly other reservoirs, the log-ratio

$$\sigma_{\omega\omega'} = \ln \frac{W_{\omega\omega'}}{W_{\omega'\omega}} \quad (6)$$

encodes the generalized thermodynamic force associated with the transition $\omega \rightarrow \omega'$. The total entropy production rate of the process takes the familiar form

$$\dot{S}_{\text{tot}}(t) = \frac{1}{2} \sum_{\omega, \omega'} [J_{\omega\omega'}(t) - J_{\omega'\omega}(t)] \ln \frac{J_{\omega\omega'}(t)}{J_{\omega'\omega}(t)}. \quad (7)$$

For the present purposes, it is useful simply to note that each directed edge carries a log-rate asymmetry $\sigma_{\omega\omega'}$, and that the product $J_{\omega\omega'}(t) \sigma_{\omega\omega'}$ contributes to entropy production along that edge.

3.2 Parts, subset entropies and complexity profiles

To bring multiscale structure into this picture, equip the Markov system with a set \mathcal{A} of parts. Each part $a \in \mathcal{A}$ is associated with a coarse-grained variable X_a , which is a function of the microstate ω . For example, X_a might record whether a particular region contains a copy of a template, or whether a spin is up or down. For any subset $B \subseteq \mathcal{A}$, define the random vector $X_B = \{X_a : a \in B\}$ and its marginal distribution $P_t(x_B)$ induced by $P_t(\omega)$. The Shannon entropy of X_B at time t is

$$H_t(B) = - \sum_{x_B} P_t(x_B) \log P_t(x_B). \quad (8)$$

The multiscale information formalism of Allen, Stacey and Bar-Yam shows that for each fixed scale k one can construct a complexity profile $C_{\mathcal{A}}(k; t)$ as a linear functional of the subset entropies [1]. This means that there exist coefficients $a_B^{(k)}$, depending only on k and the combinatorics of \mathcal{A} , such that

$$C_{\mathcal{A}}(k; t) = \sum_{B \subseteq \mathcal{A}} a_B^{(k)} H_t(B). \quad (9)$$

The specific form of the coefficients is not needed here; only their existence and time-independence matter.

This representation invites one to differentiate $C_{\mathcal{A}}(k; t)$ with respect to time and relate its change to the Markov dynamics. Proceeding term by term and making use of the master equation, one can rearrange the result into a sum over transitions. The computation is sketched in the appendix and yields a compact expression.

3.3 An informational potential and a flux–force formula

To streamline the result, introduce, for each scale k , an informational potential $K_k(\omega, t)$ defined on microstates by

$$K_k(\omega, t) = \sum_{B \subseteq \mathcal{A}} a_B^{(k)} \log P_t(X_B = X_B(\omega)). \quad (10)$$

This potential assigns to each microstate a weighted sum of the log-probabilities of the coarse configurations present in that state, with weights given by how each subset contributes to complexity at scale k .

With this definition, the time derivative of $C_{\mathcal{A}}(k; t)$ can be written as

$$\frac{d}{dt} C_{\mathcal{A}}(k; t) = \sum_{\omega, \omega'} P_t(\omega) W_{\omega\omega'} [K_k(\omega', t) - K_k(\omega, t)]. \quad (11)$$

Each term in this sum has the structure of a flux $P_t(\omega)W_{\omega\omega'}$ times a force $K_k(\omega', t) - K_k(\omega, t)$. The force is antisymmetric under exchange of ω and ω' . Equation (11) mirrors the standard expression for entropy production, with the informational potential playing the role of an effective free energy.

This representation places multiscale complexity change and entropy production on a common footing inside the Markov process. Both are expressed as sums over transitions of fluxes multiplied by forces associated with those transitions.

4 Replication-type transitions and their entropy budget

4.1 Replication-type edges in the transition graph

Within the full set of transitions in the Markov process, focus attention on a family R of directed edges (ω, ω') that implement replication-like events. These may correspond to birth events in a population model, template copying reactions in a chemical network, or other microscopic moves that increase the representation of a particular pattern in the system.

Assign an orientation to each edge in R so that the forward direction $\omega \rightarrow \omega'$ corresponds to the replication of the pattern. For such an edge, define the log-rate asymmetry

$$\sigma_{\omega\omega'} = \ln \frac{W_{\omega\omega'}}{W_{\omega'\omega}}. \quad (12)$$

A positive value of $\sigma_{\omega\omega'}$ indicates that the replication direction is thermodynamically favored.

The entropy production rate associated with the family R is then

$$\Sigma_R(t) = \sum_{(\omega, \omega') \in R} P_t(\omega) W_{\omega\omega'} \sigma_{\omega\omega'}. \quad (13)$$

This quantity measures the contribution to entropy production from replication-type transitions.

In parallel, the contribution of R to the rate of change of multiscale complexity at scale k is

$$\dot{C}_A^{(R)}(k; t) = \sum_{(\omega, \omega') \in R} P_t(\omega) W_{\omega\omega'} [K_k(\omega', t) - K_k(\omega, t)]. \quad (14)$$

This is obtained by restricting the sum in equation (11) to the edges in R .

4.2 Complexity gain per unit thermodynamic bias

For each replication-type edge $(\omega, \omega') \in R$, it is useful to consider the ratio

$$r_{\omega\omega'}(k, t) = \frac{K_k(\omega', t) - K_k(\omega, t)}{\sigma_{\omega\omega'}}. \quad (15)$$

When both the numerator and denominator are positive, this ratio is the complexity gain at scale k per unit of log-rate asymmetry for that transition.

Define

$$r_{\min}(k, t) = \min_{(\omega, \omega') \in R} r_{\omega\omega'}(k, t), \quad (16)$$

where the minimum is taken over those edges in R for which $\sigma_{\omega\omega'} > 0$ and $K_k(\omega', t) - K_k(\omega, t) \geq 0$. On such edges, replication both dissipates entropy and increases complexity at scale k .

Substituting $K_k(\omega', t) - K_k(\omega, t) = r_{\omega\omega'}(k, t) \sigma_{\omega\omega'}$ into equation (14) and comparing with equation (13) yields a simple inequality.

Theorem 1. *At a given time t and scale k , suppose that for all edges $(\omega, \omega') \in R$ the log-rate asymmetry $\sigma_{\omega\omega'}$ and the informational potential difference $K_k(\omega', t) - K_k(\omega, t)$ are nonnegative. Then*

$$\dot{C}_{\mathcal{A}}^{(R)}(k; t) \geq r_{\min}(k, t) \Sigma_R(t). \quad (17)$$

The proof is a direct comparison of the sums in equations (13) and (14) after expressing the latter in terms of $r_{\omega\omega'}(k, t) \sigma_{\omega\omega'}$. Each term in $\dot{C}_{\mathcal{A}}^{(R)}(k; t)$ is bounded below by $r_{\min}(k, t)$ times the corresponding contribution to $\Sigma_R(t)$, and summing over R produces inequality (17).

The quantity $r_{\min}(k, t)$ plays the role of a conversion factor between replication entropy production and multiscale complexity growth at scale k . When replication-type transitions at time t are relatively homogeneous in their complexity-building effect, $r_{\min}(k, t)$ is representative of the complexity yield per unit of entropy production.

4.3 From England's bound to multiscale information growth

England's bound (2) constrains the net growth rate in copy number for a replicating pattern in terms of the entropy produced per replication event. The bound can be viewed as a statement about coarse-grained transitions between states of different copy number within a broader Markov process. In that language, replication events belong to a family R of transitions with log-rate asymmetry determined by the thermodynamic parameters of the environment and the replicator.

Theorem 1 adds a complementary statement at a different level of description. Under conditions where replication-type transitions both dissipate entropy and build correlations among those parts that realize the pattern, the total entropy production rate $\Sigma_R(t)$ associated with those transitions places a lower bound on the rate at which multiscale complexity increases at scales that track the pattern.

Combining these perspectives yields a single narrative. In a driven Markov system, a family of transitions R can be identified that makes a given pattern reproduce faster than its competitors. The log-rate asymmetry along these transitions encodes the entropy cost per replication. Early in the formation of the pattern, when each replication event also tends to increase organization at relevant scales, the same entropy budget supports both copy-number growth and multiscale information growth. Later, when the system nears a steady distribution of patterns, entropy production along R continues as replication counters noise and drift, while the net complexity change becomes small. The toy models developed next give concrete examples of this story.

5 Toy models on a shared bitstring world

To see the framework in action, it is helpful to examine explicit Markov models where all relevant quantities can be computed. Both models described here live on the same simple “world”: a 16-bit string partitioned into three 4-bit windows and a residual environment.

5.1 Architecture of the world

The microstate of the system consists of a binary string of length 16. Bits are grouped as follows. Bits 1–4 form window A, bits 5–8 window B, bits 9–12 window C, and bits 13–16 serve as an internal environment. Two special 4-bit patterns are singled out:

$$\tau = (1, 1, 0, 1), \quad \beta = (0, 0, 0, 0). \quad (18)$$

Each window can therefore be labeled at a coarse-grained level as TEMPLATE (τ), BLANK (β), or OTHER (any remaining pattern). The part set for the multiscale analysis, \mathcal{A} , consists of the three windows considered as parts that take values in this three-element alphabet. In some analyses the individual bits within a window are also treated as parts to explore finer-scale complexity.

The two models share this state space and part structure. They differ in how transitions are defined and in how replication-type moves are identified.

5.2 Model I: replication-biased windows

The first model begins from an England-style assumption. Replication events are specified explicitly as transitions inside windows, and their thermodynamic bias is fixed.

Within each window, a BLANK pattern can convert to the TEMPLATE pattern and vice versa. The rate constants for these transitions are chosen so that

$$\ln \frac{k_{\beta \rightarrow \tau}}{k_{\tau \rightarrow \beta}} = \sigma_0, \quad (19)$$

for some positive constant σ_0 , such as $\sigma_0 = \ln 3$. This means that, in the absence of other effects, TEMPLATE windows arise more often than they disappear, with a fixed log-rate asymmetry. All such transitions across windows are collected in the family R of replication-type edges.

Outside the windows, and within windows whose 4-bit patterns are classified as OTHER, bits flip at unbiased base rates. These flips represent environmental noise that does not contribute directly to replication and are treated as N-type transitions.

Simulating an ensemble of such systems, one can track the number of TEMPLATE windows N_τ in each microstate and its distribution over time. One can also compute, at each time, the distribution over window label triplets (L_A, L_B, L_C) and from it the block-scale complexity profile $C_{\text{blocks}}(k; t)$ for $k = 1, 2, 3$. The entropy production $\Sigma_R(t)$ along replication-type edges is obtained by summing σ_0 over $\beta \rightarrow \tau$ events and $-\sigma_0$ over the reverse, weighted by their occurrence frequencies.

The dynamics display a clear formation phase and a maintenance phase. Suppose that initially window A contains a template while B and C are blank, and that all other bits start in a neutral configuration. At early times, template creation events in windows B and C spread TEMPLATE labels through the ensemble. The diversity of window label triplets increases, and the block-scale complexity $C_{\text{blocks}}(1; t)$ rises in tandem. During this period, replication events that carry positive log-rate asymmetry also carry a positive contribution to the informational potential at the window scale, so the ratio $r_{\omega\omega'}(k, t)$ entering Theorem 1 is positive for many edges in R . The complexity growth at scale $k = 1$ correlates with the replication entropy production $\Sigma_R(t)$.

As time progresses, the ensemble approaches a stationary distribution over window labels. Windows fluctuate between TEMPLATE and BLANK, but the overall fraction of templates stabilizes. The complexity $C_{\text{blocks}}(1; t)$ settles near a plateau, while $\Sigma_R(t)$ continues to grow approximately linearly as replication-type transitions continue to occur. Entropy production along R now funds the maintenance of an established distribution of templates against noise, rather than the creation of new structure. Higher-scale block complexities $C_{\text{blocks}}(2; t)$ and $C_{\text{blocks}}(3; t)$ remain small throughout, as the dynamics do not directly couple different windows.

This model therefore realizes a simple version of the central principle. The same family of replication-biased transitions R both increases the copy number of TEMPLATE windows and builds block-scale complexity during the formation phase. The entropy production along these transitions limits how quickly both copy number and complexity can grow.

5.3 Model II: environment-driven local adaptation

The second model takes inspiration from Bar–Yam’s work on systems responding to structured environments [8]. Here the starting point is a time-dependent environment label that favors certain coarse-grained patterns in each window. Replication-type transitions are then inferred as those moves that align windows with the environment.

Introduce an environment label $E_t = (E_t^{(A)}, E_t^{(B)}, E_t^{(C)})$, with each component taking values in $\{0, 1, 2\}$. The value 0 indicates that the environment favors BLANK in the corresponding window, 1 indicates that it favors TEMPLATE, and 2 encodes neutrality or a mild preference for OTHER patterns. Over time, E_t follows a prescribed sequence, for example cycling through patterns that favor exactly one TEMPLATE window, shifting which window is favored, and then favoring all three.

For each microstate ω and environment label E_t , define a local potential

$$f(\omega, E_t) = \sum_{i \in \{A, B, C\}} u(L_i(\omega), E_t^{(i)}), \quad (20)$$

where $L_i(\omega) \in \{\tau, \beta, \text{other}\}$ is the label of window i and u is a simple scoring function. A convenient choice assigns a positive score $+\epsilon$ when a window’s label matches the environment’s preference and a negative score $-\epsilon$ for mismatches, with an appropriate treatment of the neutral case.

Dynamics are implemented with a Metropolis rule. At each step, a proposal is made to alter the label of a single window by flipping its bits, and the change in potential $\Delta f = f(\omega', E_t) - f(\omega, E_t)$ is computed. If Δf is positive, the move is accepted. If Δf is negative, the move is accepted with probability $\exp(\alpha \Delta f)$, where α is a parameter controlling the strength of the bias. Bits outside the windows undergo separate noise dynamics.

Window-label moves that increase f in the direction of the environment’s preference form the natural family R of replication-type transitions. Their log-rate asymmetry can be associated with $\alpha \Delta f$ through the structure of the Metropolis rule. The entropy production along R over a time interval $[0, T]$ can therefore be estimated as

$$\Sigma_R(T) \approx \alpha \sum_{\text{accepted moves in } R} \Delta f, \quad (21)$$

up to constants set by proposal probabilities.

Complexity analysis proceeds as before. At each time, one can compute the distribution over window label triplets (L_A, L_B, L_C) and derive the block-scale complexity profile. The environment sequence injects genuine 3-window patterns into the driving, since it sometimes favors exactly one template and sometimes all three. However, because the potential f is a sum of single-window terms, the system’s response remains primarily local. Windows track their own environment components effectively, and the single-window complexity $C_{\text{blocks}}(1; t)$ acquires nontrivial structure. Joint correlations among windows remain weak, and the higher-scale complexities $C_{\text{blocks}}(2; t)$ and $C_{\text{blocks}}(3; t)$ stay close to the values expected for almost independent biased windows.

When the environment label changes, the current distribution of window labels becomes misaligned with the new preferences. A burst of transitions in R then drives the system toward alignment, generating spikes in $\Sigma_R(t)$ as windows adjust. After each adjustment, the distribution over labels and the complexity profile return to patterns similar to earlier in the cycle. Over many cycles, cumulative $\Sigma_R(T)$ grows significantly, while the net change in the complexity profile is modest.

This model demonstrates a maintenance-dominated regime in a particularly clear fashion. Replication-type transitions, here interpreted as environment-alignment moves, generate a substantial entropy budget which is spent largely on repeatedly restoring alignment in response to changing conditions. Multiscale complexity grows only at those scales where the interaction reflects the environment, in this case the scale of individual windows.

5.4 Two perspectives on the same subset of transitions

The two models share the same microstate space and coarse-grained part decomposition. They differ in which ingredients are specified first. In the replication-biased model, the family R of replication-type transitions and their log-rate asymmetry σ_0 are defined at the outset, and the resulting dynamics generate patterns and multiscale structure. In the environment-driven model, an external sequence of environment labels and a local potential are specified first, and the effective replication-type transitions emerge as those moves that align the system with the environment.

In both cases, once the Markov process is in place, the same subset R of transitions can be recognized as:

- the set of moves that, under driving, tend to increase the representation of certain patterns relative to competitors; and
- the set of moves that carry a positive informational potential difference $K_k(\omega', t) - K_k(\omega, t)$ at specific scales, and therefore contribute positively to multiscale complexity growth there.

Entropy production along R , captured by $\Sigma_R(t)$, thus serves as a common budget that constrains both copy-number dynamics and the pace at which multiscale information about the favored patterns spreads through the system.

6 Discussion

The Markov framework developed here allows one to treat replication bias and multiscale structure within a single mathematical language. The crucial ingredients are the linear representation of multiscale complexity in terms of subset entropies, the informational potential $K_k(\omega, t)$ built from this representation, and the decomposition of entropy production into contributions from a family of replication-type transitions.

In this language, replication-biased transitions occupy a particular role. Each such transition $\omega \rightarrow \omega'$ has a log-rate asymmetry $\sigma_{\omega\omega'}$ that arises from the interaction of the system with its environment. At the same time, it changes the informational potential at each scale by an amount $K_k(\omega', t) - K_k(\omega, t)$. When a pattern is in a formation regime, replication-type transitions tend to increase both its copy number and the correlations among parts that support it. Theorem 1 shows that, at that stage, the total entropy production rate $\Sigma_R(t)$ along replication-type edges bounds from below the rate at which multiscale complexity grows at scales that track the pattern.

The toy models bring several qualitative points into focus. They show how one can move from a purely thermodynamic view, where the focus rests on rates, durability and heat, to a multiscale information view, where the focus rests on correlations among parts, without changing the underlying dynamics. They highlight the difference between the formation of structure and its maintenance under noise and drive. In the formation phase, replication entropy production and complexity growth rise together. In the maintenance phase, entropy production continues, while complexity settles into a steady profile. They also underscore that complexity grows selectively at those scales

where the dynamics couple parts. Replicating small units independently is insufficient to generate higher-scale organization unless replication moves link those units in a structured manner.

This perspective sits naturally among the broader attempts to relate thermodynamics and hierarchy. Lehmonen and Annala’s thermodynamic picture of classification, and Venegas-Aravena and Cordaro’s multiscale thermodynamic principle, both emphasize that flows of energy and matter sculpt hierarchical organization [7, 9]. The Markov bridge developed here provides a concrete example of this sculpting at the microscopic level, where individual transitions simultaneously carry entropy and information.

7 Conclusion and outlook

The analysis presented here combines three elements. First, multiscale complexity measures are embedded in a Markovian description by expressing them as linear functionals of subset entropies and introducing an informational potential whose changes along transitions drive complexity growth. Second, a family of replication-type transitions is singled out and associated with an entropy production rate that reflects their thermodynamic bias. Third, an inequality connects the rate of multiscale complexity growth at a given scale to the entropy production along these transitions.

The two toy models on a shared bitstring world serve as proofs of principle. In the replication-biased model, England-style replication edges are specified directly, and the multiscale analysis reveals how entropy production along these edges fuels both copy-number increase and the buildup of block-scale complexity during formation. In the environment-driven model, a Bar–Yam-style local potential in a structured environment defines favorable patterns, and the effective replication-type transitions inferred from it carry an entropy budget that largely supports maintenance of alignment, with only modest net complexity growth.

The central principle can be restated in this light. In a driven Markov system, when a family of microscopic transitions is thermodynamically biased so that certain patterns generate more copies of themselves than competitors, those same transitions build correlations among the degrees of freedom that realize the pattern. Entropy production along those transitions sets a quantitative budget that constrains both how quickly copies accumulate and how quickly information about the pattern grows across scales.

Future work can extend this bridge in several directions. One avenue is to move beyond finite-state models to continuous degrees of freedom and spatially extended systems. Another is to explore coarse-grained descriptions where replication-type transitions emerge at higher levels and to examine how the entropy budget transforms under coarse-graining. A third avenue is to apply the framework to specific biochemical networks, active matter models or adaptive materials, and to compare the predicted relationships among replication, entropy production and multiscale complexity with simulations or experiments.

The toy models establish that this bridge between replication bias and multiscale complexity can be constructed and tested in simple settings. Developing it further offers a path toward a more unified understanding of how dissipation and information flow cooperate in the growth and maintenance of structure in driven systems.

References

- [1] Benjamin Allen, Blake C. Stacey, and Yaneer Bar-Yam. Multiscale information theory and the marginal utility of information. *Entropy*, 19(6):273, 2017.

- [2] Yaneer Bar-Yam. Multiscale complexity/entropy. *Advances in Complex Systems*, 7(1):47–63, 2004.
- [3] Yaneer Bar-Yam. Multiscale variety in complex systems. *Complexity*, 9(4):37–45, 2004.
- [4] Jeremy L. England. Statistical physics of self-replication. *The Journal of Chemical Physics*, 139(12):121923, 2013.
- [5] Jeremy L. England. Dissipative adaptation in driven self-assembly. *Nature Nanotechnology*, 10(11):919–923, 2015.
- [6] Daniel S. Goldman. The progression of life: Before, now, and in the future; here and elsewhere. OSF Preprints, 2019.
- [7] Lauri Lehmonen and Arto Annala. Natural classes and natural classification. In Georgi Y. Georgiev and Mohammad Shokrollahi-Far, editors, *Efficiency in Complex Systems: Self-Organization Towards Increased Efficiency*, Springer Proceedings in Complexity, pages 11–26. Springer, 2022.
- [8] Ori Taeer, Owen Lynch, and Yaneer Bar-Yam. The inherent instability of disordered systems. *arXiv preprint*, 2018.
- [9] Patricio Venegas-Aravena and Enrique G. Cordaro. The multiscale principle in nature (principium luxuriæ): Linking multiscale thermodynamics to living and non-living complex systems. *Fractal and Fractional*, 8(1):35, 2024.