

## Article

# Quantifying Knowledge Production Efficiency with Thermodynamics: A Data-Driven Study of Scientific Concepts

Artem Chumachenko \*  and Brett Buttlere 

Centre for European Regional and Local Studies (EUROREG), Science Studies Laboratory, University of Warsaw, Krakowskie Przedmieście 30, 00-927 Warsaw, Poland

\* Correspondence: a.chumachenko@uw.edu.pl

## Abstract

We develop a data-driven framework for analyzing how scientific concepts evolve through their empirical in-text frequency distributions in large text corpora. For each concept, the observed distribution is paired with a maximum entropy equilibrium reference, which takes a generalized Boltzmann form determined by two measurable statistical moments. Using data from more than 500,000 physics papers (about 13,000 concepts, 2000–2018), we reconstruct the temporal trajectories of the associated MaxEnt parameters and entropy measures, and we identify two characteristic regimes of concept dynamics, stable and driven, separated by a transition point near criticality. Departures from equilibrium are quantified using a residual-information measure that captures how much structure a concept exhibits beyond its equilibrium baseline. To analyze temporal change, we adapt the Hatano–Sasa and Esposito–Van den Broeck decomposition to discrete time and separate maintenance-like contributions from externally driven reorganization. The proposed efficiency indicators describe how concepts sustain or reorganize their informational structure under a finite representational capacity. Together, these elements provide a unified and empirically grounded description of concept evolution in scientific communication, based on equilibrium references, nonequilibrium structure, and informational work.

**Keywords:** nonequilibrium thermodynamics; thermodynamics of information; Hatano–Sasa decomposition; irreversible work; information efficiency; residual entropy; maximum entropy principle; instantaneous fixed point (IFP); entropy–energy bond; phase transition; finite systems

## 1. Introduction

Entropy occupies a central position in both physics and information theory as a measure of multiplicity, uncertainty, and disorder. Since the work of Boltzmann and Gibbs, thermodynamic entropy has quantified the number of microscopic configurations compatible with a macroscopic state. Shannon [1] extended this concept to symbolic communication, defining information entropy as a measure of uncertainty within message ensembles. Jaynes [2] later showed that the two are not merely analogous but formally connected: statistical mechanics can be reformulated as an inference procedure based on the maximum entropy (MaxEnt) principle, in which physical laws provide the relevant macroscopic constraints. This insight laid the foundation of modern statistical inference and has inspired extensive research linking informational and thermodynamic quantities across physics, chemistry, biology, and cognitive systems [3–5].



Academic Editor: Miguel Rubi

Received: 15 November 2025

Revised: 12 December 2025

Accepted: 19 December 2025

Published: 22 December 2025

**Copyright:** © 2025 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

However, the relationship between thermodynamic and information-theoretic entropy remains a topic of conceptual debate [6,7]. Thermodynamic entropy, rooted in classical physics, measures the logarithm of the number of microstates compatible with a set of macroscopic observables, whereas Shannon entropy quantifies uncertainty associated with a probability distribution. Although their mathematical forms coincide, their physical interpretation aligns only when the probabilities in Shannon's expression correspond to a MaxEnt equilibrium distribution—the least biased distribution consistent with known constraints. In this sense, Jaynes's identification of the two entropies is conditional, not universal. As several authors emphasize [8,9], entropy depends on the descriptive level adopted by the observer. Caticha's entropic-inference framework [10,11] clarifies that the quantity appearing in the exponent of a Boltzmann distribution is simply the observable whose expectation value is constrained in the MaxEnt problem; its interpretation—physical or informational—depends on the modeling context. Gao et al. [12] formalized this connection by showing that thermodynamic and Shannon entropy coincide exactly when the underlying distribution is a generalized Boltzmann distribution, while recent work demonstrates that Clausius entropy can be derived directly from Boltzmann–Gibbs–Shannon entropy under equilibrium assumptions [13]. These results delineate the conditions under which thermodynamic interpretations of informational models are mathematically grounded.

An essential insight from modern nonequilibrium thermodynamics is provided via Landauer's principle [14], which establishes that logically irreversible transformations of information have an intrinsic thermodynamic cost. In its classical formulation, reducing the Shannon entropy of a system—such as erasing one bit—requires a corresponding increase in environmental entropy, thereby ensuring non-negative total entropy production. Contemporary developments in stochastic thermodynamics [4,5,15–19] interpret this relation as a direct consequence of the second law applied to information-bearing degrees of freedom: any operation that decreases uncertainty must be offset with compensating dissipation. Along with Brillouin's notion of negentropy [20], Landauer's principle illustrates a general structure in which entropy balances link informational and thermodynamic descriptions. Although our framework does not involve physical heat flows, these results establish a coherent informational–thermodynamic perspective in which entropy balances constrain transformations of structured information. This perspective motivates the use of entropy-based measures and potential-like quantities when analyzing how empirical information is organized and reorganized in large corpora.

Within this informational–thermodynamic perspective, scientific communication can be regarded as an open informational system in which symbolic units and concepts appear with varying frequency across documents. Because each document provides only a finite representational context, the corpus distributes its limited expressive capacity unevenly among concepts: some receive substantial attention, while others appear only sparsely. This scarcity drives structural constraints that shape how knowledge propagates, leading to the empirical observation that concept-frequency distributions in scientific corpora are consistently heavy-tailed and follow power-laws [21–23]. This phenomenon is not unique to text; similar heavy-tailed patterns are endemic to a variety of complex systems, ranging from social collaboration and urban growth to the distribution of wealth and the topology of biological networks [24–26].

The maximum entropy (MaxEnt) principle provides the direct theoretical basis for a thermodynamic interpretation of these heavy-tailed distributions. The Random Group Formation (RGF) framework proposed by Baek, Bernhardsson, and Minnhagen [21] is a key model in this context. While RGF uses a complex cost function derived from a mutual entropy minimization procedure, its resulting power–exponential distribution,  $N(k) \propto e^{-bk}/k^\gamma$ , where  $N(k)$  is the number of groups with  $k$  elements, is analytically

derived from a minimal information cost constraint. Crucially, the functional form of this outcome is shown to be equivalent to applying the MaxEnt method with a constraint on the mean logarithmic group size,  $\langle \ln k \rangle$ . This logarithmic term,  $\ln k$ , associated with the group size  $k$ , quantifies the localization cost required to specify an element within it. This rigorous equivalence is formalized by Visser's analysis [27], which demonstrates that generalized RGF distributions can be precisely derived from the MaxEnt method for Shannon entropy by applying constraints to both the mean group size,  $\langle k \rangle$ , and the logarithmic group size,  $\langle \ln k \rangle$ . Together, these results show that heavy-tailed frequency distributions have a well-defined MaxEnt equilibrium form determined by the first moment and the mean logarithmic size, providing a principled baseline for interpreting empirical distributions.

A few studies have proposed thermodynamic or entropy-based frameworks for analyzing social communication or symbolic systems. Among these, the work by Peng [28] offers an important early attempt to draw an analogy between social collaboration and Boltzmann statistics. However, Peng's formulation relies on simplifying assumptions that differ from the MaxEnt-based approach adopted here. In particular, the entropy is defined with respect to a combinatorial upper bound, rather than being derived from empirical constraints on the underlying frequency distribution. The logarithmic cost term  $\ln k$  introduced in the model (where  $k$  denotes the number of edits to a given Wikipedia page) is chosen to reproduce the observed heavy-tailed pattern, but it does not emerge from a constrained entropy-maximization procedure such as those used in later RGF and MaxEnt formulations. As a consequence, Peng's model provides a useful qualitative analogy but offers limited guidance for analyzing empirical distributions over time or for characterizing their nonequilibrium evolution under external influences.

A recent approach by Giardini and daCunha [29] addresses time-based evolution by developing a Thermodynamics of Innovation model. However, this framework focuses on cumulative adoption quantities over time, constructing a canonical ensemble to model the temporal evolution of the aggregate population using Gompertz-like and Maxwell–Boltzmann-like shapes. While successful at describing time-series dynamics, this cumulative focus is orthogonal to the need for a frequency-resolved (cross-sectional) analysis of heterogeneous concepts within the literary corpus.

More broadly, the existing approaches do not provide a data-driven, frequency-resolved thermodynamic framework for symbolic communication systems. In particular, no prior work derives equilibrium reference states directly from empirical term-frequency distributions or establishes a principled way to compare these empirical states with their MaxEnt equilibrium counterparts, nor does it model their temporal evolution within an open-system, grand-canonical formalism. Consequently, current models cannot quantify nonequilibrium structure, entropy production, or the thermodynamic efficiency of informational change at the level of individual concepts.

The present work addresses this gap by developing a thermodynamic framework in which each scientific concept is treated as an open, frequency-resolved informational system. The empirical term-frequency distribution defines the concept's nonequilibrium mesoscopic state, while a MaxEnt-derived generalized Boltzmann distribution provides the corresponding equilibrium reference. The discrepancy between these two distributions—captured via residual entropy and a free-energy-like gap—quantifies the informational structure and stability of each concept. By applying a discrete-time version of the Hatano–Sasa [30] and Esposito–Van den Broeck [4] decomposition, we further separate maintenance-like (housekeeping) contributions from externally driven reorganization. This equilibrium/nonequilibrium pairing is, to our knowledge, new to the study

of scientific communication and forms the backbone of our analysis of concept evolution, stabilization, and informational efficiency.

The remainder of the paper is organized as follows. Section 2 introduces the data and formalizes the thermodynamic framework, including the MaxEnt construction of equilibrium reference states, the definition of residual entropy and free-energy measures, and the Hatano–Sasa/Esposito–Van den Broeck decomposition for nonequilibrium evolution. Section 3 presents an empirical analysis of more than 11,000 scientific concepts, detailing their trajectories in thermodynamic state space, identifying characteristic equilibration and driving regimes, and evaluating dissipation and efficiency patterns. Section 4 discusses the theoretical and empirical implications of the proposed framework for understanding concept stabilization and semantic innovation.

## 2. Materials and Methods

### 2.1. Database Description

We used a corpus of 451,524 English-language research articles from the High Energy Physics and Astronomy sections of arXiv (2000–2018). Each document contains full text and standard metadata and was annotated with term-frequency counts for 13,945 scientific concepts drawn from a curated ontology [31–33].

Related studies have analyzed this corpus for concept co-occurrence structure, innovation emergence, and topic organization [22,34,35]. Here, we use the same concept-usage data but develop a thermodynamic, MaxEnt-based description, focusing on equilibrium reference states, residual information, and nonequilibrium informational work.

### 2.2. Thermodynamic Framework

To formalize our model, we consider the scientific concept  $c$  and define its empirical support at time  $t$  as the collection of relevant documents that contain at least one mention of the concept. The number of such documents is denoted as  $N_c(t)$  among the total number of  $N_D > N_c$  documents in the corpus. If no further information is known beyond whether a concept appears in the document or not [28], then a uniform probability,  $1/N_c(t)$ , can be assigned to each document. This yields a maximum entropy value of  $\ln N_c(t)$ , representing the highest possible uncertainty under this minimal description [36,37]. In this representation, each document is associated with a single indistinguishable *microstate* of a concept.

More detailed information on a concept state arises from the in-text term frequency analysis, where each frequency class,  $k \in \mathbb{Z}^+$ , defines a *mesostate*, a coarse-grained representation of the underlying microstates of individual concept mentions [38]. The corresponding empirical probability,  $p(k, t) = N_c(k, t)/N_c(t)$ , is proportional to the number,  $N_c(k, t)$ , of documents that mention the concept  $c$  exactly  $k$  times until time  $t$ . The Shannon entropy of this mesoscopic description,

$$S(t) = - \sum_{k=1}^{\infty} p(k, t) \ln p(k, t), \quad (1)$$

reflects additional knowledge about the structural organization of  $p(k, t)$  and, therefore, is typically smaller than the maximal value  $\ln N_c(t)$  corresponding to uniform ignorance.

In the early stage of its appearance, a concept is supported by only a small number of documents, and its mesostate,  $p(k, t)$ , is correspondingly sparse and almost identical to its microstate configuration. As the concept accumulates sufficient usage across documents, the empirical distribution  $p(k, t)$  stabilizes and develops a broad, approximately heavy-tailed form. This enables a power–exponential fit,  $p(k, t) \propto k^{-\beta_t} e^{-\lambda_t k}$ , defining the concept’s instantaneous fixed point (IFP) [38]. Extensive empirical comparisons with other heavy-tailed distributions performed in our previous work [23] show that this distribution

provides the tightest upper envelope (maxentropic bound) across the entire observed range of mesostate Shannon entropies.

As was shown by Visser [27], the distribution parameters  $\beta$  and  $\lambda$  can be obtained from the maximization of Shannon entropy  $S$  under two empirical constraints: the mean frequency  $\langle k \rangle_p$  and the logarithmic moment  $\langle \ln k \rangle_p$ , both calculated from the observed empirical distribution  $p(k, t)$ . This constrained maximization yields the following [22]:

$$\pi(k, t; \beta, \lambda) = \frac{1}{Z} \frac{e^{-\lambda k}}{k^\beta}, \quad Z = \sum_{k=1}^{\infty} \frac{e^{-\lambda k}}{k^\beta} = Li_\beta(e^{-\lambda}), \quad (2)$$

where  $Z$  is the normalization constant expressed through the polylogarithm function  $Li_\beta(e^{-\lambda})$ , and  $\beta, \lambda > 0$  are Lagrange multipliers determined as follows:

$$\langle k \rangle_p = \langle k \rangle_\pi = \frac{Li_{\beta-1}(e^{-\lambda})}{Li_\beta(e^{-\lambda})}, \quad \langle \ln k \rangle_p = \langle \ln k \rangle_\pi = -\frac{\partial_\beta Li_\beta(e^{-\lambda})}{Li_\beta(e^{-\lambda})}. \quad (3)$$

The entropy of the resulting macrostate then follows as [22]

$$\begin{aligned} S_{\text{IFP}} &= - \sum_{k=1}^{\infty} \pi(k, t) \ln \pi(k, t) \\ &= \ln Z + \langle \ln k \rangle_p + \lambda \langle k \rangle_p. \end{aligned} \quad (4)$$

The distribution  $\pi(k, t)$  belongs to the class of generalized Boltzmann distributions, that is, exponential-family distributions obtained as MaxEnt solutions under constraints on  $\langle \ln k \rangle$  and  $\langle k \rangle$ . In the entropic–inference viewpoint [2,10,11], the observables that appear in the exponent play the role of effective “energies”, and the associated Lagrange multipliers are their conjugate intensive parameters. In our case, the constrained quantity  $E(k) = \ln k$  acts as an internal informational energy, while  $k$  controls the occupancy of the concept. The mean values

$$U = \langle \ln k \rangle_\pi, \quad N = \langle k \rangle_\pi$$

therefore represent, respectively, the internal informational energy and the mean number of logical particles (concept mentions) in the ensemble. The normalization constant  $Z$  plays the role of a generalized partition function for this open informational system.

For systems described by generalized Boltzmann distributions, the Gibbs–Shannon entropy coincides (up to an additive constant) with the thermodynamic entropy in the formal sense demonstrated in Refs. [12,13,39]. Under this formal equivalence, it is convenient to introduce effective intensive parameters

$$T = \frac{1}{\beta}, \quad \mu = -\frac{\lambda}{\beta},$$

which mirror, at the level of Legendre structure, the roles played by temperature and chemical potential in physical thermodynamics. We emphasize that  $T$  and  $\mu$  here are *informational* intensive parameters, not physical quantities.

With this notation, the thermodynamic-like potential takes the standard grand-canonical form

$$S_{\text{therm}} = k_B S_{\text{IFP}} = k_B (\ln Z + \beta U - \mu \beta N), \quad (5)$$

providing a consistent set of state variables for analyzing the equilibrium and nonequilibrium properties of concept–frequency distributions.

The notion of internal equilibrium is essential: although concept usage evolves over time and the empirical distribution is rarely at equilibrium globally, within each time

window, the IFP defines a locally equilibrated reference state that maximizes entropy under the empirical constraints. This is precisely the regime in which generalized Boltzmann distributions acquire a consistent thermodynamic structure [12,39,40].

### 2.3. Residual Entropy and Free Energy

We refer to a concept as being in *equilibrium* when its empirical entropy,  $S$ , coincides with the corresponding MaxEnt value  $S_{\text{IFP}}$ . In this regime, the macroscopic descriptors  $(U, N, T, \mu)$  inferred from the MaxEnt distribution remain effectively constant, and the informational structure of the concept is stabilized.

Departures from this equilibrium reference are quantified by the residual entropy

$$R(t) = S_{\text{IFP}}(t) - S(t) = D_{\text{KL}}(p(t) \| \pi(t)) = \sum_k p(k, t) \ln \frac{p(k, t)}{\pi(k, t)} \geq 0, \quad (6)$$

which measures the additional information required to specify the empirical state relative to its MaxEnt equilibrium projection [22]. In this sense,  $R$  plays the role of an information distance from the IFP manifold.

To analyze this nonequilibrium structure, we introduce the corresponding *informational* grand potentials,

$$\Phi_{\text{IFP}} = -T \ln Z, \quad \Phi = U_p - TS - \mu N_p, \quad (7)$$

which follow the usual Legendre form of generalized Boltzmann ensembles. These quantities are not physical energies or work; rather, they are formal thermodynamic potentials arising from the MaxEnt representation of the concept distribution.

Using Equations (3), (6) and (7), the difference between equilibrium and non-equilibrium informational grand potentials at fixed  $(T, \mu)$  or, equivalently,  $(U, N)$  satisfies the following:

$$\Phi - \Phi_{\text{IFP}} = T R \geq 0. \quad (8)$$

This is the exact information-theoretic analogue of the nonequilibrium free-energy identity in stochastic thermodynamics [4]: the excess informational grand potential equals the residual entropy multiplied by the intensive parameter  $T$ . The identity in Equation (8) has the same mathematical structure as the Landauer-type relation: reductions in uncertainty (here represented as  $R$ ) require a corresponding decrease in a potential-like quantity. In physical systems, this manifests as heat dissipation; in our formal, informational setting it quantifies the minimal “informational work” needed to maintain structure beyond the MaxEnt equilibrium. Thus, the relation provides a Landauer-type interpretation without assuming any physical energy flows.

Across a finite interval,  $\Delta t$ , between initial ( $i$ ) and final ( $f$ ) states, the change

$$\Delta R = \Delta S_{\text{IFP}} - \Delta S = R^f - R^i \quad (9)$$

captures how the concept’s nonequilibrium organization evolves. This form is directly analogous to the generalized second-law expressions of Esposito and Van den Broeck, which separate changes in entropy into equilibrium and nonequilibrium components in Landauer-like identities. As the mesostate distribution  $p(k, t)$  is updated through new documents,  $\Delta R$  typically decreases, but  $S$  and  $S_{\text{IFP}}$  may vary non-monotonically during semantic innovation or attention shifts. Temporary increases in  $R$  reflect transient reorganization, rather than a violation of the long-term relaxation trend.

At a fixed  $T$ , the corresponding change in excess potential satisfies

$$\Delta(\Phi - \Phi_{\text{IFP}})_T \approx T \Delta R, \quad (10)$$



so that decreases in  $R$  correspond to decreases in the excess informational grand potential, consistent with relaxation toward the IFP manifold whenever the empirical state moves closer to its MaxEnt reference.

#### 2.4. Dissipation Towards a Stationary Reference State

As a baseline, we first consider the regime in which the instantaneous fixed point  $\pi$  is stationary and the thermodynamic control parameters  $(T, \mu)$  remain constant. Under these conditions, the empirical distribution relaxes towards its IFP,  $p \rightarrow \pi$ , in the absence of external driving.

For such isothermal, fixed- $\mu$  transitions, the integrated second law can be written in the Esposito–Van den Broeck form [4]:

$$W_{\text{irr}} = T \Delta S_i + T \Delta R, \quad (11)$$

where  $\Delta S_i \geq 0$  is the irreversible entropy production [4,15]. This identity follows from the standard stochastic thermodynamic decomposition of entropy production and work for open systems [4,15]; a detailed derivation adapted to our notation is provided in Appendix B.

When  $\pi$  is stationary, local detailed balance implies (see Appendix C)

$$\frac{dS_i}{dt} = - \frac{dR}{dt},$$

which integrates over a relaxation interval to

$$\Delta S_i = -\Delta R, \quad \Rightarrow \quad W_{\text{irr}} = 0. \quad (12)$$

This result characterizes the simplest form of conceptual dynamics: In the absence of external driving, any departure of  $p$  from its fixed point is dissipated entirely as entropy production, and the residual entropy  $R(t)$  monotonically decreases. In this regime,  $R(t)$  behaves as a Lyapunov function, governing the convergence of the empirical distribution to the IFP manifold [15,18,41,42].

In physical systems, this regime corresponds to standard isothermal relaxation processes: with fixed intensive parameters  $(T, \mu)$ , probability mass is redistributed across microstates until the stationary distribution is reached, while the irreversible work  $W_{\text{irr}}$  vanishes over the full relaxation process. In our informational setting, an analogous behavior is observed for concepts whose empirical distributions remain close to a fixed IFP and whose inferred intensive parameters  $(T, \mu)$  vary only weakly over time. Such concepts operate in a regime of thermodynamic buffering, where incoming documents primarily refine the sampling of an essentially stationary reference distribution, rather than driving the system away from it.

As we show below, this buffering behavior can be understood in terms of the geometry of the MaxEnt manifold. Concepts whose fitted parameters  $(\beta, \lambda)$  place them near the heavy-tailed region of the parameter space ( $\beta \approx 3/2$ ,  $\lambda \approx 0$ ) exhibit markedly increased thermodynamic response coefficients. In this region, the curvature of the MaxEnt surface becomes large, making the macroscopic quantities  $(U, N)$  highly sensitive to small perturbations of  $(T, \mu)$ . Empirical concepts rarely reach the singular power-law limit  $\lambda \rightarrow 0$ , but many lie close enough to it to display a characteristic “critical-like” enhancement of heat capacity and susceptibility. Paradoxically, this heightened responsiveness enables certain concepts to keep their intensive parameters nearly constant: fluctuations in  $(U, N)$  are absorbed without substantial movement in  $(T, \mu)$ , producing the observed thermodynamic

buffering. In this sense, the equilibrated portion of the vocabulary acts as an effective informational reservoir.

Within this framework, the cutoff  $\lambda$  sets the effective extent of accessible frequency states, while  $\beta$  controls their scaling and thereby governs the proximity to the heavy-tailed regime. To quantify the resulting degrees of responsiveness, we consider the standard thermodynamic response coefficients associated with the MaxEnt model.

The heat capacities

$$C_\mu = \left( \frac{\partial U}{\partial T} \right)_\mu, \quad \alpha_U = \left( \frac{\partial U}{\partial \mu} \right)_T, \quad (13)$$

describe how the informational energy  $U$  responds to variations in the intensive parameters, while the susceptibilities

$$\chi_T = \left( \frac{\partial N}{\partial \mu} \right)_T, \quad \alpha_\mu = \frac{1}{N} \left( \frac{\partial N}{\partial T} \right)_\mu, \quad (14)$$

characterize the corresponding response of the mean concept usage  $N$ .

Because all of these coefficients are derived from the same grand-potential function, they share the same qualitative structure and exhibit parallel signatures near the heavy-tailed region of the MaxEnt manifold. For empirical analysis, it is, therefore, sufficient to work with a reduced pair of combined differential coefficients that capture the joint variations of both  $U$  and  $N$ . To this end, we define the effective energy  $E = U - \mu N$ , whose differential

$$dE = \left( \frac{\partial E}{\partial T} \right)_\mu dT + \left( \frac{\partial E}{\partial \mu} \right)_T d\mu = C_\mu^E dT + \alpha_E d\mu, \quad (15)$$

provides a compact summary of responsiveness through the coefficients  $(C_\mu^E, \alpha_E)$ . These quantities behave analogously to the full set of heat capacities and susceptibilities but offer a clearer geometric interpretation on the  $(T, \mu)$  manifold.

Large values of  $(C_\mu^E, \alpha_E)$  identify concepts with enhanced responsiveness—characteristic of systems approaching the heavy-tailed boundary—where extensive quantities exhibit substantial fluctuations while intensive parameters vary only weakly. In contrast, dynamically driven concepts typically appear *before* reaching these critical regions: their IFP trajectories  $(T(t), \mu(t))$  begin to deviate from the stationary manifold while the response coefficients are still rising but have not yet diverged. This pre-critical regime reflects the onset of semantic reorganization, where changes in topic context or conceptual usage start to generate irreversible work,  $W_{\text{irr}} > 0$ , before full buffering behavior is lost. The next subsection examines this driven regime in detail and quantifies how much of the resulting dissipation supports structural adaptation versus the maintenance of the existing conceptual organization.

## 2.5. Dynamic Efficiency and Non-Equilibrium Work Decomposition (Driving Case)

To analyze concept evolution under external informational driving, we now make explicit the discrete-time structure that was implicit in the preceding discussion. We represent the corpus as a sequence of snapshots,  $t_0 < t_1 < \dots < t_f$ , with each corresponding to an observational window (e.g., one year). At every snapshot,  $t_i$ , we compute the empirical mesostate  $p_{t_i}(k)$  and its instantaneous fixed point (IFP),  $\pi_{t_i}(k)$ , obtained from the MaxEnt construction described above. Each step,  $t_i \rightarrow t_{i+1}$ , therefore constitutes a transition between coarse-grained nonequilibrium steady states (NESSs) in the sense of Hatano–Sasa [30] and the Esposito–Van den Broeck framework for Markovian thermodynamics [4,15]. The resulting trajectory,  $\{p_t, \pi_t\}$ , captures how the internal energy,  $U(t)$ , and particle number,  $N(t)$ , reorganize in response to external informational inputs.



In the Esposito–Van den Broeck framework, the total irreversible entropy production generated via a finite driven evolution admits the decomposition

$$\Delta S_{\text{tot}} = \Delta S_{\text{na}} + \Delta S_{\text{a}} \geq 0. \quad (16)$$

The non-adiabatic component  $\Delta S_{\text{na}}$  quantifies entropy generated via changes in the IFP parameters (structural or contextual adaptation), while the adiabatic component  $\Delta S_{\text{a}}$  corresponds to the steady “house-keeping” dissipation required to maintain the nonequilibrium structure when the parameters are held fixed. In the absence of driving (stationary  $\pi_t$ ), this reduces to the intrinsic entropy production discussed earlier.

For discrete-time dynamics, the non-adiabatic contribution can be written in terms of changes in residual entropy change, together with an excess term associated with parameter driving. Following the Esposito–Van den Broeck Markov formulation (see Appendix D for derivation), we obtain

$$\Delta S_{\text{na}} = -\Delta R + \sum_t \hat{Y}_{t \rightarrow t+1}, \quad \Delta S_{\text{a}} = \Delta S_{\text{tot}} - \Delta S_{\text{na}}, \quad (17)$$

where  $-\Delta R \geq 0$  is the boundary term produced via changes in the KL divergence  $D_{\text{KL}}(p_t \| \pi_t)$ , and

$$\hat{Y}_{t \rightarrow t+1} = \sum_k p_t(k) \ln \frac{\pi_t(k)}{\pi_{t+1}(k)} \quad (18)$$

is the ensemble-averaged excess contribution generated purely via changes in the IFP parameters  $(T(t), \mu(t))$ . This term is the discrete counterpart of the Hatano–Sasa excess functional for driven steady states.

Using (16) and (17) yields the driving/housekeeping split [30]:

$$W_{\text{driving}} = T_{\text{ref}} \Delta S_{\text{na}}, \quad W_{\text{hk}} = T_{\text{ref}} \Delta S_{\text{a}}, \quad W_{\text{irr}} = W_{\text{ex}} + W_{\text{hk}}. \quad (19)$$

Thus,  $W_{\text{ex}}$  quantifies the energetic cost of structural adaptation to changing semantic conditions, while  $W_{\text{hk}}$  represents steady-state dissipation required to maintain the nonequilibrium distribution.

The reference temperature  $T_{\text{ref}}$  specifies the statistical environment (bath) for evaluating work quantities. We use either (i) a global reference  $T_{\star}$  obtained from the ontology-wide mode, or (ii) a topic-specific temperature  $T_{\text{core}}$  estimated from equilibrated core concepts. Because efficiencies depend only on entropy-production ratios, all reported efficiency measures are dimensionless and independent of the choice of  $T_{\text{ref}}$ .

To evaluate how dissipation partitions between maintenance and adaptation, we introduce the dimensionless efficiencies

$$\eta_{\text{hk}} = \frac{\Delta S_{\text{a}}}{\Delta S_{\text{tot}}}, \quad \eta_{\text{ex}} = \frac{\Delta S_{\text{na}}}{\Delta S_{\text{tot}}} = 1 - \eta_{\text{hk}}. \quad (20)$$

In addition, we define the residual–information ratio

$$\rho_R = \frac{[-\Delta R]_+}{\Delta S_{\text{tot}}}, \quad (21)$$

which quantifies the fraction of total entropy production that is effectively used to reduce the residual informational structure. This measure is analogous in spirit to the information–erasure efficiencies introduced by Allahverdyan et al. [43].

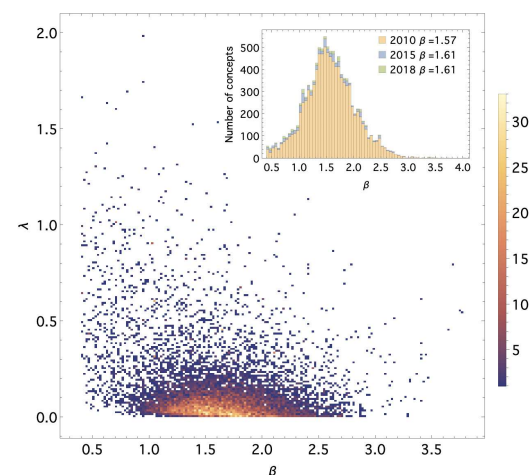
High values of  $\eta_{\text{hk}}$  indicate stabilized concepts whose evolution is dominated by steady maintenance, while high  $\eta_{\text{ex}}$  identifies adaptive concepts whose informational structure is being reorganized over the interval  $\Delta t$ .

These two regimes manifest equally clearly in the behavior of the residual–information ratio  $\rho_R$ . In the absence of external driving, relaxation proceeds entirely through the elimination of residual entropy, so  $\rho_R$  attains its maximal value (Equation (12)). Under driving, however, part of the dissipation is diverted toward accommodating changes in the evolving IFP, leading to reduced values of  $\rho_R$ . Thus,  $\rho_R \approx 1$  signals simple relaxation toward an equilibrium, whereas smaller values reveal periods of genuine adaptive reorganization in response to shifts in the semantic environment.

### 3. Results

For each of the 13,945 concepts in the ontology, we estimated the MaxEnt parameters  $(\beta, \lambda)$  of the macrostate (IFP) distribution (Equation (2)) using maximum-likelihood estimation. To ensure internal consistency, each fitted pair was required to reproduce the empirical moment constraints  $\langle k \rangle$  and  $\langle \ln k \rangle$  (Equation (3)) within a numerical tolerance of  $10^{-5}$ . Concepts with a small number of mentions provided insufficient information for stable parameter inference and were excluded; over 2000–2018, this affected 2208 early-stage concepts, leaving 11,737 analyzable cases.

Figure 1 shows the empirical distribution of fitted  $(\beta, \lambda)$  values across all concepts for the period of 2000–2018. Most estimates lie within the region  $\beta \in [1, 2]$  and  $\lambda \in [0, 0.15]$ , with a persistent mode near  $\beta \simeq 1.5$ . Time-averaged values of  $\beta$  remain remarkably stable ( $\bar{\beta} \approx 1.51$  through 2010 and  $\bar{\beta} \approx 1.61$  in 2018), consistent with the heavy-tailed usage patterns typical of scientific terminology [22].



**Figure 1.** Empirical distribution of fitted MaxEnt parameters  $(\beta, \lambda)$  for 11,737 scientific concepts across 2000–2018. Panels show the corpus-wide distribution of  $\beta$  over selected time periods and highlight the persistent concentration of values near  $\beta \simeq 1.5$ .

To characterize broad equilibration patterns, we examined how the fitted parameters relate to the following: (i) the time at which a concept first appears and (ii) the number of documents supporting it. Three robust empirical regularities emerge.

First, concepts introduced earlier in the corpus more frequently exhibit small residual entropy  $R$ , indicating close agreement between empirical usage distributions and the MaxEnt prediction. Recently introduced concepts also reach this regime once sufficient document support accumulates.

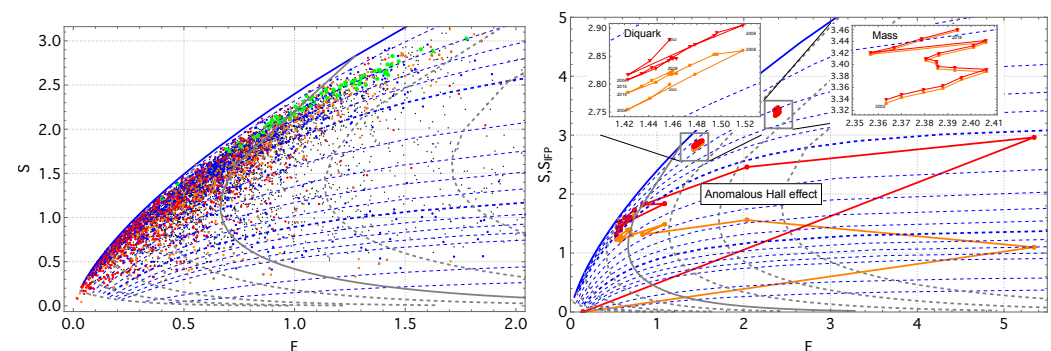
Second, the number of supporting documents  $N_c$  is a strong predictor of equilibration. Across all years, most concepts with  $N_c \gtrsim 10^3$  satisfy  $R \approx 0$ , and the empirical entropy  $S$  nearly coincides with the macrostate entropy  $S_{\text{IFP}}$ . As support increases, the fitted parameter  $\lambda$  typically decreases, yielding heavier-tailed usage consistent with the  $\lambda \rightarrow 0$  limit.

Third, equilibrated high-support concepts tend to exhibit higher mesoscopic entropy  $S$  and internal energy  $U$ , while their grand potential  $\Phi$  is lower. This reflects a reduction in the residual information potential  $TR$  as concepts stabilize. Newly introduced or sparsely used concepts remain underdetermined in short windows but become analyzable as more evidence accumulates; for example, the *Anomalous Hall effect* required approximately four years of accumulation (2000–2004) before the empirical moments could be reliably matched.

Taken together, these findings reveal a coherent empirical structure underlying concept evolution. Most concepts converge toward a narrow region of parameter space centered near  $\beta \simeq 1.5$ ; the exponential cutoff parameter  $\lambda$  decreases systematically with increasing support; and the residual entropy  $R$  falls sharply once a concept exceeds roughly  $10^3$  supporting documents. A subset of long-established, high-support concepts forms a stable background whose distributions already satisfy the MaxEnt constraints, providing a statistically coherent reference against which more adaptive concepts evolve.

### 3.1. Energy–Entropy Diagram

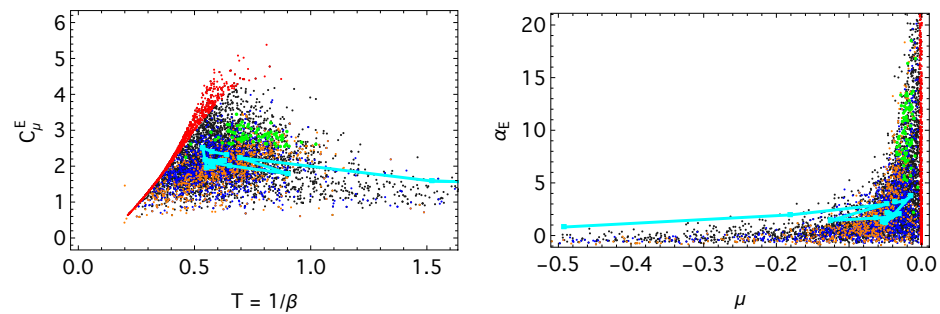
To examine the global organization of concept states, we embed each concept in the energy–entropy ( $E$ – $S$ ) plane using the effective energy  $E = \langle \ln k - \mu k \rangle$  and the empirical mesostate entropy  $S$ . Figure 2 (left) shows the resulting diagram. All concepts lie below the theoretical maximum for the macrostate entropy  $S_{\text{IFP}}(E)$ , represented by the solid blue line and computed from Equation (4) in the near–power-law limit  $\lambda = 10^{-3}$  as a function of effective energy  $E$ . The dashed blue curves show the corresponding theoretical envelopes  $S_{\text{IFP}}(E)$  for larger cutoff values  $\lambda > 10^{-3}$  (in increments of 0.01), illustrating how finite-size effects progressively suppress the maximum attainable entropy at a fixed  $E$ .



**Figure 2.** (left) Energy–entropy diagram for 11,737 concepts (2018). The green points denote stationary (equilibrated) concepts. The blue ( $R < 0.04$ ) and orange ( $R < 0.004$ ) points identify concepts with low residual entropy, while the red points highlight concepts with a near power-law term-frequency spectrum ( $\lambda < 0.04$ ). The black points represent all remaining concepts. The dashed blue and gray curves are theoretical MaxEnt envelopes,  $S_{\text{IFP}}(E; \lambda, \beta)$ , for fixed cutoff values,  $\lambda > 0$  and fixed  $\beta > 0$ , respectively. The dashed gray curves show isotherms for  $\beta \in \{0.5, 1, 2\}$ , with the solid gray curve marking  $\beta = 1.5$ . The solid blue curve represents the heavy-tailed limit  $\lambda \rightarrow 0$ , here evaluated at  $\lambda = 10^{-3}$ . (right) The trajectories of three representative concepts—*Diquark*, *Mass*, and the *Anomalous Hall Effect*. The orange lines show the empirical mesostate evolution, while the red lines show the corresponding IFP trajectory.

Figure 2 (right) illustrates representative trajectories of concept state evolution. *Diquark* and *Mass* both follow nearly stationary paths for more than a decade, maintaining almost constant effective energy  $E$  and macrostate entropy  $S_{\text{IFP}}$ . The *Diquark* trajectory shows a gradual decrease in both empirical mesostate entropy  $S$  and residual entropy  $R$ , accompanied by small oscillations in the inferred temperature. In contrast, *Mass*, probably the most basic concepts in this dataset, remains extremely close to its IFP at all times ( $\Delta R \simeq 0$ ), exhibiting only a modest increase in  $S$  and  $S_{\text{IFP}}$  while keeping  $E$  essentially unchanged.

The *Anomalous Hall effect* concept illustrates a typical emergence trajectory. Initially, both  $T$  and  $\mu$  fluctuate strongly, indicating a driving regime in which the concept adapts to a rapidly changing topical environment. As the concept stabilizes, these fluctuations diminish, and the trajectory approaches the  $\beta \simeq 1.5$  isotherm, entering a non-driving regime in which the evolution becomes effectively autonomous. This region coincides with maximal heat capacity and susceptibility (Figure 3), reflecting a buffered state where intensive variables remain stable despite continued information exchange.



**Figure 3.** Heat capacity,  $C_\mu^E$ , and susceptibility,  $\alpha_E$ , diagrams from Equation (15). The cyan curve traces the trajectory of the *Anomalous Hall Effect* concept; the trajectory terminates at the year 2018, when its estimated temperature reaches  $T \simeq 0.6$  and  $\mu \simeq -0.05$ . The color of the points denotes concept classes, as in Figure 2.

Figure 3 displays the effective response functions  $C_\mu^E$  and  $\alpha_E$ , which are representative of the broader family of thermal and chemical susceptibilities introduced in Equations (13) and (14). To relate these responses to the global entropic structure of the system, we introduce a dimensionless intensive entropy–energy ratio that empirically remains almost constant for equilibrated concepts. This ratio was previously discussed as an *entropy reduction ratio* by Peng et al. [28]:

$$TQ^{\text{eff}} = \frac{TS_{\text{IFP}}}{E} = \frac{E - \Phi_{\text{IFP}}}{E}, \quad (22)$$

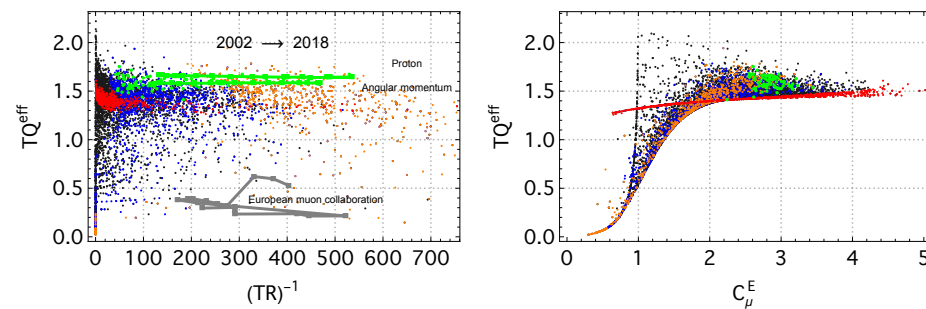
which quantifies how much equilibrium entropy a concept produces per-unit effective energy,  $E$ , times the local temperature,  $T$ . The second equality follows directly from Equations (5) and (7), using the thermodynamic identity  $\Phi_{\text{IFP}} = E - TS_{\text{IFP}}$ .

Empirically, all equilibrated concepts cluster around a characteristic value,  $TQ^{\text{eff}} \simeq 1.5$  (green points in Figure 4). Their trajectories in the  $((TR)^{-1}, TQ^{\text{eff}})$  plane remain close to this level and typically follow nearly horizontal iso- $TQ^{\text{eff}}$  curves. Such behavior reflects monotonic relaxation toward the IFP manifold at a constant temperature,  $T$ : the residual entropy decreases ( $\Delta R < 0$ ), together with the non-equilibrium free-energy difference  $\Delta(\Phi - \Phi_{\text{IFP}})$  (Equation (10)), and the irreversible entropy production remains positive. In such a stationary-reference regime, the relation  $\Delta S_i = -\Delta R$  holds, so these contributions cancel, and the net irreversible work satisfies  $W_{\text{irr}} = 0$  over the full relaxation interval.

By contrast, driven concepts exhibit deviations of various amplitudes away from the  $TQ^{\text{eff}} \simeq 1.5$  band, during which both  $R$  and  $Q^{\text{eff}}$  can temporarily deviate from their relaxation trend. These episodes signal adaptive reorganization of the usage distribution and are accompanied by Hatano–Sasa excess entropy production, indicating genuine non-equilibrium driving.

For concepts whose IFP distributions approach the power-law limit (red points in Figure 4), the quantity  $TQ^{\text{eff}}$  approaches a well-defined empirical plateau. Across the corpus, this plateau lies near  $TQ_{\text{crit}}^{\text{eff}} \approx 1.5$ , and it is associated with concepts whose thermodynamic parameters lie close to the empirical critical region around  $\beta \simeq \beta_c = 1.5$  (cf. Figure 3). To

highlight this connection, Figure 4 (right) presents the state diagram in  $(C_\mu^E, TQ^{\text{eff}})$  coordinates, where the high- $TQ^{\text{eff}}$  plateau aligns with the region of maximal response coefficients.



**Figure 4.** (left) The values of the  $TQ^{\text{eff}}$  as a function of the (inverse) residual free energy  $(TR)^{-1}$  (Equation (8)). Trajectories for the equilibrium concepts of the *Proton* and the *Angular momentum* are shown in green; the trajectory for the driven concept *European muon collaboration* is shown in a gray color. (right) The intensive entropy–energy ratio  $TQ^{\text{eff}}$  as a function of the heat capacity  $C_\mu^E$  (Equation (15)). The scatter plot for the period of 2002–2018 with the color scheme as that in Figure 2.

Most equilibrated concepts (green points) align with this plateau. This deviation reflects finite-size effects associated with strictly positive cutoff values  $\lambda$ , which limit the accessible frequency range and keep equilibrated concepts away from the heavy-tailed regime where the plateau is attained.

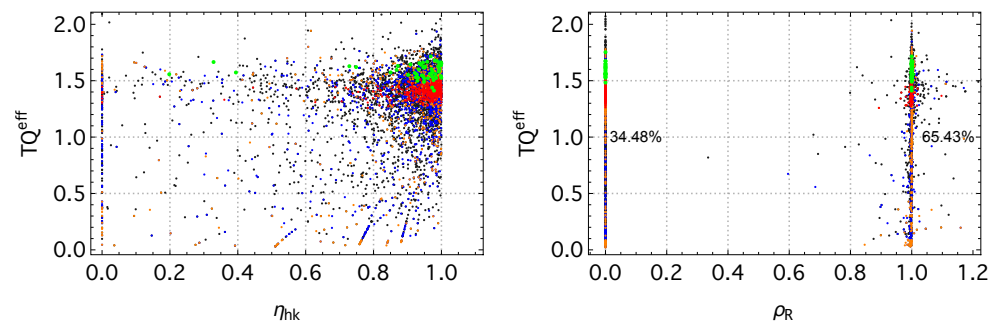
A useful way to interpret the observed plateau in the entropy–energy ratio  $TQ^{\text{eff}}$  is by analogy with classical scaling relations in statistical physics, where entropy and energy often exhibit approximately linear dependencies over characteristic regions of the state space. For example, in bounded thermodynamic systems and in hadronic matter near the Hagedorn transition, the entropy grows proportionally to energy up to a characteristic scale determined by the underlying distribution of microstates [44,45]. In our informational setting, the empirical concentration of equilibrated concepts around a nearly constant value of  $TQ^{\text{eff}}$  reflects an analogous finite-size scaling behavior: the informational “capacity” of concepts stabilizes, with equilibrated concepts generating equilibrium entropy in a near-constant proportion to their effective energy. Departures from this level marks adaptive reorganization and non-equilibrium driving, during which both  $R$  and  $TQ^{\text{eff}}$  may temporarily increase or decrease.

### 3.2. Dissipative Regimes and Efficiency of Informational Maintenance

The entropy–production decomposition introduced in Section 2.5 allows us to quantify how irreversible dissipation is partitioned between two components: (i) the housekeeping contribution  $W_{\text{hk}}$ , associated with sustaining the current nonequilibrium structure and (ii) the excess contribution  $W_{\text{driving}}$ , arising from changes in the instantaneous fixed-point (IFP) parameters  $(T, \mu)$  during externally driven evolution. The corresponding dimensionless ratios from Equation (20) characterize the relative weight of these two types of dissipation, while the residual–information ratio  $\rho_R$  (Equation (21)) measures the extent to which entropy production offsets changes in the residual entropy,  $R$ .

Figure 5 (left) displays the empirical distribution of  $\eta_{\text{hk}}$  for the 2017–2018 interval. A large fraction of concepts cluster near  $\eta_{\text{hk}} \simeq 1$ , indicating that most irreversible dissipation in this period is classified as housekeeping, rather than excess. Equilibrated concepts (green points) concentrate mostly near  $\eta_{\text{hk}} = 1$ . Concepts with nearly power-law usage spectra ( $\lambda < 0.005$ ) also predominantly lie in this region. The residual entropy value alone does not correlate strongly with  $\eta_{\text{hk}}$ , reflecting that  $\eta_{\text{hk}}$  depends on the relative magnitudes of adiabatic and non-adiabatic entropy production, rather than on the instantaneous distance from equilibrium.





**Figure 5.** (left) The product  $TQ^{\text{eff}}$  plotted against the housekeeping efficiency  $\eta_{\text{hk}}$  for the 2017–2018 interval. (right) The intensive energy–entropy ratio  $TQ^{\text{eff}}$  versus the residual–information ratio  $\rho_R$ . Color denotes the concept class, as in previous figures.

The residual–information ratio  $\rho_R$  exhibits a characteristic clustering around two values, approximately 0 and 1 (Figure 5 (right)). When  $\rho_R \approx 0$ , the residual entropy remains essentially unchanged ( $\Delta R \approx 0$ ), so no dissipation is required to compensate for changes in informational structure; all entropy production contributes to  $\Delta S_{\text{tot}}$ . In contrast, the ridge at  $\rho_R = 1$  corresponds to the case  $\Delta S_{\text{tot}} = |\Delta R|$ , characteristic of non-driven relaxation where entropy production precisely offsets the decrease in residual entropy.

In the 2017–2018 window, most concepts lie on the  $\rho_R \approx 1$  ridge (65.4%), with the remainder at  $\rho_R \approx 0$ . Equilibrated concepts appear on both ridges: some exhibit  $\Delta R \approx 0$  and thus fall at  $\rho_R \approx 0$ , while others undergo slow, undriven relaxation and satisfy  $\Delta S_{\text{tot}} = |\Delta R|$ , placing them at  $\rho_R = 1$ . Concepts with large accessible frequency ranges (high  $k_{\text{max}}$ , i.e., small  $\lambda$ ) are also largely confined to the two ridges. More adaptively evolving concepts show limited dispersion around  $\rho_R \simeq 1$ , indicating intermittent deviations from non-driven relaxation dynamics.

Overall, the empirical patterns indicate that concept evolution can be grouped into three broad operational regimes based on dissipation characteristics: (i) a *maintenance-dominated regime* ( $\eta_{\text{hk}} \approx 1$ ,  $\rho_R \approx 0$ ), where concepts exhibit minimal changes in residual structure; (ii) a *balanced relaxation regime* ( $\rho_R = 1$ ), in which the entropy production and the residual entropy change are of a comparable magnitude; and (iii) an *adaptive regime* ( $\eta_{\text{hk}} < 1$ ,  $\rho_R \neq 1$ ), characterized by significant excess dissipation and pronounced variation in the IFP parameters. This classification offers a concise, data-driven picture of how concepts allocate irreversible dissipation across different phases of their evolution.

#### 4. Discussion

The frequency distributions of scientific concepts exhibit the heavy-tailed behavior characteristic of many symbolic and social systems. Following Baek et al. [21], Visser [27], and Martini et al. [22], we model these distributions using a maximum entropy (MaxEnt) framework constrained by the empirical moments  $\langle k \rangle$  and  $\langle \ln k \rangle$ . The logarithmic moment reflects the information cost of locating a concept within a frequency class, and under these constraints, the MaxEnt solution takes the geometric–power-law form

$$p(k) \propto k^{-\beta} e^{-\lambda k},$$

which can be written as a generalized Boltzmann distribution with an effective internal energy of  $U(k) = \ln k$  and chemical potential  $\mu$ . This provides a principled equilibrium reference state for each concept, grounded directly in the observed frequency data.

The empirical frequency distribution defines a concept’s nonequilibrium mesoscopic state, while the corresponding MaxEnt solution represents its equilibrium projection. Their divergence, quantified by the residual entropy (Equation (6)), captures informational



structure not explained by the MaxEnt constraints. Residual entropy, therefore, reflects heterogeneity, historical imprint, and stabilization patterns in concept usage. A monotonic decrease in  $R$  signals relaxation toward the equilibrium reference, whereas increases or fluctuations indicate externally driven, reorganizing dynamics.

Thermodynamic-like quantities derived from MaxEnt formalism—generalized free energy, susceptibility, and heat capacity—provide interpretable summaries of these nonequilibrium states without attributing physical meaning to them. In particular, the identity

$$\Phi - \Phi_{\text{IFP}} = TR$$

links the residual entropy to a free-energy-like measure of the informational structure, consistent with relations appearing in nonequilibrium thermodynamics [4].

An analysis of state diagrams and temporal trajectories reveals two characteristic regimes:

- **Driven regime.** Concepts undergo substantial variations in  $(\beta, \lambda)$ , corresponding to changes in both the localization structure and the accessible frequency range. These parameter shifts generate non-adiabatic entropy production, captured by the Hatano–Sasa and Esposito–Van den Broeck decomposition. Such episodic, externally induced changes resemble jump-like updates in stochastic thermodynamics [18]. Concepts in this regime exhibit large residual entropy,  $R$ , and often correspond to emerging, debated, or rapidly evolving topics [22].
- **Buffered (non-driven) regime.** Concepts fluctuate around their equilibrium reference with monotonically decreasing residual entropy. Dissipation is dominated by the housekeeping component, and the system approaches or remains near its fixed-point distribution. Well-established, stable concepts populate this regime.

Empirically, most concepts transition from the driven to the buffered regime near an inverse-temperature value,  $\beta \approx 1.5$ , where heat capacity and susceptibility reach pronounced peaks. Around this point, the intensive entropy–energy ratio  $TS/E$  attains a value close to 1.5, consistent across years and across concepts. Concepts that approach this plateau typically enter a slow, buffered evolution, indicating that their frequency distributions have stabilized. Notably, while the intensive ratio remains approximately constant in the buffered regime, the residual free-energy gap  $\Phi - \Phi_{\text{IFP}}$  continues to decrease over time.

To quantify how concepts evolve across these regimes, we apply a discrete-time Hatano–Sasa/Esposito–Van den Broeck decomposition, separating entropy production into housekeeping and driving components. The resulting efficiency measures (Equations (20) and (21)) characterize how effectively a concept maintains or reorganizes its informational structure relative to total entropy production.

Most concepts operate with high housekeeping efficiency, indicating weakly driven or stabilized dynamics. Only a minority—typically those undergoing rapid semantic or topical change—exhibit substantial non-adiabatic entropy production and corresponding reductions in housekeeping efficiency. These patterns suggest that emerging or transformative concepts reside temporarily in far-from-equilibrium states before stabilizing into the buffered regime.

The framework developed here provides a unified, data-driven thermodynamic perspective on concept evolution. Pairing empirical nonequilibrium distributions with MaxEnt equilibrium references enables us to distinguish systematic structure from noise, identify stabilization and reorganization patterns, and quantify dissipation and efficiency in conceptual dynamics. Although the analogy is formal and carries no physical interpretation, the resulting thermodynamic quantities offer a coherent language for describing how scientific concepts emerge, evolve, and stabilize within the finite informational capacity of a research field.

Overall, our findings suggest that conceptual change follows measurable thermodynamic-like trajectories: from exploratory, driven evolution to buffered stability marked by entropy–energy plateaus and reduced dissipation. This perspective may support science-of-science research by identifying concepts with high transformative potential, characterizing semantic stabilization, and illuminating the collective dynamics of knowledge production.

**Author Contributions:** Conceptualization, A.C. and B.B.; methodology, A.C.; software, A.C.; validation, A.C.; formal analysis, A.C.; investigation, A.C.; resources, A.C.; data curation, A.C.; writing—original draft preparation, A.C.; writing—review and editing, A.C.; visualization, A.C.; supervision, A.C.; project administration, A.C.; funding acquisition, B.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** The University of Warsaw supported this publication under Priority Research Area V of the “Excellence Initiative—Research University” program.

**Data Availability Statement:** All materials related to the research are available and can be accessed through Zenodo: 10.5281/zenodo.17611701.

**Acknowledgments:** We would like to acknowledge that this manuscript overlaps with the preprint [46] available on Qeios. The preprint version of this work was previously made available by the authors. We also extend our gratitude to the University of Warsaw Center for European Regional and Local Studies (EUROREG) for its valuable support.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Nomenclature

$tf$	Term frequency: number of times $k \in \mathbb{Z}^+$ a concept appears in a document.
$N_c(t)$	Number of documents containing concept $c$ up to time $t$ .
$N_c(k, t)$	Number of documents containing $c$ exactly $k$ times up to time $t$ .
$p(k, t)$	Observed probability that concept $c$ is cited $k$ times up to $t$ .
$p_c(\{k\}, t) \equiv p_t$	Mesostate probability mass function at time $t$ .
$\pi^{\text{eq}}(k, t; \beta, \lambda)$	Reference probability mass function at time $t$ (thermal bath).
$\pi(k, t) \equiv \pi_t$	Macrostate PMF or instantaneous fixed point (IFP) of the mesostate $p_c$ .
$\beta$	Lagrange multiplier; inverse temperature ( $\beta = 1/(k_B T)$ ).
$\lambda$	Lagrange multiplier related to the chemical potential.
$\mu$	Chemical potential ( $\mu = -\lambda/\beta$ ).
$k_B$	Boltzmann constant (set $k_B = 1$ in numerical computations).
$Z$	Grand partition function that normalizes the macrostate PMF.
$U = \langle \ln k \rangle$	Internal (information) energy of a concept per document.
$N = \langle k \rangle$	Average number of logical particles (concept mentions).
$E = U - \mu N$	The generalized energy function.
$S$	Mesostate entropy per document derived from $p(k, t)$ .
$S_{\text{macro}} \equiv S_{\text{IFP}}$	Macrostate (thermodynamic) entropy per document.
$Q^{\text{eff}} = S/E$	Entropy efficiency.
$R$	Residual entropy, $R = S_{\text{IFP}} - S$ .
$\Phi$	Grand canonical potential, $\Phi = U - TS - \mu N$ .
$\Delta R$	Change in residual entropy over a finite interval $\Delta t$ .
$TR$	Residual free energy ( $\Phi - \Phi_{\text{IFP}}$ ).
$C_\mu$	Heat capacity at constant chemical potential, $C_\mu = (\partial U / \partial T)_\mu$ .
$C_N$	Heat capacity at constant document support, $C_N = (\partial U / \partial T)_N$ .
$\chi_T$	Particle-number susceptibility, $\chi_T = (\partial N / \partial \mu)_T$ .
$\alpha_\mu$	Thermal expansion coefficient, $\alpha_\mu = (1/N)(\partial N / \partial T)_\mu$ .
$\alpha_E$	Energy–chemical potential coupling coefficient, $\alpha_E = (\partial \langle E \rangle / \partial \mu)_T$ .
$C_\mu^E$	Heat capacity at constant chemical potential for effective energy $E$ .
$\eta_{\text{hk}}$	Housekeeping efficiency, $\eta_{\text{hk}} = \Delta S_a / \Delta S_{\text{tot}}$ .

$\eta_{\text{ex}}$	Driving (adaptive) efficiency, $\eta_{\text{ex}} = \Delta S_{\text{na}} / \Delta S_{\text{tot}} = 1 - \eta_{\text{hk}}$ .
$\rho_R$	Residual-information retention ratio, $\rho_R = 1 - [-\Delta R]^+ / \Delta S_{\text{tot}}$ .

## Appendix A. Derivation of the TR Identity for the MaxEnt Reference State

We consider MaxEnt distributions of the exponential-family form at time  $t$ :

$$\pi_t(k) = \frac{1}{Z_t} \exp[-\beta_t E_t(k)] = \frac{1}{Z_t} k^{-\beta_t} e^{-\lambda_t k}, \quad \lambda_t = \beta_t \mu_t, \quad (\text{A1})$$

where the exponent involves the constraint function

$$E(k) = \ln k - \mu k.$$

Let  $p_t(k)$  denote the empirical (nonequilibrium) distribution at time  $t$ . We define the associated informational potential,

$$\Phi[p; T, \mu] = \langle E \rangle_p - T S[p] = \sum_k p_k (\ln k - \mu k) - T \left( - \sum_k p_k \ln p_k \right), \quad (\text{A2})$$

and its MaxEnt (IFP) counterpart,

$$\Phi_{\text{IFP}}(T, \mu) = \Phi[\pi; T, \mu] = -T \ln Z. \quad (\text{A3})$$

For any  $p$  and its corresponding IFP  $\pi$  at the same  $(T, \mu)$ ,

$$\Phi[p; T, \mu] - \Phi_{\text{IFP}}(T, \mu) = T D_{\text{KL}}(p \| \pi) = T R(p \| \pi) \geq 0, \quad (\text{A4})$$

with equality only when  $p = \pi$ .

**Proof.** The Kullback–Leibler divergence is

$$R(p \| \pi) = \sum_k p_k \ln \frac{p_k}{\pi_k} = \sum_k p_k \ln p_k - \sum_k p_k \ln \pi_k.$$

Since  $\ln \pi_k = -\beta E(k) - \ln Z$ , we obtain

$$R(p \| \pi) = \sum_k p_k \ln p_k + \beta \sum_k p_k E(k) + \ln Z.$$

Multiplying by  $T$  and using  $T\beta = 1$  yields

$$TR = T \sum_k p_k \ln p_k + \sum_k p_k E(k) + T \ln Z = \Phi[p; T, \mu] - \Phi_{\text{IFP}}(T, \mu).$$

□

The identity (A4) requires only that  $(T, \mu)$  be fixed; it does not assume fixed moments such as  $\langle k \rangle$  or  $\langle \ln k \rangle$ . Differences in these moments are fully encoded in the KL divergence.

On a discrete time grid,  $\{t_n\}$ ,

$$\Phi(p_{t_n}) - \Phi_{\text{IFP}}(t_n) = T_{t_n} R(p_{t_n} \| \pi_{t_n}) \geq 0. \quad (\text{A5})$$

Thus, the decomposition

$$\Phi[p; T, \mu] = \Phi_{\text{IFP}}(T, \mu) + T R(p \| \pi) \quad (\text{A6})$$

makes explicit the contribution of the residual information  $R$  as the nonequilibrium excess over the MaxEnt reference.

## Appendix B. Irreversible Work Identity for Informational Grand Potentials

The purpose of this appendix is to show that, for exponential family (MaxEnt) distributions with constrained observables  $\langle \ln k \rangle$  and  $\langle k \rangle$ , the algebraic structure of the Legendre transform leads to an identity formally equivalent to the nonequilibrium free-energy relations derived by Esposito and Van den Broeck [4].

We define the informational grand potential

$$\Phi_t \equiv U_t - T_t S_t - \mu_t N_t,$$

where  $U_t = \langle \ln k \rangle_{p_t}$ ,  $N_t = \langle k \rangle_{p_t}$ , and  $S_t$  is the Shannon entropy. Differentiating gives the exact identity

$$d\Phi_t = dU_t - T_t dS_t - S_t dT_t - \mu_t dN_t - N_t d\mu_t. \quad (\text{A7})$$

In classical nonequilibrium thermodynamics [47,48], entropy changes are decomposed as

$$dS_t = dS_{e,t} + dS_{i,t}, \quad dS_{i,t} \geq 0,$$

where  $dS_{e,t}$  is the reversible contribution induced via the protocol, and  $dS_{i,t}$  is entropy production.

To ensure consistency with the Legendre structure of  $\Phi_t$ , we define the reversible part using the identity

$$T_t dS_{e,t} \equiv dU_t - \mu_t dN_t,$$

which reduces to the usual reversible relation when  $p_t = \pi_t$ . Thus,

$$T_t dS_{i,t} \equiv dU_t - \mu_t dN_t - T_t dS_t. \quad (\text{A8})$$

Insert Equation (A8) into the differential of  $\Phi_t$  (Equation (A7)), and simplify. Introducing the formal forcing term

$$\delta W_{\text{form}} \equiv dU_t - \mu_t dN_t,$$

one obtains the exact identity

$$T_t dS_{i,t} = \delta W_{\text{form}} - d\Phi_t - S_t dT_t - N_t d\mu_t. \quad (\text{A9})$$

Equation (A9) is a formal rewriting of the Hatano–Sasa/Esposito decomposition applied to the informational potential, rather than to a physical free energy.

Differentiating the identity  $\Phi_t = \Phi_{\text{IFP},t} + T_t R_t$  yields

$$d\Phi_t = d\Phi_{\text{IFP},t} + T_t dR_t + R_t dT_t. \quad (\text{A10})$$

Substituting Equation (A10) into Equation (A9) gives

$$T_t dS_{i,t} = \delta W_{\text{form}} - d\Phi_{\text{IFP},t} - T_t dR_t - R_t dT_t - S_t dT_t - N_t d\mu_t.$$

Integrating from  $t_i$  to  $t_f$  gives the total irreversibility:

$$W_{\text{irr}} \equiv \int_{t_i}^{t_f} \delta W_{\text{form}} - W_{\text{rev}} = \int_{t_i}^{t_f} T_t dS_{i,t} + \int_{t_i}^{t_f} T_t dR_t, \quad (\text{A11})$$

where  $W_{\text{rev}}$  denotes the reversible contribution obtained when the protocol is restricted to the IFP manifold ( $R_t = 0$ ,  $S_{i,t} = 0$ ).

For protocols that keep  $(T, \mu)$  fixed, Equation (A11) reduces to

$$W_{\text{irr}} = T \Delta S_i + T \Delta R. \quad (\text{A12})$$

Equation (A12) is the *informational* analog of the iso- $(T, \mu)$  irreversible work identity in nonequilibrium thermodynamics [4]. Here,  $T \Delta S_i$  measures irreversible dispersion generated via the protocol, and  $T \Delta R$  measures the change in the nonequilibrium structure relative to the MaxEnt baseline.

### Appendix C. Entropy Production Under Instantaneous Fixed-Point Reference

In this appendix, we demonstrate that, when the reference distribution is chosen as the instantaneous fixed point of the system, the time derivative of the residual entropy equals the negative of the entropy production rate:

$$\dot{R}(t) = -\dot{S}_i(t). \quad (\text{A13})$$

Let  $p_i(t)$  denote the empirical (mesoscopic) probability of observing the microstate  $i$  at time  $t$ , corresponding in the main text to the document-level frequency distribution  $p(k, t)$  defined in Equation (1). The instantaneous fixed point or macrostate of the system is the distribution that maximizes the Shannon entropy  $S_{\text{macro}}(t) = -\sum_i \pi_i(t) \ln \pi_i(t)$  under the current empirical constraints  $\langle f_j \rangle_{p(t)} = \langle f_j \rangle_{\pi(t)}$ . This constrained maximization yields an exponential-family form (cf. Equations (2) and (3) in the main text):

$$\pi_i(t) = \frac{1}{Z(t)} \exp \left[ -\sum_j \theta_j(t) f_j(i) \right], \quad Z(t) = \sum_i \exp \left[ -\sum_j \theta_j(t) f_j(i) \right], \quad (\text{A14})$$

where the Lagrange multipliers  $\theta_j(t)$  (identified in the main text as  $\beta(t)$  and  $\lambda(t)$ ) are determined at each instant by the empirical moments  $\langle f_j \rangle_{p(t)}$ .

The residual entropy—the informational distance between the empirical distribution and its instantaneous MaxEnt reference—is the Kullback–Leibler divergence (cf. Equation (6) in the main text):

$$R(t) = D_{\text{KL}}(p(t) \parallel \pi(t)) = \sum_i p_i(t) \ln \frac{p_i(t)}{\pi_i(t)} \geq 0. \quad (\text{A15})$$

The equality  $R = 0$  occurs only when the mesoscopic state coincides with its IFP,  $p(t) = \pi(t)$ , that is, when the system is locally in equilibrium with respect to the chosen constraints.

Taking the time derivative of Equation (A15) gives

$$\begin{aligned} \dot{R}(t) &= \sum_i \dot{p}_i \ln \frac{p_i}{\pi_i} + \sum_i p_i \frac{d}{dt} \left[ \ln \frac{p_i}{\pi_i} \right] \\ &= \sum_i \dot{p}_i \ln \frac{p_i}{\pi_i} + \sum_i p_i \left( \frac{\dot{p}_i}{p_i} - \frac{\dot{\pi}_i}{\pi_i} \right). \end{aligned} \quad (\text{A16})$$

The first sum in the second term vanishes because the probabilities are normalized,  $\sum_i \dot{p}_i = 0$ . Hence,

$$\dot{R}(t) = \sum_i \dot{p}_i \ln \frac{p_i}{\pi_i} - \sum_i p_i \frac{\dot{\pi}_i}{\pi_i}. \quad (\text{A17})$$

From Equation (A14), we have  $\ln \pi_i = -\ln Z - \sum_j \theta_j f_j(i)$ , so that

$$\frac{\dot{\pi}_i}{\pi_i} = -\partial_t \ln Z - \sum_j \dot{\theta}_j f_j(i). \quad (\text{A18})$$

Taking the expectation value with respect to the empirical distribution  $p_i$ :

$$\sum_i p_i \frac{\dot{\pi}_i}{\pi_i} = -\partial_t \ln Z - \sum_j \dot{\theta}_j \langle f_j \rangle_p. \quad (\text{A19})$$

For an exponential-family distribution, the partition function satisfies  $\partial_{\theta_j} \ln Z = -\langle f_j \rangle_\pi$ , so that  $\partial_t \ln Z = \sum_j \partial_{\theta_j} \ln Z \dot{\theta}_j = -\sum_j \dot{\theta}_j \langle f_j \rangle_\pi$ . Substituting this into Equation (A19) gives

$$\sum_i p_i \frac{\dot{\pi}_i}{\pi_i} = -\sum_j \dot{\theta}_j (\langle f_j \rangle_p - \langle f_j \rangle_\pi).$$

By definition of the instantaneous fixed point, these expectation values coincide ( $\langle f_j \rangle_p = \langle f_j \rangle_\pi$ ), and therefore,

$$\sum_i p_i \frac{\dot{\pi}_i}{\pi_i} = 0. \quad (\text{A20})$$

With Equation (A20), the residual-entropy rate (A17) reduces to

$$\dot{R}(t) = \sum_i \dot{p}_i(t) \ln \frac{p_i(t)}{\pi_i(t)}. \quad (\text{A21})$$

For a continuous-time Markov process with transition rates  $W_{ij}$ , the entropy production rate is defined as [18]

$$\dot{S}_i(t) = \frac{1}{2} \sum_{i,j} J_{ij}(t) \ln \frac{p_i(t) W_{ij}}{p_j(t) W_{ji}}, \quad J_{ij} = p_i W_{ij} - p_j W_{ji}. \quad (\text{A22})$$

If the transition rates satisfy detailed balance with respect to the instantaneous reference distribution  $\pi_i(t)$ ,

$$\pi_i(t) W_{ij}(t) = \pi_j(t) W_{ji}(t),$$

then

$$\begin{aligned} \dot{S}_i(t) &= \frac{1}{2} \sum_{i,j} (p_i W_{ij} - p_j W_{ji}) \ln \left( \frac{p_i W_{ij}}{p_j W_{ji}} \right) \\ &= \frac{1}{2} \sum_{i,j} (p_i W_{ij} - p_j W_{ji}) \ln \left( \frac{p_i / \pi_i}{p_j / \pi_j} \right). \end{aligned}$$

Define  $f_i = p_i / \pi_i$ . Using the antisymmetry of  $J_{ij} = p_i W_{ij} - p_j W_{ji}$  and the symmetry of the sum, we obtain the following:

$$\dot{S}_i(t) = \sum_{i,j} p_i W_{ij} \ln \left( \frac{f_i}{f_j} \right). \quad (\text{A23})$$

Now, consider the expression  $-\sum_i \dot{p}_i \ln f_i$ . Using the master equation  $\dot{p}_i = \sum_j (p_j W_{ji} - p_i W_{ij})$ :

$$\begin{aligned} -\sum_i \dot{p}_i \ln f_i &= -\sum_{i,j} \ln f_i (p_j W_{ji} - p_i W_{ij}) \\ &= \sum_{i,j} p_i W_{ij} \ln f_i - \sum_{i,j} p_j W_{ji} \ln f_i. \end{aligned}$$



We relabel indices  $i \leftrightarrow j$  in the second sum:

$$-\sum_i \dot{p}_i \ln f_i = \sum_{i,j} p_i W_{ij} (\ln f_i - \ln f_j) = \sum_{i,j} p_i W_{ij} \ln \left( \frac{f_i}{f_j} \right).$$

We can see that this matches Equation (A23) for  $\dot{S}_i(t)$ . Therefore, the right-hand side of Equation (A22) can be rewritten as

$$\dot{S}_i(t) = -\sum_i \dot{p}_i(t) \ln \frac{p_i(t)}{\pi_i(t)}.$$

Comparing with Equation (A21) immediately yields the desired identity

$$\dot{R}(t) = -\dot{S}_i(t). \quad (\text{A24})$$

Equation (A24) shows that, when the reference state is the instantaneous MaxEnt fixed point  $\pi(t)$  (determined self-consistently from the data at each time), the residual entropy  $R(t)$  decreases exactly at the rate at which the entropy is produced. In fact, the reference state need not be the IFP of the empirical mesostate; Equation (A21) holds when  $\dot{\pi} = 0$  and its temperature and chemical potential remained constant ( $\dot{\theta}_j = 0$ ) for the reference state. To see this, we need to compare Equations (A21) and (A32).

In both cases, the system relaxes monotonically toward  $\pi$ , and the relative entropy  $R(t)$  plays the role of a Lyapunov function:

$$\dot{R}(t) \leq 0, \quad \dot{S}_i(t) \geq 0.$$

The identification of  $R$  as a measure of “informational distance from equilibrium” justifies its interpretation as an internal measure of the system’s nonequilibrium organization, and it explains why, under IFP dynamics, entropy production precisely quantifies the rate of information dissipation.

**Mid-point approximation for an entropy-production estimate.** In the main text, we distinguish two informational quantities: (i) the residual entropy  $R(t) = D_{\text{KL}}(p_t \| \pi_t)$ , which measures the instantaneous separation between the empirical mesostate  $p_t$  and its MaxEnt reference (IFP)  $\pi_t$ , and (ii) the entropy production  $\Delta S_i$ , which measures the total irreversibility accumulated along the transition between two time points. These quantities coincide only in the special case of relaxation toward a *fixed* reference distribution; in general,  $R$  is a state function (defined at each  $t$ ), while  $\Delta S_i$  is path-dependent because the reference  $\pi_t$  changes over time.

To formalize this distinction without invoking physical heat flows, we use the standard Hatano–Sasa/Esposito–Van den Broeck informational balance for a time-dependent reference distribution  $\pi_t$ :

$$\Delta S_i = \Delta S + \Delta S_{\text{ex}}, \quad \Delta S \equiv S[p_{t_f}] - S[p_{t_i}], \quad (\text{A25})$$

where the *excess* contribution is defined as

$$\Delta S_{\text{ex}} \equiv -\int_{t_i}^{t_f} dt \sum_k p(k, t) \partial_t \ln \pi(k, t). \quad (\text{A26})$$

Equation (A25) is the informational analogue of the entropy production decomposition used in stochastic thermodynamics:  $\Delta S_i$  captures irreversibility, while  $\Delta S_{\text{ex}}$  accounts for the fact that the reference (IFP) itself moves in time under the external protocol.

Our data are available only at discrete times  $t_i$  and  $t_f$ . We therefore approximate the integral in Equation (A26) by a midpoint (Stratonovich) rule [49], i.e.,

$$\Delta S_{\text{ex}} \approx - \sum_k p_*(k) \left[ \ln \pi_f(k) - \ln \pi_i(k) \right], \quad p_*(k) \equiv \frac{1}{2} [p_i(k) + p_f(k)], \quad (\text{A27})$$

where  $p_i \equiv p(\cdot, t_i)$  and  $p_f \equiv p(\cdot, t_f)$ , and likewise  $\pi_i \equiv \pi(\cdot, t_i)$ ,  $\pi_f \equiv \pi(\cdot, t_f)$ .

In our MaxEnt construction the IFP has the generalized Boltzmann form  $\pi(k, t) \propto k^{-\beta(t)} e^{-\lambda(t)k}$ . To implement the midpoint rule consistently at the level of the protocol parameters, we also introduce midpoint intensives

$$\beta_* = \frac{1}{2} [\beta(t_i) + \beta(t_f)], \quad \mu_* = - \frac{\lambda(t_i) + \lambda(t_f)}{\beta(t_i) + \beta(t_f)}, \quad (\text{A28})$$

which correspond to the midpoint of  $(\beta, \mu)$  under the identity  $\mu(t) = -\lambda(t)/\beta(t)$ . (Equivalently, one may work with  $(\beta, \lambda)$  directly; we use  $(\beta, \mu)$  only to match the grand-canonical notation adopted in the main text.)

Finally, combining Equations (A25) and (A27), we obtain the discrete-time estimate used in the Results:

$$\Delta S_i \approx [S[p_f] - S[p_i]] - \sum_k p_*(k) \left[ \ln \pi_f(k) - \ln \pi_i(k) \right]. \quad (\text{A29})$$

This estimator makes explicit why  $\Delta S_i$  is not, in general, equal to  $\Delta R$  or to the endpoint KL divergences: it depends on the *change of the reference* along the interval, approximated here by a midpoint rule.

## Appendix D. Dynamic Efficiency and Work Decomposition in Concept Evolution

We present a derivation of the discrete-time decomposition of entropy production into non-adiabatic (excess) and adiabatic (housekeeping) parts, as given by Equations (18) and (17) of the manuscript. This derivation follows the frameworks of Hatano and Sasa [30] and Esposito–Van der Broeck [15].

In stochastic thermodynamics, the total entropy production rate  $\dot{S}_{\text{tot}}(t)$  can be split into two non-negative contributions:

$$\dot{S}_{\text{tot}} = \dot{S}_{\text{na}} + \dot{S}_{\text{a}}. \quad (\text{A30})$$

Here,  $\dot{S}_{\text{na}}$  is the non-adiabatic (excess) entropy production rate due to the change of control parameters, and  $\dot{S}_{\text{a}}$  is the adiabatic (housekeeping) entropy production rate required to maintain the non-equilibrium steady state, even when the control parameters are fixed.

Following [15], the non-adiabatic part can be expressed as

$$\dot{S}_{\text{na}} = - \frac{d}{dt} D(p_t \| \pi_t) + \dot{S}_{\text{ex}}, \quad (\text{A31})$$

where the excess entropy production rate  $\dot{S}_{\text{ex}}$  is given by

$$\dot{S}_{\text{ex}} = - \sum_k p_t(k) \partial_t \ln \pi_t(k). \quad (\text{A32})$$

Integrating Equation (A31) from an initial time,  $t_i$ , to a final time,  $t_f$ , yields

$$\Delta S_{\text{na}} = - [D(p_{t_f} \| \pi_{t_f}) - D(p_{t_i} \| \pi_{t_i})] + \int_{t_i}^{t_f} \dot{S}_{\text{ex}} dt = -\Delta R + \Delta S_{\text{ex}}, \quad (\text{A33})$$

with  $\Delta R = R_{t_f} - R_{t_i}$  (see Equation (6) from the main text) and  $\Delta S_{\text{ex}} = \int_{t_i}^{t_f} \dot{S}_{\text{ex}} dt$ .

To obtain a form suitable for the analysis of discrete-time data, we approximate the integral of  $\dot{S}_{\text{ex}}$  using a sum over discrete time steps. Let the time interval be divided into steps of size  $\Delta t = 1$  (without a loss in generality). For a step from  $t$  to  $t + 1$ , we approximate the excess entropy production using the functional introduced by Hatano and Sasa [30], which, in our notations, reads as  $\phi_t(k) = -\ln \pi_t(k)$ . For a discrete-time trajectory  $p_t \rightarrow p_{t+1}$ , we define the excess functional

$$\hat{Y}_{t \rightarrow t+1} = \sum_k p_t(k) [\phi_{t+1}(k) - \phi_t(k)] = \sum_k p_t(k) \ln \frac{\pi_t(k)}{\pi_{t+1}(k)}, \quad (\text{A34})$$

which satisfies  $\langle e^{-\hat{Y}} \rangle = 1$  and, therefore,  $\langle \hat{Y} \rangle \geq 0$ .

To see the connection, note that, in the continuous limit,

$$\ln \frac{\pi_t(k)}{\pi_{t+1}(k)} \approx -\partial_t \ln \pi_t(k) \Delta t,$$

so that

$$\hat{Y}_{t \rightarrow t+1} \approx \Delta t \dot{S}_{\text{ex}}(t).$$

Summing over all steps from  $t_i$  to  $t_f - 1$  gives

$$\sum_{t=t_i}^{t_f-1} \hat{Y}_{t \rightarrow t+1} \approx \int_{t_i}^{t_f} \dot{S}_{\text{ex}} dt = \Delta S_{\text{ex}}. \quad (\text{A35})$$

Substituting this discrete approximation into Equation (A33) yields

$$\Delta S_{\text{na}} \approx -\Delta R + \sum_{t=t_i}^{t_f-1} \hat{Y}_{t \rightarrow t+1} > 0. \quad (\text{A36})$$

The adiabatic (housekeeping) entropy production is then the remainder:

$$\Delta S_{\text{a}} = \Delta S_{\text{tot}} - \Delta S_{\text{na}} \geq 0. \quad (\text{A37})$$

## References

- Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [\[CrossRef\]](#)
- Jaynes, E. Information Theory and Statistical Mechanics I. *Phys. Rev.* **1957**, *106*, 620–630. [\[CrossRef\]](#)
- Goldt, S.; Seifert, U. Stochastic Thermodynamics of Learning. *Phys. Rev. Lett.* **2017**, *118*, 010601. [\[CrossRef\]](#) [\[PubMed\]](#)
- Esposito, M.; Broeck, C.V.D. Second law and Landauer principle far from equilibrium. *EPL Europhys. Lett.* **2011**, *95*, 40004. [\[CrossRef\]](#)
- Parrondo, J.M.R.; Horowitz, J.M.; Sagawa, T. Thermodynamics of information. *Nat. Phys.* **2015**, *11*, 131–139. [\[CrossRef\]](#)
- Bera, M.N.; Winter, A.; Lewenstein, M. Thermodynamics from Information. In *Thermodynamics in the Quantum Regime*; Binder, F., Correa, L.A., Gogolin, C., Anders, J., Adesso, G., Eds.; Fundamental Theories of Physics; Springer International Publishing: Cham, Switzerland, 2018; Volume 195, pp. 799–820. [\[CrossRef\]](#)
- Paglietti, A. Why Thermodynamic Entropy and Statistical Entropy are Two Different Physical Quantities. *Curr. Phys. Chem.* **2023**, *13*, 233–245. [\[CrossRef\]](#)
- Wallace, D. *Philosophy of Physics: A Very Short Introduction*; Oxford University Press: Oxford, UK, 2021.
- Caticha, A. Entropy, Information, and the Updating of Probabilities. *Entropy* **2021**, *23*, 895. [\[CrossRef\]](#)
- Caticha, A. Lectures on Probability, Entropy, and Statistical Physics. *arXiv* **2007**, arXiv:0808.0012.
- Caticha, A.; Knuth, K.H.; Caticha, A.; Center, J.L.; Giffin, A.; Rodriguez, C.C. Information and Entropy. *AIP Conf. Proc.* **2007**, *954*, 11–22. [\[CrossRef\]](#)
- Gao, X.; Gallicchio, E.; Roitberg, A.E. The generalized Boltzmann distribution is the only distribution in which the Gibbs-Shannon entropy equals the thermodynamic entropy. *J. Chem. Phys.* **2019**, *151*, 034113. [\[CrossRef\]](#) [\[PubMed\]](#)

13. Natal, J.; Ávila, I.; Tsukahara, V.B.; Pinheiro, M.; Maciel, C.D. Entropy: From Thermodynamics to Information Processing. *Entropy* **2021**, *23*, 1340. [\[CrossRef\]](#)
14. Landauer, R. Irreversibility and Heat Generation in the Computing Process. *IBM J. Res. Dev.* **1961**, *5*, 183–191. [\[CrossRef\]](#)
15. Esposito, M.; Van den Broeck, C. Three detailed fluctuation theorems. *Phys. Rev. Lett.* **2010**, *104*, 090601. [\[CrossRef\]](#)
16. Rao, R.; Esposito, M. Nonequilibrium Thermodynamics of Chemical Reaction Networks: Wisdom from Stochastic Thermodynamics. *Phys. Rev. X* **2016**, *6*, 041064. [\[CrossRef\]](#)
17. Parrondo, J.M.R. Thermodynamics of Information. *Encycl. Condens. Matter Phys.* **2023**, *4*, 522–535. [\[CrossRef\]](#)
18. Seifert, U. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* **2012**, *75*, 126001. [\[CrossRef\]](#)
19. Crooks, G.E. Measuring Thermodynamic Length. *Phys. Rev. Lett.* **2007**, *99*, 100602. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Brillouin, L. *Science and Information Theory*; Academic Press: New York, NY, USA, 1956.
21. Baek, S.K.; Bernhardsson, S.; Minnhagen, P. Zipf’s law unzipped. *New J. Phys.* **2011**, *13*, 043004. [\[CrossRef\]](#)
22. Martini, A.; Cardillo, A.; Rios, P.D.L. Entropic selection of concepts unveils hidden topics in documents corpora. *arXiv* **2018**, arXiv:1705.06510 [\[CrossRef\]](#)
23. Chumachenko, A.; Kreminskyi, B.; Mosenkis, I.; Yakimenko, A. Dynamical entropic analysis of scientific concepts. *J. Inf. Sci.* **2022**, *48*, 561–569. [\[CrossRef\]](#)
24. Hernando, A.; Plastino, A. The thermodynamics of urban population flows. *Phys. Rev. E* **2012**, *86*, 066105. [\[CrossRef\]](#)
25. Gabaix, X. Zipf’s law for cities: An explanation. *Q. J. Econ.* **1999**, *114*, 739–767. [\[CrossRef\]](#)
26. Clauset, A.; Shalizi, C.R.; Newman, M.E.J. Power-Law Distributions in Empirical Data. *SIAM Rev.* **2009**, *51*, 661–703. [\[CrossRef\]](#)
27. Visser, M. Zipf’s law, power laws, and maximum entropy. *New J. Phys.* **2013**, *15*, 043021. [\[CrossRef\]](#)
28. Peng, H.K.; Zhang, Y.; Pirolli, P.; Hogg, T. Thermodynamic Principles in Social Collaborations. *arXiv* **2012**, arXiv:1204.3663. [\[CrossRef\]](#)
29. Giardini, G.S.Y.; daCunha, C.R. Thermodynamics of Innovation: A Statistical Mechanics Framework of Social Adoption. *arXiv* **2025**, arXiv:2508.18560 [\[CrossRef\]](#)
30. Hatano, T.; ichi Sasa, S. Steady-State Thermodynamics of Langevin Systems. *Phys. Rev. Lett.* **2001**, *86*, 3463–3466. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Prokofyev, R.; Demartini, G.; Boyarsky, A.; Ruchayskiy, O.; Cudré-Mauroux, P. Ontology-Based Word Sense Disambiguation for Scientific Literature. In *Advances in Information Retrieval*; Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J.M., Mattern, F., Mitchell, J.C., Naor, M., Nierstrasz, O., Pandu Rangan, C., Steffen, B., et al., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2013; Volume 7814, pp. 594–605. [\[CrossRef\]](#)
32. Martini, A.; Lutov, A.; Gemmetto, V.; Magalich, A.; Cardillo, A.; Constantin, A.; Palchykov, V.; Khayati, M.; Cudré-Mauroux, P.; Boyarsky, A.; et al. ScienceWISE: Topic Modeling over Scientific Literature Networks. *arXiv* **2016**, arXiv:1612.07636. [\[CrossRef\]](#)
33. Palchykov, V.; Gemmetto, V.; Boyarsky, A.; Garlaschelli, D. Ground truth? Concept-based communities versus the external classification of physics manuscripts. *EPJ Data Sci.* **2016**, *5*, 28. [\[CrossRef\]](#)
34. Brodiuk, S.; Palchykov, V.; Holovatch, Y. Embedding technique and network analysis of scientific innovations emergence in an arXiv-based concept network. *arXiv* **2020**, arXiv:2003.10289. [\[CrossRef\]](#)
35. Palchykov, V.; Krasnytska, M.; Mryglod, O.; Holovatch, Y. Network of Scientific Concepts: Empirical Analysis and Modeling. *Adv. Complex Syst.* **2021**, *24*, 2140001. [\[CrossRef\]](#)
36. Cover, T.; Thomas, J. *Elements of Information Theory*; Wiley: Hoboken, NJ, USA, 2012.
37. Caticha, A.; Mohammad-Djafari, A.; Bercher, J.F.; Bessiere, P. Entropic Inference. *AIP Conf. Proc.* **2011**, *1305*, 20–29. [\[CrossRef\]](#)
38. Strasberg, P.; Esposito, M. Non-Markovianity and Negative Entropy Production Rates. *Phys. Rev. E* **2019**, *99*, 012120. [\[CrossRef\]](#)
39. Gujrati, P.D. Comment on “The generalized Boltzmann distribution is the only distribution in which the Gibbs-Shannon entropy equals the thermodynamic entropy” by X. Gao, E. Gallicchio and A.E. Roitberg [*J. Chem. Phys.* **2019**, *151*, 034113]. *arXiv* **2020**, arXiv:2004.07795.
40. Hill, T. *An Introduction to Statistical Thermodynamics*; Addison-Wesley Series in Chemistry; Dover Publications: Mineola, NY, USA, 1986.
41. Schnakenberg, J. Network theory of microscopic and macroscopic behavior of master-equation systems. *Rev. Mod. Phys.* **1976**, *48*, 571–585. [\[CrossRef\]](#)
42. Csiszár, I. Information-type measures of difference of probability distributions and indirect observation. *Stud. Sci. Math. Hung.* **1967**, *2*, 299–318.
43. Allahverdyan, A.E.; Janzing, D.; Mahler, G. Thermodynamic efficiency of information and heat flow. *J. Stat. Mech. Theory Exp.* **2009**, *2009*, P09011. [\[CrossRef\]](#)
44. Hagedorn, R. Statistical Thermodynamics of Strong Interactions at High Energies. *Suppl. Nuovo C.* **1965**, *3*, 147–186.
45. Frautschi, S.C. Statistical Bootstrap Model of Hadrons. *Phys. Rev. D* **1971**, *3*, 2821–2834. [\[CrossRef\]](#)

46. Chumachenko, A.; Buttlere, B. Quantifying Knowledge Evolution with Thermodynamics: A Data-Driven Study of Scientific Concepts. *Qeios* **2024**. [[CrossRef](#)]
47. de Groot, S.R.; Mazur, P. *Non-Equilibrium Thermodynamics*; North-Holland: Amsterdam, The Netherlands, 1962.
48. Prigogine, I. *Introduction to Thermodynamics of Irreversible Processes*; American Lecture Series Publication; Thomas: Springfield, IL, USA, 1955.
49. Stratonovich, R.L. A new representation for stochastic integrals and equations. *SIAM J. Control* **1966**, *4*, 362–371. [[CrossRef](#)]

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