

PAPER • OPEN ACCESS

## Thermodynamics of exponential Kolmogorov–Nagumo averages

To cite this article: Pablo A Morales *et al* 2023 *New J. Phys.* **25** 073011

View the [article online](#) for updates and enhancements.



You may also like

- [Thermodynamic Investigations and Modeling of  \$\text{Co}\_3\text{O}\_4\$ -Based Conversion Anode Materials](#)  
Nicolas Alexander Mayer, Petronela Gotcu-Freis, Damian Marlon Cupid *et al.*
- [Thermodynamics of Irreversible Processes](#)  
P T Landsberg
- [Thermochemical Investigations on Functional Energy Storage Materials](#)  
Torsten Markus



## PAPER

## Thermodynamics of exponential Kolmogorov–Nagumo averages

Pablo A Morales<sup>1</sup> , Jan Korbel<sup>2,3,\*</sup>  and Fernando E Rosas<sup>4,5,6,7</sup> <sup>1</sup> Research Division, Araya Inc., Tokyo 107-6019, Japan<sup>2</sup> Section for Science of Complex Systems, CeMSIIS, Medical University of Vienna, Spitalgasse, 23, 1090 Vienna, Austria<sup>3</sup> Complexity Science Hub Vienna, Josefstädter Strasse 39, 1080 Vienna, Austria<sup>4</sup> Department of Informatics, University of Sussex, Brighton BN1 9RH, United Kingdom<sup>5</sup> Centre for Psychedelic Research, Department of Brain Science, Imperial College London, London SW7 2DD, United Kingdom<sup>6</sup> Centre for Complexity Science, Imperial College London, London SW7 2AZ, United Kingdom<sup>7</sup> Centre for Eudaimonia and Human Flourishing, University of Oxford, Oxford OX3 9BX, United Kingdom

\* Author to whom any correspondence should be addressed.

E-mail: [jan.korbel@meduniwien.ac.at](mailto:jan.korbel@meduniwien.ac.at)**Keywords:** Kolmogorov–Nagumo average, Rényi entropy, H-theorem, Bregman divergence, fluctuation theorem, multifractals

## RECEIVED

23 February 2023

## REVISED

19 June 2023

## ACCEPTED FOR PUBLICATION

6 July 2023

## PUBLISHED

31 July 2023

Original Content from  
this work may be used  
under the terms of the  
[Creative Commons  
Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/).Any further distribution  
of this work must  
maintain attribution to  
the author(s) and the title  
of the work, journal  
citation and DOI.**Abstract**

This paper investigates generalized thermodynamic relationships in physical systems where relevant macroscopic variables are determined by the exponential Kolmogorov–Nagumo average. We show that while the thermodynamic entropy of such systems is naturally described by Rényi's entropy with parameter  $\gamma$ , an ordinary Boltzmann distribution still describes their statistics under equilibrium thermodynamics. Our results show that systems described by exponential Kolmogorov–Nagumo averages can be interpreted as systems originally in thermal equilibrium with a heat reservoir with inverse temperature  $\beta$  that are suddenly quenched to another heat reservoir with inverse temperature  $\beta' = (1 - \gamma)\beta$ . Furthermore, we show the connection with multifractal thermodynamics. For the non-equilibrium case, we show that the dynamics of systems described by exponential Kolmogorov–Nagumo averages still observe a second law of thermodynamics and the H-theorem. We further discuss the applications of stochastic thermodynamics in those systems—namely, the validity of fluctuation theorems—and the connection with thermodynamic length.

**1. Introduction**

In ancient Greece, three classic types of averages were extensively studied: arithmetic mean  $\frac{1}{n} \sum_i x_i$ , geometric mean  $(\prod_i x_i)^{1/n}$ , and harmonic mean  $n(\sum_i \frac{1}{x_i})^{-1}$ , which played various roles in physics, geometry, and music. These so-called *Pythagorean means* found a natural generalisation via functional analysis and measure theory into the well-known one-parametric class of *Hölder means*  $(\frac{1}{n} \sum_i x_i^p)^{1/p}$ . A different generalization of the notion of average was independently proposed by Kolmogorov [1] and Nagumo [2] in 1930, which took the form of  $f^{-1}(\frac{1}{n} \sum_i f(x_i))$  for any continuous and injective function  $f$ . These *Kolmogorov–Nagumo means*—also known as quasi-arithmetic means or f-means—have triggered numerous theoretical developments by several researchers, including de Finetti [3], Jessen [4], Kitagawa [5], Aczél [6], or Fodor and Roubens [7].

Kolmogorov–Nagumo averages have found applications in many fields, including machine learning [8], random fields [9], or fuzzy sets [10]. A particularly important application of Kolmogorov–Nagumo averages is the introduction of *Rényi entropy* [11], which in turn have found many applications in quantum systems [12–14], strongly coupled or entangled systems [15–17], phase transitions [12, 18, 19], multifractal thermodynamics [20, 21], and time series analysis [22, 23], among others. Another important application of generalized means have been the development of the thermodynamics of complex systems [24]. In addition to concepts such as deformed calculus [25, 26], escort means [27–29], or non-linear dynamics [30, 31], Kolmogorov–Nagumo means served as a natural framework for generalized entropies [32]—with the Rényi entropy being a special case. Furthermore, connections between the maximum entropy principle and Rényi entropy have been established in [17, 33].

In this paper we investigate the thermodynamics of systems whose relevant macroscopic variables are determined by the exponential Kolmogorov–Nagumo average. The pioneering work of Czachor and Naudts [32] on the thermodynamics of exponential Kolmogorov–Nagumo averages discusses aspects of equilibrium thermodynamics, such as a generalized internal energy and its relationship with Tsallis entropy [34]. Bagci and Tirnakli [35] studied a generalized maximum entropy principle for the exponential Kolmogorov–Nagumo averages, while leaving its thermodynamic interpretation unclear. A first thermodynamic interpretation of Rényi entropy was given by Baez [36], who noted that the Rényi entropy of a Boltzmann distribution is equal to the change of Helmholtz free energy divided by the change of temperature. The first goal of this paper is to develop a framework that unifies and extends these previous results, enabling a unified interpretation of all thermodynamic quantities as well as the maximum entropy principle in such scenarios.

A second set of relevant results is related to the Légendre structure of thermodynamics [37]. Scarfone *et al* [38] showed that the Légendre structure of equilibrium thermodynamics remains valid for Kolmogorov–Nagumo averages on the level of macroscopic quantities. Additionally, Wong [39, 40] introduced a generalized Légendre structure that corresponds to the Rényi entropy. A second goal of this paper clarify the thermodynamic interpretation of this generalized Légendre structure. Furthermore, Peng *et al* [41] have recently discussed non-equilibrium thermodynamics with non-extensive quantities. Investigations on non-equilibrium thermodynamics and multifractals related to dielectric breakdown was carried out by Enciso *et al* [42]. We will show that these results are closely related to non-equilibrium stochastic thermodynamics based on exponential Kolmogorov–Nagumo averages, while deriving a non-equilibrium version of the second law of thermodynamics, a generalized H-theorem, and proving the validity of detailed and integrated fluctuation theorem. Finally, several authors have studied generalized statistical mechanics from the point of information geometry [43–46]. In particular, Eguchi *et al* have investigated the geometry of generalized e-geodesic and m-geodesic in terms of Kolmogorov–Nagumo means [47]. We extend their results to the case of a generalized divergence, and calculate relevant quantities such as Fisher–Rao information and thermodynamic length.

The rest of the paper is organized as follows. Section 2 defines the exponential Kolmogorov–Nagumo means and summarizes its main properties. Section 3 then establishes equilibrium thermodynamics based on Kolmogorov–Nagumo averages of both entropy and internal energy. Section 4 describes the application of this framework to multifractal systems. Section 5 is focused to the discussion on non-equilibrium thermodynamics. Section 6 is then focused on thermodynamic length. Finally, section 7 summarizes our main conclusions.

## 2. Exponential Kolmogorov–Nagumo averages

Our line of inquiry focuses on systems governed by constraints that can be expressed in terms of non-linear averages. A natural extension of linear (arithmetic) averages is the *Kolmogorov–Nagumo average* [1, 2]

$$\langle X \rangle_f = f^{-1} \left( \sum_i p_i f(x_i) \right), \quad (1)$$

where  $f$  is a continuous injective function. Without loss of generality, we focus on cases where  $f$  is an increasing function. Note that the average is invariant to affine transformations of the function  $f(x) \mapsto f_{a,b}(x) = af(x) + b$ , i.e.  $\langle X \rangle_f = \langle X \rangle_{f_{a,b}}$ .

The linear average is recovered by setting  $f(x) = x$ , which can be shown to be the only one that satisfies two properties:

1. *Homogeneity*:  $\langle aX \rangle = a \langle X \rangle$ ,
2. *Translation invariance*:  $\langle X + c \rangle = \langle X \rangle + c$ .

In [48], it is shown that the first condition alone leads to functions of the form  $f(x) = x^p$ , which corresponds to the well-known class of Hölder averages. In contrast, the second property has been shown to lead to the class of so-called *exponential Kolmogorov–Nagumo averages* [49], which corresponds to  $f(x) = \exp_\gamma(x) = (e^{\gamma x} - 1)/\gamma$  and inverse function  $\ln_\gamma(x) = \frac{1}{\gamma} \ln(1 + \gamma x)$ .

The property of translation invariance is of particular interest for statistical mechanics, as it guarantees that thermodynamic relations are independent of the specific value of the ground state energy. Also, the

standard arithmetic mean is recovered for  $\gamma = 0$ . These considerations make us focus on the Kolmogorov–Nagumo averages, which lead to the following type of averages:

$$\langle X \rangle_\gamma = \ln_\gamma \left( \sum_i p_i \exp_\gamma(x_i) \right) = \frac{1}{\gamma} \ln \left( \sum_i p_i e^{\gamma x_i} \right). \tag{2}$$

A key property of the exponential mean is a weaker version of the additivity of the expectation value, which reads

$$\langle X + Y \rangle_\gamma = \langle X \rangle_\gamma + \langle Y \rangle_\gamma \Leftarrow X \perp\!\!\!\perp Y, \tag{3}$$

where  $\perp\!\!\!\perp$  denotes statistical independence. In the general case, it is possible to express the expected value of sums as

$$\langle X + Y \rangle_\gamma = \langle X + \langle Y|X \rangle_\gamma \rangle_\gamma, \tag{4}$$

where  $\langle Y|X \rangle_\gamma$  is the conditional average, i.e.

$$\langle Y|X \rangle_\gamma = \frac{1}{\gamma} \ln \sum_i p_i \exp(\gamma \langle Y|X = x_i \rangle_\gamma). \tag{5}$$

### 2.1. Connection with cumulants

The exponential Kolmogorov–Nagumo averages are closely related to the cumulant-generating functions, which are given by

$$M_\gamma(X) = \ln \langle e^{\gamma X} \rangle = \gamma \cdot \langle X \rangle_\gamma. \tag{6}$$

By considering the well-known Taylor expansion of the cumulant-generating function, one can find that

$$\langle X \rangle_\gamma = \sum_{n=1}^{\infty} \kappa_n(X) \frac{\gamma^{n-1}}{n!}, \tag{7}$$

where  $\kappa_n(X)$  is the  $n$ th cumulant. This shows that the exponential Kolmogorov–Nagumo averages combine all cumulants weighted by the factor  $\gamma^{n-1}/n!$ .

The relationship between cumulants and Kolmogorov–Nagumo averages can be used to provide a complementary view of the properties of the latter. For example, the additivity property of the Kolmogorov–Nagumo average can be understood as a consequence of the additivity of the cumulant generating function for independent random variables.

### 2.2. Connections with large deviation theory

When considering the sum  $S_n = \frac{1}{n} \sum_{i=1}^n X_i$  of  $n$  i.i.d. random variables  $X_1, \dots, X_n$ , large deviation theory [50] states that the probability of observing  $S_n = s$  can be expressed as

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log P(S_n = s) = I(s) \tag{8}$$

with  $I(s)$  corresponding to the so-called rate function. Cramér’s theorem states that the rate function can be obtained from the cumulant-generating function [50], which—via equation (6)—can be expressed in terms of  $\gamma$ -average using  $\gamma = nk$  as follows:

$$I(s) = \sup_k k(s - \langle S_n \rangle_{nk}) \tag{9}$$

where the  $nk$ -average can be expressed as

$$\langle S_n \rangle_{nk} = \frac{1}{nk} \ln \int ds e^{nks} P(S_n = s). \tag{10}$$

A connection between large deviation theory and statistical mechanics can then be drawn using equation (9) by considering  $k = -\beta$  being the inverse temperature and  $X_i = h_i$  being the energy of  $i$ th subsystem. Then, one can find that

$$\langle S_n \rangle_\beta = -\frac{1}{\beta n} \langle e^{-\beta n h_n} \rangle = \Psi(\beta), \tag{11}$$

with  $\Psi = S - \beta U$  being the free entropy (also known as Massieu function), with  $S$  being the thermodynamic entropy and  $U$  being the internal energy. For more details about this relationship, we refer the interested reader to [50].

### 2.3. Rényi entropy as the exponential Kolmogorov–Nagumo average

The Rényi entropy can be naturally formulated in terms of the exponential Kolmogorov–Nagumo average of Hartley information  $\ln 1/p_k$  [51], also known as the Shannon pointwise entropy [52]. To show this, let us consider  $X$  to be a random variable with probability distribution  $p_k$  and calculate the following:

$$\frac{1}{1-\gamma} \left\langle \ln \frac{1}{p_k} \right\rangle_\gamma = \frac{1}{1-\gamma} \ln_\gamma \left( \sum_i p_i e_\gamma(\ln 1/p_i) \right) = \frac{1}{\gamma(1-\gamma)} \ln \sum_i p_i^{1-\gamma} =: R_\gamma(X), \quad (12)$$

where  $R_\gamma(X)$  denotes the Rényi entropy of order  $1-\gamma$ . This result should not be surprising since the Rényi entropy appears in the Campbell coding theorem as the minimal price that one must pay to encode a message, where the price is the exponential function of the message length [53].

Note that the definition of the Rényi entropy that we are using includes a factor  $1/(1-\gamma)$  that is often not considered. We include this factor in this paper as its addition greatly simplifies the calculations presented in the next sections, which in turn will endow it with a clear thermodynamic meaning. Additionally, by including this factor the limit  $\gamma \rightarrow 1$  of  $R_\gamma(X)$  leads to the well-known Burg entropy  $R_1 = -\sum_i \ln p_i$  [54].

Let us finalize this section by noting that the joint Rényi entropy can be decomposed as follows:

$$\begin{aligned} R_\gamma(X, Y) &= \frac{1}{1-\gamma} \left\langle \ln \frac{1}{p_{ij}} \right\rangle_\gamma = \frac{1}{\gamma} \left\langle \ln \frac{1}{p_i} + \left\langle \ln \frac{1}{p_{ji}} \middle| p_i \right\rangle_\gamma \right\rangle_\gamma \\ &= \frac{1}{\gamma(1-\gamma)} \ln \sum_i p_i^{1-\gamma} + \sum_i p_i \frac{1}{\gamma(1-\gamma)} \ln \sum_j p_{ji}^{1-\gamma} \\ &= R_\gamma(X) + R_\gamma(Y|X), \end{aligned} \quad (13)$$

where  $R_\gamma(Y|X)$  is known as the conditional Rényi entropy.

### 3. Equilibrium thermodynamics

Let us now consider a system whose internal energy at state  $i$  is given by  $\epsilon_i$ . Then, the Kolmogorov–Nagumo  $\gamma$ -average energy of the system is calculated as  $U = \langle \epsilon \rangle_\gamma = \ln \sum_i p_i \exp(\epsilon_i)$ . However, as discussed in the previous section, this quantity is not invariant to rescaling by a factor (i.e.  $\epsilon_i \mapsto \lambda \epsilon_i$ ), and furthermore, it does not have properly defined units of energy. Therefore, it will be convenient to focus our analysis on the following rescaled internal energy:

$$U_\gamma^\beta := \frac{1}{\beta} \langle \beta \epsilon \rangle_\gamma = \frac{1}{\beta} \ln_\gamma \left( \sum_i p_i e_\gamma(\beta \epsilon_i) \right), \quad (14)$$

where  $\beta$  is the inverse temperature of the system. Note that the units of  $\beta$  are  $1/\text{Joules}$ , and hence  $\beta \epsilon_i$  is dimensionless, making  $U_\gamma^\beta$  a properly defined mean energy. This type of average energy function has already been considered in previous research, e.g. in [32]. A summary of thermodynamic quantities studied in this work is presented in table 1

Similarly, our analysis will consider the thermodynamic entropy as defined by the Kolmogorov–Nagumo average of  $\ln 1/p_k$ , which gives us Rényi entropy, as discussed in the previous section. Interestingly, for the case  $\gamma = 0$  this formalism recovers the standard definitions of average energy and Boltzmann–Gibbs entropy, while for  $\gamma \neq 1$  it accounts for different scenarios—which we study in the following.

#### 3.1. Maximum entropy principle

Let us now focus on distributions that correspond to a given value of mean energy as given by  $U_\gamma^\beta$ , according to the maximum entropy principle. The distribution  $\pi_i$  that maximizes the Rényi entropy while satisfying a given  $\gamma$ -average level of energy can be found by using the method of Lagrange multipliers on the following Lagrange function:

$$\mathcal{L} = R_\gamma - \alpha_0 \sum_i p_i - \alpha_1 \frac{1}{\beta} \langle \beta \epsilon \rangle_\gamma. \quad (15)$$

A direct calculation shows that  $\pi_i$  is the solution of the following equation:

$$\frac{1}{\gamma} \frac{\pi_i^{-\gamma}}{\sum_k \pi_k^{1-\gamma}} - \alpha_0 - \frac{\alpha_1}{\beta \gamma} \frac{e^{\gamma \beta \epsilon_i}}{\sum_k \pi_k e^{\gamma \beta \epsilon_k}} = 0. \quad (16)$$

**Table 1.** Comparison of the relevant quantities for Shannon’s and Rényi’s framework. Quantities from standard the Shannon framework can be recovered by taking  $\gamma \rightarrow 0$ .

Quantity	Shannon thermodynamics	Rényi thermodynamics
Entropy	$S(p) = -\sum_k p_k \ln p_k$	$R_\gamma(p) = \frac{1}{\gamma(1-\gamma)} \ln \sum_k p_k^{1-\gamma}$
Cross-entropy	$S(p, q) = -\sum_k p_k \log q_k$	$R_\gamma(p, q) = -\frac{1}{\gamma} \ln \sum_k p_k q_k^{-\gamma} - \gamma R_\gamma(q)$
Csiszár divergence	$\mathcal{D}_{\text{KL}}(p  q) = \sum_k p_k \ln(p_k/q_k)$	$\mathfrak{D}_\gamma(p  q) = \frac{1}{\gamma} \ln \sum_k p_k^{\gamma+1} q_k^{-\gamma}$
Bregmann divergence	$\mathcal{D}_{\text{KL}}(p  q) = -S(p) - S(p, q)$	$\mathcal{D}_\gamma(p  q) = -R_\gamma(p) - R_\gamma(p, q)$
Link function	$C(x, y) = \sum_k x_k y_k$	$C_\gamma(x, y) = \frac{1}{\gamma} \ln(1 + \gamma \sum_k x_k y_k)$
Kolmogorov–Nagumo function	$g(x) = x = g^{-1}(x)$	$g(x) = \frac{e^{\gamma x} - 1}{\gamma}, \quad g^{-1}(x) = \frac{1}{\gamma} \ln(1 + \gamma x)$
Internal energy	$U = \sum_k p_k \epsilon_k$	$U_\gamma = \frac{1}{\gamma} \ln \left( \sum_k p_k e^{\gamma \epsilon_k} \right)$
Heat rate	$\dot{Q} = \sum_k \dot{p}_k \epsilon_k$	$\dot{Q}_\gamma^\beta = \frac{1}{\beta \gamma} \frac{\sum_k \dot{p}_k e^{\gamma \beta \epsilon_k}}{\sum_k p_k e^{\gamma \beta \epsilon_k}}$
Work rate	$\dot{W} = \sum_k p_k \dot{\epsilon}_k$	$\dot{W}_\gamma^\beta = \frac{\sum_k p_k \dot{\epsilon}_k e^{\gamma \beta \epsilon_k}}{\sum_k p_k e^{\gamma \beta \epsilon_k}}$
Equilibrium free energy	$F(\pi) = -\frac{1}{\beta} \ln \sum_k e^{-\beta \epsilon_k}$	$F_\gamma(\pi) = \frac{1}{\beta(\gamma-1)} \ln \sum_k e^{(\gamma-1)\beta \epsilon_k}$
Non-equilibrium free energy	$F(p) = F(\pi) + \frac{1}{\beta} \mathcal{D}_{\text{KL}}(p  \pi)$	$F_\gamma(p) = F_\gamma(\pi) + \frac{1}{\beta} \mathfrak{D}_\gamma(p  \pi)$
Entropy production	$\Delta\Sigma = \mathcal{D}_{\text{KL}}(P(x(t))  \tilde{P}(\tilde{x}(t)))$	$\Delta\Sigma_\gamma = \mathfrak{D}_\gamma(P(x(t))  \tilde{P}(\tilde{x}(t)))$

By multiplying the equation by  $\pi_i$  and summing over  $i$ , one obtains that  $\alpha_0 = \frac{1-\alpha_1}{\gamma}$ , which leads to

$$\frac{\pi_i^{-\gamma}}{\sum_k \pi_k^{1-\gamma}} - 1 = \frac{\alpha_1}{\beta} \left( \frac{e^{\gamma \beta \epsilon_i}}{\sum_k \pi_k e^{\gamma \beta \epsilon_k}} - 1 \right). \tag{17}$$

Above, the Lagrange parameter  $\alpha_1$  can be chosen such that one recovers standard thermodynamic relationships. To this end, we identify that  $\alpha_1 = \beta$  (which is the standard relation between the Lagrange multiplier and inverse temperature), which gives us

$$\pi_i = \frac{(\sum_k \pi_k e^{\gamma \beta \epsilon_k})^{1/\gamma}}{(\sum_k \pi_k^{1-\gamma})^{1/\gamma}} \exp(-\beta \epsilon_i) = \frac{\exp(-\beta \epsilon_i)}{Z^\beta} \tag{18}$$

which is just the Boltzmann distribution with inverse temperature  $\beta$ . We can rewrite the previous relationship as

$$\ln \pi_i = -\beta(\epsilon_i - U_\gamma^\beta(\pi)) - (1 - \gamma)R_\gamma(\pi) \tag{19}$$

and using the fact that  $\sum_k \pi_k = 1$  one finds that

$$\begin{aligned} \ln \sum_k e^{-\beta \epsilon_k} &= \frac{1}{\gamma} \ln \sum_k \pi_k^{1-\gamma} - \frac{1}{\gamma} \ln \sum_k \pi_k e^{\gamma \beta \epsilon_k} \\ &= (1 - \gamma)R_\gamma - \beta U_\gamma^\beta \\ &= \Psi_\gamma - \gamma R_\gamma, \end{aligned} \tag{20}$$

where  $\Psi_\gamma = R_\gamma - \beta U_\gamma^\beta$  is the free entropy (also called Massieu function). Finally, one can derive the free energy by plugging in the equilibrium distribution into internal energy and Rényi entropy as

$$F_\gamma^\beta(\pi) = U_\gamma^\beta(\pi) - \frac{1}{\beta} R_\gamma(\pi) = \frac{1}{(\gamma - 1)\beta} \ln \sum_k e^{(\gamma-1)\beta \epsilon_k}. \tag{21}$$

Thus, one finds that  $F_\gamma^\beta(\pi) = F^{\beta(1-\gamma)}(\pi)$ . This result shows that the Kolmogorov–Nagumo average is effectively a rescaling of the free energy from inverse temperature  $\beta$  to  $(1 - \gamma)\beta$ . Note that equation (21) recapitulates the standard relationship between the free energy and the partition function for  $\gamma = 0$ .

These results reveal that, perhaps surprisingly, the obtained equilibrium distribution in equation (18) (obtained via the maximum entropy principle) is Boltzmann, while the thermodynamic quantities at play are

nonetheless different from the case of ordinary thermodynamics based on Shannon entropy and linear averages. In fact, equation (20) implies that the free entropy and the logarithm of the partition function generally are not equal but differ by the term  $-\gamma R_\gamma$ , which vanishes only for  $\gamma = 0$ .

### 3.2. Thermodynamic interpretation

Let us now focus on the thermodynamic interpretation of equilibrium thermodynamics with the exponential Kolmogorov–Nagumo average. We denote the equilibrium versions of thermodynamic potentials by calligraphic symbols, i.e.

$$\mathcal{U}_\gamma^\beta = U_\gamma^\beta(\pi), \tag{22}$$

$$\mathcal{R}_\gamma^\beta = R_\gamma(\pi), \tag{23}$$

$$\mathcal{F}_\gamma^\beta = \mathcal{U}_\gamma^\beta - \frac{1}{\beta} \mathcal{R}_\gamma^\beta. \tag{24}$$

Let us start with equation (23). As already shown in [36], the equilibrium Rényi entropy can be expressed as

$$\begin{aligned} \mathcal{R}_\gamma^\beta &= \frac{1}{\gamma(1-\gamma)} \ln \sum_i \left( \frac{e^{-\beta \epsilon_i}}{Z^\beta} \right)^{1-\gamma} = \frac{1}{\gamma(1-\gamma)} \ln \sum_i e^{-(1-\gamma)\beta} - \frac{1}{\gamma} \ln Z^\beta \\ &= -\frac{\beta}{\gamma} \left( \mathcal{F}^{(1-\gamma)\beta} - \mathcal{F}^\beta \right). \end{aligned} \tag{25}$$

By defining

$$\beta' = (1-\gamma)\beta \quad \Rightarrow \quad \gamma = 1 - \frac{\beta'}{\beta}, \tag{26}$$

we obtain

$$\mathcal{R}_\gamma^\beta = \beta^2 \frac{\mathcal{F}^{\beta'} - \mathcal{F}^\beta}{\beta' - \beta}, \tag{27}$$

which is the  $\beta$  rescaling of the free energy difference. This can be interpreted as the maximum amount of work the system can perform by quenching the system from inverse temperature  $\beta$  to inverse temperature  $\beta'$ . Note that  $\gamma \rightarrow 0$  corresponds to  $\beta' \rightarrow \beta$  and we recover the ordinary relation between entropy and free energy

$$\mathcal{S}^\beta = \beta^2 \left( \frac{\partial \mathcal{F}^\beta}{\partial \beta} \right). \tag{28}$$

where  $\mathcal{S}^\beta$  is the ordinary thermodynamic entropy.

Since  $\mathcal{F}_\gamma^\beta = \mathcal{F}^{(1-\gamma)\beta}$ , the Kolmogorov–Nagumo energy can be expressed as

$$\mathcal{U}_\gamma^\beta = \mathcal{F}_\gamma^\beta + \frac{1}{\beta} \mathcal{R}_\gamma^\beta = \mathcal{F}^{\beta'} + \beta \frac{\mathcal{F}^{\beta'} - \mathcal{F}^\beta}{\beta' - \beta} = \frac{\beta' \mathcal{F}^{\beta'} - \beta \mathcal{F}^\beta}{\beta' - \beta}. \tag{29}$$

By denoting free entropy as  $\Psi^\beta = -\beta \mathcal{F}^\beta = \mathcal{S}^\beta - \beta \mathcal{U}^\beta$  we find that

$$\mathcal{U}_\gamma^\beta = -\frac{\Psi^{\beta'} - \Psi^\beta}{\beta' - \beta}. \tag{30}$$

Again, in the limit  $\gamma \rightarrow 0$  we recover the classic relationship

$$\mathcal{U}^\beta = -\left( \frac{\partial \Psi^\beta}{\partial \beta} \right). \tag{31}$$

## 4. Connection with multifractal thermodynamics

Let us now focus on the connection of Kolmogorov–Nagumo averages with multifractals. The multifractal analysis provides a powerful tool for investigating the self-similarity of complex systems, including physical and chemical systems [55], weather forecast [56] or financial systems [57]. It has been shown that the Rényi entropy plays a crucial role in the theory of multifractals [20]. We will show that the presented framework based on Kolmogorov–Nagumo averages establishes a connection between two distinct multifractal formalisms.

### 4.1. Fundamentals in multifractal theory

Following the standard approach in the theory of multifractals [58], let us consider a physical system whose state space is parceled into distinct regions  $k_i(s)$  indexed by  $i \in \mathcal{I}$ , with the partition depending on a characteristic scale  $s$ . Such partitions can be studied with respect to positional, spatiotemporal, or energetic state space. Consider the probability of observing the system within region  $k_i(s)$ , which is denoted by  $p_i(s)$ . Let us assume that this probability observes a scaling property of the form

$$p_i(s) = \frac{s^{\alpha_i}}{z(s)}, \quad (32)$$

where  $\alpha_i$  is a scaling exponent and  $z(s) = \sum_i s^{\alpha_i}$  is a normalization constant.

For small scales, i.e.  $s \rightarrow 0$ , let us assume that the frequency of the scaling exponent  $\alpha_i$  is given by a continuous probability distribution, whose density  $\rho$  has the form

$$\rho(\alpha, s) d\alpha = c(\alpha) s^{-f(\alpha)} d\alpha. \quad (33)$$

Above,  $c(\alpha)$  is a slowly varying function of  $\alpha$  and  $f(\alpha)$  is the so-called *multifractal spectrum* of the system, being the fractal dimension of subset which scales with exponent  $\alpha$ . This means that the number of sets  $k_i(s)$  that have the scaling exponent  $\alpha_j$  scale as

$$\text{card}\{k_i(s) | p_i(s) = s^\alpha / z(s)\} = N(s) s^{-f(\alpha)} \quad (34)$$

where  $\text{card}$  denotes cardinality (i.e. the number of sets) and  $N(s)$  is the normalization constant.

Systems that satisfy equation (33) are called multifractals, and the scaling exponent of the Rényi entropy, denoted by  $D_\gamma$ , is known as the generalized dimension [58], i.e. [59]

$$\lim_{s \rightarrow 0} \frac{(1-\gamma)R_\gamma(s)}{\ln s} = D_\gamma. \quad (35)$$

### 4.2. Multifractals and Rényi's entropy

A direct calculation shows that the Rényi entropy can be rewritten as

$$R_\gamma(s) = \frac{1}{\gamma(1-\gamma)} \ln \frac{\int s^{(1-\gamma)\alpha} c(\alpha) s^{-f(\alpha)} d\alpha}{\left(\int s^\alpha c(\alpha) s^{-f(\alpha)} d\alpha\right)^{1-\gamma}}. \quad (36)$$

In the limit of  $s \rightarrow 0$ , both integrals can be approximated by the steepest descent approximation, i.e. it is possible to find a value  $\alpha_\gamma$  (resp.  $\alpha_1$ ) that maximizes the exponents in the integrals. To this end, we define

$$\alpha_\gamma = \arg \min_{\alpha} \{(1-\gamma)\alpha - f(\alpha)\}, \quad (37)$$

$$\alpha_1 = \arg \min_{\alpha} \{\alpha - f(\alpha)\}, \quad (38)$$

and obtain that, for small  $s$ , the Rényi entropy can be approximated as

$$R_\gamma(s) = \left[ \frac{\alpha_\gamma - \alpha_1}{\gamma} - \frac{f(\alpha_\gamma) - (1-\gamma)f(\alpha_1)}{\gamma(1-\gamma)} \right] \ln s + \mathcal{O}(1), \quad (39)$$

where we omitted the constant term coming from the normalization function  $c(\alpha)$ . By introducing the Legendre transform of  $f(\alpha)$ ,

$$\tau_\gamma = (1-\gamma)\alpha_\gamma - f(\alpha_\gamma), \quad (40)$$



one can express the generalized dimension as

$$R_\gamma = \frac{\tau_\gamma - (1 - \gamma)\tau_1}{\gamma(1 - \gamma)} \ln s + \mathcal{O}(1). \quad (41)$$

As a result, we find that the connection between the multifractal spectrum and generalized dimension can be established as

$$D_\gamma = \frac{\tau_\gamma - (1 - \gamma)\tau_1}{\gamma}. \quad (42)$$

Thus, by calculating Rényi entropy, one can obtain the multifractal spectrum and vice versa.

### 4.3. Turbulence cascades

Let us now focus on an alternative approach to multifractal based on turbulence cascades on the scaling of the energy field [60]. The fundamental assumption at the base of this approach is that the average energy  $\epsilon(s)$  scales as  $\langle \epsilon(s)^\gamma \rangle \sim s^{M_\gamma(\epsilon)}$ , which can be formally written as

$$\lim_{s \rightarrow 0} \frac{\ln \langle \epsilon(s)^\gamma \rangle}{\ln s} = M_\gamma(\epsilon) = \gamma \langle \epsilon \rangle_\gamma = \gamma U_\gamma. \quad (43)$$

Interestingly, this approach can be connected with the previous one by noting that the distribution obtained from maximization of Rényi entropy with the constraint on the cumulant generating function leads to  $p(\epsilon_i) \equiv \pi_i$ , i.e. Boltzmann distribution (18). By comparing equations (18) and (32), we find the following correspondence:

$$\alpha_i \ln s = -\beta \epsilon_i. \quad (44)$$

Therefore,  $\epsilon_i = \alpha_i$  relates the energy to the characteristic scaling exponent, and  $\beta = -\ln s$  connects the inverse temperature with the characteristic scale. Furthermore, the parameter  $\gamma$  plays the role of the rescaling of the characteristic scale, where  $\beta' = (1 - \gamma)\beta$  leads to

$$s' = s^{1-\gamma}. \quad (45)$$

This implies that using the  $\gamma$ -exponential Kolmogorov averages implies a change of the characteristic scale  $s \mapsto s^{1-\gamma}$ . Finally, we obtained that

$$\beta U_\beta^\gamma(s) = \frac{\tau_\gamma - \tau_1}{\gamma} \ln s + \mathcal{O}(1), \quad (46)$$

$$\beta F_\beta^\gamma(s) = \frac{\tau_\gamma}{\gamma - 1} \ln s + \mathcal{O}(1), \quad (47)$$

$$M_\gamma = \gamma U_\gamma^\beta(s) = \tau_1 - \tau_\gamma + o(1). \quad (48)$$

We note that the relation between cumulant generating function has been described in [60].

These results show that the connection between the two formalisms naturally leads to thermodynamics with exponential Kolmogorov–Nagumo averages. Our results imply that, for the case of multifractal systems, the maximization of Rényi entropy under the constraint of Kolmogorov–Nagumo average naturally gives us the Boltzmann distribution, where the energy can be translated to the scaling exponent of the equilibrium distribution.

## 5. Non-equilibrium thermodynamics

In this section we investigate the thermodynamics of non-equilibrium systems subject to constraints in the form of Kolmogorov–Nagumo averages. We focus on their Legendre structure,  $H$ -theorem, entropy production, and detailed and integrated fluctuation theorems.

### 5.1. Légendre structure

The Légendre transform establishes several key relations in thermodynamics, connecting the internal energy, entropy, temperature, and Helmholtz free energy—which is defined as  $F = U - TS$ . It also establishes a natural link between extensive variables (e.g. energy, entropy) with intensive variables (e.g. temperature). Interestingly, while the Légendre transform arises naturally in the classic Boltzmann–Gibbs framework, it has been shown that a more general Légendre structure can still be applied to more general scenarios [37] in the context of thermodynamics and [61] in information geometry.

Following this line of reasoning, we now derive thermodynamic relationships that lead directly to a generalized Légendre structure. The Légendre structure for the case of Kolmogorov–Nagumo averages has also been investigated by Scarfone *et al* [38]. Let us start by calculating the change of Helmholtz free energy from the equilibrium distribution to an arbitrary state. Using equation (21), this can be expressed as

$$F_\gamma(p) - F_\gamma(\pi) = \Delta U_\gamma^\beta - \frac{1}{\beta} \Delta R_\gamma. \quad (49)$$

The difference between non-equilibrium and equilibrium internal energy can be expressed as

$$\begin{aligned} U_\gamma^\beta(p) - U_\gamma^\beta(\pi) &= \frac{1}{\beta\gamma} \ln \frac{\sum_i p_i e^{\gamma\beta_0\epsilon_i}}{\sum_i \pi_i e^{\gamma\beta\epsilon_i}} = \frac{1}{\beta} \gamma \ln \left( 1 + \frac{\sum_i (p_i - \pi_i) \pi_i^{-\gamma}}{\sum_i \pi_i^{1-\gamma}} \right) \\ &= \frac{1}{\beta\gamma} \ln (1 + \gamma \nabla R_\gamma \cdot (p - \pi)), \end{aligned} \quad (50)$$

where  $\nabla R_\gamma$  is the vector of partial derivatives of the Rényi entropy expressed in the equilibrium state, which can be written as

$$\frac{\partial R_\gamma}{\partial \pi_i} = \frac{1}{\gamma \sum_i (\pi_i)^{1-\gamma} \pi_i^{-\gamma}}. \quad (51)$$

Thus, the free energy difference can be expressed as

$$F_\gamma(p) - F_\gamma(\pi) = \frac{1}{\beta} R_\gamma(\pi) - \frac{1}{\beta} R_\gamma(p) + \frac{1}{\beta\gamma} \ln (1 + \gamma \nabla R_\gamma \cdot (p - \pi)) \equiv \frac{1}{\beta} \mathcal{D}_\gamma(p||\pi), \quad (52)$$

where  $\mathcal{D}_\gamma(p||\pi)$  is the Rényi–Bregmann divergence [39]. Hence, the divergence between non-equilibrium and equilibrium distributions can be re-written as

$$\mathcal{D}_\gamma(p||\pi) = R_\gamma(\pi) - R_\gamma(p) + C(\nabla R_\gamma, (p - \pi)) = R_\gamma(\pi) - R_\gamma(p) + \frac{1}{\gamma} \ln \sum p_i \Pi_i^{(\gamma)}, \quad (53)$$

where  $\Pi_i^{(\gamma)} = \frac{\pi_i^{-\gamma}}{\sum_i \pi_i^{1-\gamma}}$ , and  $C(x, y) = \frac{1}{\gamma} \ln (1 + \gamma x \cdot y)$  is a generalized *link function*, which becomes the ordinary dot product  $x \cdot y$  for  $\gamma \rightarrow 0$ . The fact that  $\mathcal{D}_\gamma(p||\pi) \geq 0$  [40] implies that  $\Delta F_\gamma$  is always positive, and therefore the free energy is minimized by the equilibrium distribution—generalizing the classical result to the cases of non-linear averages. The implications of this result for thermodynamic scenarios are developed in the next section.

To conclude, let us note that while the standard free energy is obtained by a regular Légendre transform on the macroscopic level (i.e. as  $F = U - TS$ ), the more general free energy  $F_\gamma$  is obtained as a generalized Légendre transform on the mesoscopic (i.e. probability) level. This can be seen by comparing equation (21), where free energy difference is obtained as  $\Delta F_\gamma = \Delta U_\gamma^\beta - \beta^{-1} \Delta R_\gamma$ , with equation (52), where the free energy difference is obtained as  $\Delta F = \beta^{-1} [-\Delta R_\gamma + \ln_\gamma(\nabla R_\gamma \cdot \Delta p)]$ .

### 5.2. Second law and H-theorem

Let us now focus on the case of non-equilibrium stochastic thermodynamics of systems that follow constraints given by Kolmogorov–Nagumo averages. Stochastic thermodynamics [62] recently became an important topic of non-equilibrium statistical physics. While there have been several attempts to derive stochastic thermodynamics associated with generalized entropies (see, e.g. [30, 31, 41]), there have been no studies focused on the stochastic thermodynamics of Kolmogorov–Nagumo averages.

Before starting, let us consider a general formula for a time derivative of a  $\gamma$ -exponential average:

$$\frac{d}{dt} \langle Y(t) \rangle_\gamma = \frac{d}{dt} \left( \frac{1}{\gamma} \ln \sum_i p_i(t) e^{\gamma y_i(t)} \right) = \frac{1}{\gamma} \frac{\sum_i \dot{p}_i(t) e^{\gamma y_i(t)} + \sum_i p_i(t) \gamma \dot{y}_i(t) e^{\gamma y_i(t)}}{\sum_i p_i(t) e^{\gamma y_i(t)}}. \quad (54)$$

With this expression at our disposal, let us start investigating non-equilibrium thermodynamics.

We begin by focusing on the second law of thermodynamics. For this, let us use equation (54) to re-state the first law of thermodynamics for the total energy measured as a Kolmogorov–Nagumo average, which reads as follows:

$$\frac{dU_\gamma^\beta}{dt} = \dot{Q}_\gamma^\beta + \dot{W}_\gamma^\beta, \tag{55}$$

where the work and heat flow into the system of interest are given by

$$\dot{W}_\gamma^\beta = \frac{\sum_i \dot{\epsilon}_i p_i(t) e^{\gamma\beta\epsilon_i}}{\sum_i p_i(t) e^{\gamma\beta\epsilon_i}} \quad \text{and} \quad \dot{Q}_\gamma^\beta = \frac{1}{\beta\gamma} \frac{\sum_i \dot{p}_i(t) e^{\gamma\beta\epsilon_i}}{\sum_i p_i(t) e^{\gamma\beta\epsilon_i}}. \tag{56}$$

Using these expressions, we can now derive the second law of thermodynamics for the case of Rényi entropy and an arbitrary non-equilibrium process driven by linear Markov dynamics. To this end, let us consider a standard master equation given by

$$\dot{p}_i(t) = \sum_j \left( w_{ij}(t) p_j(t) - w_{ji}(t) p_i(t) \right), \tag{57}$$

where  $\dot{p}_i(t) := \frac{dp_i(t)}{dt}$  is the time derivative of  $p_i(t)$ ,  $w_{ij}$  is the transition rate between states  $i$  and  $j$ . Let us consider a control protocol  $\lambda(t)$  which controls the energy spectrum  $\epsilon_i(t) \equiv \epsilon_i(\lambda(t))$ . Let us also introduce the Boltzmann distribution as  $\pi_i(t) = \frac{1}{Z} \exp(-\beta(\epsilon_i(t)))$ . We assume that the transition rates satisfy detailed balance, i.e.

$$\frac{w_{ij}(t)}{w_{ji}(t)} = \frac{\pi_i(t)}{\pi_j(t)} = e^{\beta(\epsilon_j(t) - \epsilon_i(t))}. \tag{58}$$

The time derivative of entropy and heat flow can then be expressed as

$$\dot{R}_\gamma = \frac{1}{\gamma} \frac{\sum_i (p_i)^{-\gamma} \dot{p}_i}{\sum_i p_i^{1-\gamma}} = \frac{1}{\gamma} \sum_{ij} (w_{ij} p_j - w_{ji} p_i) P_i^{(\gamma)}, \tag{59}$$

$$\beta \dot{Q}_\gamma^\beta = \frac{1}{\gamma} \frac{\sum_i \dot{p}_i e^{\gamma\beta\epsilon_i}}{\sum_i p_i e^{\gamma\beta\epsilon_i}} = \frac{1}{\gamma} \sum_{ij} (w_{ij} p_j - w_{ji} p_i) \Phi_i^{(\gamma)}, \tag{60}$$

where we are using  $P_i^{(\gamma)}(t) = \frac{p_i^{-\gamma}(t)}{\sum_i p_i^{1-\gamma}(t)}$  and  $\Phi_i^{(\gamma)}(t) = \frac{\pi_i^{-\gamma}(t)}{\sum_i p_i(t) \pi_i^{-\gamma}(t)}$  as shorthand notations. In the rest of the text, we will often omit the explicit dependence on time for the sake of clarity. Then, the entropy production rate can be calculated as

$$\begin{aligned} \dot{\Sigma}_\gamma &:= \dot{R}_\gamma - \beta \dot{Q}_\gamma^\beta = \frac{1}{\gamma} \sum_{ij} (w_{ij} p_j - w_{ji} p_i) (P_i^{(\gamma)} - \Phi_i^{(\gamma)}) \\ &= \frac{1}{2\gamma} \sum_{ij} (w_{ij} p_j - w_{ji} p_i) ([P_i^{(\gamma)} - P_j^{(\gamma)}] - [\Phi_i^{(\gamma)} - \Phi_j^{(\gamma)}]). \end{aligned} \tag{61}$$

Thus, it is possible to write the entropy production rate as

$$\dot{\Sigma}_\gamma = \frac{1}{2} \sum_{ij} J_{ij} F_{ij}, \tag{62}$$

where  $J_{ij} := w_{ij} p_j - w_{ji} p_i$  are the thermodynamic fluxes and  $F_{ij} = \sigma_{ij}^{\text{bath}} + \sigma_{ij}^{\text{sys}}$  the thermodynamic forces with

$$\sigma_{ij}^{\text{bath}} := -\frac{1}{\gamma} \left( \Phi_i^{(\gamma)} - \Phi_j^{(\gamma)} \right) = \frac{\Phi_i^{(\gamma)}}{\gamma} \left[ \exp \left( \gamma \ln \frac{w_{ij}}{w_{ji}} \right) - 1 \right], \tag{63}$$

$$\sigma_{ij}^{\text{sys}} := \frac{1}{\gamma} \left( P_i^{(\gamma)} - P_j^{(\gamma)} \right) = \frac{P_j^{(\gamma)}}{\gamma} \left[ \exp \left( \gamma \ln \frac{p_j}{p_i} \right) - 1 \right]. \tag{64}$$

For the case of  $\gamma \rightarrow 0$ , the thermodynamic force reduces to the ordinary thermodynamic force  $F_{ij} = \ln \frac{w_{ij} p_j}{w_{ji} p_i}$ . Using this derivation, one can show that the entropy production satisfies the second law of thermodynamics, which in this case, reads as

$$\dot{\Sigma}_\gamma = \dot{R}_\gamma - \beta \dot{Q}_\gamma^\beta \geq 0. \tag{65}$$

Let us now define a generalized  $H$ -function as  $H_\gamma(p) := \mathcal{D}_\gamma(p||\pi) = \beta(F_\gamma(p) - F_\gamma(\pi))$ , where the second equality is based on equation (52). Then, one can show that for the case of pure relaxation, i.e. where  $w_{ij}$  does not depend on time and satisfies the detailed balance

$$\frac{w_{ij}}{w_{ji}} = \frac{\pi_i}{\pi_j} = e^{\beta(\epsilon_j - \epsilon_i)}. \tag{66}$$

$H_\gamma$  fulfills a generalized H-theorem given by the fact that  $H_\gamma(p) \geq 0$  and

$$\dot{H}_\gamma = -\frac{1}{\gamma} \sum_{ij} (w_{ij}p_j - w_{ji}p_i) \left( P_i^{(\gamma)} - \Pi_i^{(\gamma)} \right) \leq 0, \tag{67}$$

with  $\dot{H}_\gamma = 0$  if and only if  $p = \pi$ . This implies that  $H_\gamma(p)$  is a Lyapunov function of the dynamics, i.e. a non-negative quantity that monotonically decreases with time until equilibrium is attained. The proof of both the second law and H-theorem is provided in [appendix](#).

Interestingly, the H-function (i.e. the distance from equilibrium distribution) and the entropy production (i.e. the maximum amount of reversible work) generally differ if  $\gamma \neq 0$ , contrary to the case of the standard Boltzmann–Gibbs framework ( $\gamma = 0$ ) where both are equal to the Kullback–Leibler divergence. The difference between  $\dot{\Sigma}_\gamma$  (equation (65)) and  $\dot{H}_\gamma$  (equation (67)) is in the replacement of  $\Phi_i^{(\gamma)}$  by  $\Pi_i^{(\gamma)}$ , i.e. a different averaging in the denominator. In fact, by combining the first and second laws of thermodynamics, one can find that

$$\begin{aligned} \dot{W}_\gamma^\beta &= \dot{F}_\gamma^\beta + \frac{1}{\beta} \dot{\Sigma}_\gamma^\beta = \frac{1}{\beta} (\dot{H}_\gamma^\beta + \dot{\Sigma}_\gamma^\beta) = \frac{1}{\beta\gamma} \sum_{ij} (w_{ij}p_j - w_{ji}p_i) \left( \Pi_i^{(\gamma)} - \Phi_i^{(\gamma)} \right) \\ &= \frac{1}{\beta\gamma} \sum_{ij} (w_{ij}p_j - w_{ji}p_i) \left( \frac{\pi_i^{-\gamma}}{\sum_i \pi_i \pi_i^{-\gamma}} - \frac{\pi_i^{-\gamma}}{\sum_i p_i \pi_i^{-\gamma}} \right). \end{aligned} \tag{68}$$

To conclude, it is important to note that both equations (65) and (67) reveal a family of inequalities—indexed by  $\gamma$ —that hold for any stochastic system whose dynamics are governed by a master equation (as in equation (57)) and satisfy detailed balance (as in equation (58)). Said differently, our results reveal that any process following a master equation and satisfying detailed balance will satisfy those inequalities for all values of  $\gamma$ . It is then the constraints of the system that determine *which values of  $\gamma$  are physically meaningful*: systems obeying linear constraints yield to a Shannon-type second law (with  $\gamma = 0$ ), while non-linear Kolmogorov–Nagumo constraints lead to a Rényi-type second law (with  $\gamma \neq 0$ ).

### 5.3. Fluctuation theorems

We now show that trajectory thermodynamics remains in the same functional form as in ordinary stochastic thermodynamics. Throughout this section, we will be denoting trajectories and their functionals by bold symbols. We will be denoting probabilities of observing a trajectory by uppercase, bold symbol  $\mathbf{P}$ , which is the probability of observing a state  $x$  at time  $t$  by a lowercase symbol  $p$ .

Let us consider a trajectory  $\mathbf{x} = (x_0, t_0; x_1, t_1, \dots; x_f, t_f)$  starting at  $t_0$  and finishing at  $t_f$ . We denote the trajectory state at time  $t$  as  $\mathbf{x}(t)$ . The trajectory dwells in state  $\mathbf{x}(t) = x_{i-1}$  for  $t_{i-1} \leq t < t_i$  and then jumps to a state  $x_i$  at time  $t_i$ . Let us define  $j(\mathbf{x})$  as the number of trajectory jumps. The probability of observing a trajectory  $\mathbf{x}$  can be obtained from the master equation (57) as

$$\mathbf{P}[\mathbf{x}] = p_{x_0}(t_0) \prod_{j \in j(\mathbf{x})} e^{-\int_{t_{j-1}}^{t_j} w_{x_{j-1}x_{j-1}}(\tau) d\tau} w_{x_jx_{j-1}}(t_j). \tag{69}$$

The energy corresponding to the state  $x$  is controlled by a time-dependent protocol  $\lambda(t)$ , so  $\epsilon_x(t) \equiv \epsilon_x(\lambda(t))$ . Similarly to the previous section, we assume that the transition rates satisfy the detailed balance (58).

Let us now introduce an operation of time reversal  $\tilde{t} = t_f - (t - t_0)$ . The time-reversed trajectory can be defined as  $\tilde{\mathbf{x}}(t) = \mathbf{x}(\tilde{t}) = \mathbf{x}(t_f - (t - t_0))$ . This concept is crucial for understanding the connection between the irreversibility of mesoscopic systems and entropy production. By considering the time-reversed dynamics with time-reversed control protocol  $\tilde{\lambda}(t)$ , the probability of observing the time-reversed trajectory  $\tilde{\mathbf{x}}$  in the time-reversed dynamics determined by the time-reversed protocol  $\tilde{\lambda}$  can be expressed as

$$\tilde{\mathbf{P}}[\tilde{\mathbf{x}}] = p_{x_f}(t_f) \prod_{j \in j(\mathbf{x})} e^{-\int_{t_{j-1}}^{t_j} w_{x_{j-1}x_{j-1}}(\tau) d\tau} w_{x_{j-1}x_j}(t_j), \tag{70}$$

where  $\tilde{P}$  denotes the fact that the probability is calculated in the time-reversed dynamics. By calculating the log ratio of the trajectory probabilities, we get that the waiting probabilities  $e^{-\int_{t_{j-1}}^{t_j} w_{x_{j-1}x_j}(\tau) d\tau}$  cancel out and we end with

$$\ln \frac{P[\mathbf{x}]}{\tilde{P}[\tilde{\mathbf{x}}]} = \ln p(x(t_0)) - \ln p(x(t_f)) + \ln \prod_{j \in j(\mathbf{x})} \frac{w_{x_j x_{j-1}}(t_j)}{w_{x_j x_{j-1}}(t_j)}. \tag{71}$$

Let us now define the entropy of state  $x$  at time  $t$  as the Hartley information of  $p_x(t)$ , i.e.  $s_x(t) = -\log p_x(t)$ . From this definition, it is possible to define a trajectory entropy  $\mathbf{s}[\mathbf{x}](t) \equiv s_{\mathbf{x}(t)}(t) = -\ln p_{\mathbf{x}(t)}(t)$ . By using the condition of detailed balance, we can express the log-ratio as

$$\ln \frac{P[\mathbf{x}]}{\tilde{P}[\tilde{\mathbf{x}}]} = s_{x_f}(t_f) - s_{x_0}(t_0) + \sum_{j \in j(\mathbf{x})} \beta (\epsilon_{x_j}(t_j) - \epsilon_{x_{j-1}}(t_j)). \tag{72}$$

Above, the first difference is equal to the change of the trajectory entropy from the beginning of the trajectory to the end of the trajectory, and is denoted by  $\Delta s[\mathbf{x}] := s_{x_f}(t_f) - s_{x_0}(t_0)$ . Please note that while  $\Delta s$  formally depends on the trajectory, it actually only depends on the starting and ending point (and hence we do not use the bold symbol for it). The second term is equal to the  $\beta$  times the heat exchanged with the reservoir during the trajectory and is denoted by  $\mathbf{q}[\mathbf{x}] := \sum_{j \in j(\mathbf{x})} (\epsilon_{x_j}(t_j) - \epsilon_{x_{j-1}}(t_j))$ . Thus, the log ratio of forward and reversed probabilities is equal to

$$\ln \frac{P[\mathbf{x}]}{\tilde{P}[\tilde{\mathbf{x}}]} = \Delta s[\mathbf{x}] + \beta \mathbf{q}[\mathbf{x}] = \boldsymbol{\sigma}[\mathbf{x}], \tag{73}$$

which is the trajectory entropy production.

To show the relation between trajectory quantities and ensemble quantities, we calculate the time derivative of the entropy production. The time derivative of trajectory entropy can be expressed as

$$\frac{\partial \mathbf{s}[\mathbf{x}](t)}{\partial t} = -\frac{\dot{p}_{\mathbf{x}(t)}}{p_{\mathbf{x}(t)}} - \sum_{j \in j(\mathbf{x})} \delta(t - t_j) \ln \frac{p_{x_j}(t)}{p_{x_{j-1}}(t)}, \tag{74}$$

where the first term is due to the change in the probability distribution, and the second term is due to trajectory jumps. Similarly, the time derivative of trajectory heat is

$$\begin{aligned} \frac{\partial \mathbf{q}[\mathbf{x}](t)}{\partial t} &= \sum_{j \in j(\mathbf{x})} \delta(t - t_j) (\epsilon_{x_j}(t) - \epsilon_{x_{j-1}}(t)) \\ &= \frac{1}{\beta} \sum_{j \in j(\mathbf{x})} \delta(t - t_j) \ln \frac{w_{x_j x_{j-1}}(t)}{w_{x_{j-1} x_j}(t)}. \end{aligned} \tag{75}$$

In both cases, the time derivative depends only on  $x^- = \mathbf{x}(t^-)$  and  $x^+ = \mathbf{x}(t^+)$ . By introducing

$$\dot{s}_{x^- x^+} = \log \frac{p_{x^+}(t)}{p_{x^-}(t)}, \tag{76}$$

$$\dot{q}_{x^- x^+} = \frac{1}{\beta} \log \frac{w_{x^+ x^-}(t)}{w_{x^- x^+}(t)}, \tag{77}$$

the ensemble entropy production rate (equation (62)) can be expressed as

$$\dot{\Sigma}_\gamma = \frac{1}{2} \sum_{x^- x^+} J_{x^- x^+} \left( \Phi_{x^-}^{(\gamma)} \frac{\exp(\gamma \beta \dot{q}_{x^- x^+}) - 1}{\gamma} + P_{x^+}^{(\gamma)} \frac{\exp(\gamma \dot{s}_{x^- x^+}) - 1}{\gamma} \right). \tag{78}$$

In this case, the relation between trajectory quantities and ensemble quantities is not so straightforward due to more complicated averaging.

Let us now focus on another aspect of the entropy production, i.e. the measure of irreversibility. We define a Kolmogorov–Nagumo average of all trajectories for a functional  $\mathbf{G}[\mathbf{x}]$  defined for each probability as

$$\langle\langle \mathbf{G} \rangle\rangle_\gamma = \ln_\gamma \int \mathcal{D}\mathbf{x} \mathbf{P}[\mathbf{x}] \exp_\gamma(\mathbf{G}[\mathbf{x}]), \tag{79}$$

where  $\mathcal{D}\mathbf{x}$  is the path integral measure. Using this, one can find the ensemble entropy production as given by

$$\begin{aligned} \langle\langle \sigma \rangle\rangle_\gamma &= \frac{1}{\gamma} \ln \int \mathcal{D}\mathbf{x} \mathbf{P}[\mathbf{x}] \exp \left( \gamma \ln \frac{\mathbf{P}[\mathbf{x}]}{\tilde{\mathbf{P}}[\tilde{\mathbf{x}}]} \right) = \frac{1}{\gamma} \ln \int \mathcal{D}\mathbf{x} (\mathbf{P}[\mathbf{x}])^{\gamma+1} (\tilde{\mathbf{P}}[\tilde{\mathbf{x}}])^{-\gamma} \\ &= \mathfrak{D}_\gamma(\mathbf{P}[\mathbf{x}] || \tilde{\mathbf{P}}[\tilde{\mathbf{x}}]) \geq 0, \end{aligned} \tag{80}$$

where  $\mathfrak{D}_\gamma(p||q) = \frac{1}{\gamma} \ln \int dx p(x)^{\gamma+1} q(x)^{-\gamma}$  is the Rényi–Csiszár divergence.

Furthermore, one can use equation (73) to show the validity of a detailed fluctuation theorem that holds in the common form [63]. Let us define the probability of observing trajectory entropy production as

$$P(\sigma) = \int \mathcal{D}\mathbf{x} \mathbf{P}[\mathbf{x}] \delta(\sigma - \sigma_x) \tag{81}$$

i.e. we sum over all trajectories that result in entropy production equal to  $\sigma$ . By a simple manipulation, we obtain

$$\begin{aligned} P(\sigma) &= \int \mathcal{D}\mathbf{x} \mathbf{P}[\mathbf{x}] \delta \left( \sigma - \ln \frac{\mathbf{P}[\mathbf{x}]}{\tilde{\mathbf{P}}[\tilde{\mathbf{x}}]} \right) = e^\sigma \int \mathcal{D}\mathbf{x} \tilde{\mathbf{P}}[\tilde{\mathbf{x}}] \delta \left( -\Delta\sigma - \ln \frac{\tilde{\mathbf{P}}[\tilde{\mathbf{x}}]}{\mathbf{P}[\mathbf{x}]} \right) \\ &= e^{\Delta\sigma} \tilde{P}(-\Delta\sigma). \end{aligned} \tag{82}$$

Here  $\tilde{P}$  again denotes that probability is calculated for the time-reversed dynamics. As a result, we obtain the detailed fluctuation theorem

$$\frac{P(\sigma)}{\tilde{P}(-\sigma)} = e^\sigma, \tag{83}$$

Therefore, on the trajectory level, the relations remain exactly the same as in the case of the ordinary Shannon–Boltzmann–Gibbs framework. Finally, the integrated fluctuation theorem can then be formulated as

$$\int d\sigma P(\sigma) e^\sigma = 1, \tag{84}$$

where the integral takes place over the values of trajectory entropy production. An application of Jensen’s inequality yields  $\int d\sigma P(\sigma) \sigma \geq 0$ .

### 6. Thermodynamic length

Thermodynamic length is a well-known metric that characterizes the distance between thermodynamic states. More specifically, this metric is related to the dissipation in a thermodynamic system due to finite-time transformations [64, 65], and has important connections with the Jensen–Shannon divergence [66], Fisher information and Rao’s entropy differential metric [67]. Therefore, thermodynamic length is of great interest for out-of-equilibrium analyses.

Let us consider a collection of thermodynamic states that can be parametrized by  $(\theta_1, \dots, \theta_n) \in \Theta$ . Then, the thermodynamic length of a path  $\mathbf{s}(t) = (s_1(t), \dots, s_n(t)) : [0, \tau] \rightarrow \Theta$  can be calculated as

$$\mathcal{L} := \int_0^\tau \sqrt{\left( \frac{d\mathbf{s}}{dt} \right)^2} dt = \int_0^\tau \sqrt{\sum_{ij} g_{ij} \dot{s}_i(t) \dot{s}_j(t)} dt, \tag{85}$$

$\dot{s}_i(t) := \frac{ds_i(t)}{dt}$  is the time derivative of  $s_i(t)$  where  $g_{ij}$  is a metric tensor corresponding to the well-known Fisher metric [43], which is given by

$$g_{ij}(q) = - \left. \frac{\partial^2 \mathcal{D}(p||q)}{\partial p_i \partial q_j} \right|_{p=q}. \tag{86}$$

Above,  $\mathcal{D}(p||q)$  corresponds to a divergence, a central quantity in information geometry from which all other geometrical properties—including the metric tensor (86) and connections—can be derived. It is worth noting that there are two fundamental types of divergences:

(1) Csiszár divergences, of the form:

$$\mathcal{D}_{f,g}(p||q) = g\left(\sum_i q_i f\left(\frac{p_i}{q_i}\right)\right).$$

(2) Bregmann divergence, of the form:

$$\mathcal{D}_{f,g}(p||q) = S_{f,g}(p) - S_{f,g}(q) - g(\langle \nabla S_{f,g}(q), p - q \rangle).$$

Both families of divergences are closely related to a trace-class generalized entropy  $S_{f,g}(p) = g(\sum_i f(p_i))$  [68], where  $f$  is an increasing function and  $g$  is a concave function. For the case of  $g(x) = x$  and  $f(x) = x \ln x$ , both divergences reduce to the well-known Kullback–Leibler divergence. This shows that the Fisher–Rao metric corresponding to the Csiszár divergence is equivalent to the Fisher metric corresponding to the Kullback–Leibler divergence [45]. Furthermore, as elaborated in equation (52) (see also [44]), the Bregmann divergence corresponds to the difference of free energies—i.e. the amount of reversible work between two states. These facts motivate us to focus on the Rényi–Bregman divergence.

To recapitulate, equation (52) shows that the Rényi–Bregmann divergence can be expressed as

$$\mathcal{D}_\gamma(p||q) = -R_\gamma(p) - R_\gamma(p, q), \tag{87}$$

where  $R_\gamma(p, q)$  is the Rényi–Bregmann cross entropy given by

$$R_\gamma(p, q) = -\frac{1}{\gamma} \ln \sum_i p_i q_i^{-\gamma} - \gamma R_\gamma(q). \tag{88}$$

This generalizes the standard relationship between Kullback–Leibler divergence, Shannon entropy, and cross-entropy  $\mathcal{D}_0(p||q) = -H(p) - H_{\text{cross}}(p, q)$ , which corresponds to the case of  $\gamma = 0$ . By leveraging the structure of the Rényi–Bregmann divergence, one can find the following expression for the metric tensor valid for arbitrary  $\gamma$ :

$$g_{ij}^\gamma(q) = Q_i^{(\gamma)} Q_j^{(\gamma)} + \frac{Q_i^{(\gamma)}}{q_i} \delta_{ij} \tag{89}$$

where  $Q_i^{(\gamma)} = \frac{q_i^{-\gamma}}{\sum_i q_i^{1-\gamma}}$  has been adopted for brevity. Using this expression, one can then evaluate  $ds^2$  as follows:

$$ds^2 = \sum_i \frac{Q_i^{(\gamma)}}{q_i} (dq_i)^2 + \sum_{ij} Q_i^{(\gamma)} Q_j^{(\gamma)} dq_i dq_j. \tag{90}$$

## 7. Conclusions

This paper presents a description of the thermodynamics of systems that follow non-linear constraints in the form of Kolmogorov–Nagumo averages. Our results provide a first step towards a deeper understanding of the thermodynamics of such systems, opening the door to the analysis of systems with long-range correlations and/or multifractal properties, which are naturally described by the Rényi entropy [17, 20]. Furthermore, recent applications of non-equilibrium thermodynamics based on Rényi entropy (and consequently exponential Kolmogorov–Nagumo averages) to utility theory have been investigated [69], opening the possibility to apply the presented formalism in the context of game theory.

The thermodynamics of Kolmogorov–Nagumo averages was found to be naturally centered around the notion of Rényi’s entropy and a generalized Legendre transform, which lead to a novel form of free entropy. Our results show that this free energy, in turn, is directly related to the entropy production of the system. The presented framework allows to extend of most thermodynamic relations to these non-linear systems—including the second law of thermodynamics and fluctuation theorems—if their relationships are adequately recast in terms of the deformed Legendre transform, as summarized in table 1.

In the context of current attempts to apply generalized entropies known mainly from information theory in thermodynamics, it is worth mentioning that generalized entropies—such as the Rényi entropy—do not only emerge in the case of systems under non-arithmetic means as constraints but also emerge in the description of systems that obey a non-linear master equation with ordinary arithmetic average [31]. In the context of equilibrium thermodynamics, it has been argued—perhaps surprisingly—that different entropies

and constraints can lead to the same equilibrium distribution [70]. Our results can be seen as extending these ideas for non-equilibrium processes, as they imply that different combinations of particular dynamics (i.e. the precise form of the Fokker–Planck/master equation), detailed balance (identifying stationary distribution with the equilibrium distribution), and energetic constraints can also lead to the same entropic functional. The precise choice of thermodynamic and dynamic relations depends on the particular choice of a physical system that one intends to describe. The investigation of how to better characterize classes of dynamics and constraints that lead to similar phenomena is an important topic that deserves further investigation in future work.

### Data availability statement

No new data were created or analysed in this study.

### Acknowledgments

P A M acknowledges support by JSPS KAKENHI Grant Number 23K168550001. J K acknowledges support by the Austrian Science Fund (FWF) Project No. P 34994. F R was supported by the Fellowship Programme of the Institute of Cultural and Creative Industries of the University of Kent.

### Appendix. Proof of the second law of thermodynamics and H-theorem for systems described by Kolmogorov–Nagumo averages

The time derivative of the entropy and heat rate over temperature can be found to be given by

$$\dot{R}_\gamma = \frac{1}{\gamma} \frac{\sum_i (p_i)^{-\gamma} \dot{p}_i}{\sum_i p_i^{1-\gamma}} = \frac{1}{\gamma} \sum_{ij} (w_{ij} p_j - w_{ji} p_i) \frac{p_i^{-\gamma}}{\sum_k p_k^{1-\gamma}} \quad (\text{A1})$$

$$\beta \dot{Q}_\gamma = \frac{1}{\gamma} \frac{\sum_i \dot{p}_i e^{\gamma \beta \epsilon_i}}{\sum_i p_i e^{\gamma \beta \epsilon_i}} = \frac{1}{\gamma} \sum_{ij} (w_{ij} p_j - w_{ji} p_i) \frac{e^{\gamma \beta \epsilon_i}}{\sum_k p_k e^{\gamma \beta \epsilon_k}}. \quad (\text{A2})$$

Above, the explicit dependence of  $p_i(t)$  on time is omitted for simplicity. The entropy production rate can then be expressed as

$$\begin{aligned} \dot{S}_\gamma &= \frac{1}{\gamma} \sum_{ij} (w_{ij} p_j - w_{ji} p_i) \times \left[ \frac{p_i^{-\gamma}}{\sum_k p_k^{1-\gamma}} - \frac{e^{\gamma \beta \epsilon_i}}{\sum_k p_k e^{\gamma \beta \epsilon_k}} \right] \\ &= \frac{1}{\gamma} \sum_{ij} w_{ji} p_i \left( \frac{\pi_i p_j}{\pi_j p_i} - 1 \right) \times \left[ \frac{p_i^{-\gamma}}{\sum_k p_k^{1-\gamma}} - \frac{\sum_k \pi_k e^{\gamma \beta \epsilon_k}}{\sum_k p_k e^{\gamma \beta \epsilon_k}} \frac{\pi_i^{-\gamma}}{\sum_k \pi_k^{1-\gamma}} \right]. \end{aligned} \quad (\text{A3})$$

Furthermore, by using equation (51) one can find that

$$\frac{\sum_k \pi_k e^{\gamma \beta \epsilon_k}}{\sum_k p_k e^{\gamma \beta \epsilon_k}} \frac{\pi_i^{-\gamma}}{\sum_k \pi_k^{1-\gamma}} = \frac{\pi_i^{-\gamma}}{\sum_k p_k \pi_k^{-\gamma}}. \quad (\text{A4})$$

By denoting  $P_i^{(\gamma)}(t) = \frac{p_i^{-\gamma}(t)}{\sum_i p_i(t) p_i^{-\gamma}(t)}$  and  $\Phi_i^{(\gamma)}(t) = \frac{\pi_i^{-\gamma}(t)}{\sum_i p_i(t) \pi_i^{-\gamma}(t)}$ , we can then find that

$$\dot{S}_\gamma = \frac{1}{\gamma} \sum_{ij} w_{ji} p_i \left( \left[ \frac{\Phi_i^{(\gamma)} P_j^{(\gamma)}}{\Phi_j^{(\gamma)} P_i^{(\gamma)}} \right]^{-1/\gamma} - 1 \right) \left( P_i^{(\gamma)} - \Phi_i^{(\gamma)} \right). \quad (\text{A5})$$

Similarly, an analogous derivation shows that the time derivative of H-function  $\dot{H}_\gamma$  can be expressed as

$$\dot{H}_\gamma = -\frac{1}{\gamma} \sum_{ij} w_{ji} p_i \left( \left[ \frac{\Pi_i^{(\gamma)} P_j^{(\gamma)}}{\Pi_j^{(\gamma)} P_i^{(\gamma)}} \right]^{-1/\gamma} - 1 \right) \left( P_i^{(\gamma)} - \Pi_i^{(\gamma)} \right). \quad (\text{A6})$$



To conclude the derivation, let us consider now, instead, consider a more general expression

$$\dot{Y}_\gamma = \frac{1}{\gamma} \sum_{ij} w_{ji} p_i \left( \left[ \frac{\Upsilon_i^{(\gamma)} P_j^{(\gamma)}}{\Upsilon_j^{(\gamma)} P_i^{(\gamma)}} \right]^{-1/\gamma} - 1 \right) (P_i^{(\gamma)} - \Upsilon_i^{(\gamma)}), \tag{A7}$$

and show that it is non-negative for  $\Upsilon_i^{(\gamma)} = \Pi_i^{(\gamma)}$  and  $\Upsilon_i^{(\gamma)} = \Phi_i^{(\gamma)}(t)$ . By using the inequality  $\gamma(x^{-1/\gamma} - 1) \geq \log(1/x)$  for  $\gamma > 0$ , one can find that

$$\dot{Y}_\gamma \geq \frac{1}{\gamma^2} \sum_{ij} w_{ji} p_i \log \left( \frac{\Upsilon_j^{(\gamma)} P_i^{(\gamma)}}{\Upsilon_i^{(\gamma)} P_j^{(\gamma)}} \right) (P_i^{(\gamma)} - \Upsilon_i^{(\gamma)}) =: \mathcal{Y}. \tag{A8}$$

For both  $\Upsilon_i^{(\gamma)} = \Pi_i^{(\gamma)}$  or  $\Upsilon_i^{(\gamma)} = \Phi_i^{(\gamma)}$  cases one can use the fact that the processes satisfies the detailed balance, i.e.

$$\log \left( \frac{\Phi_i^{(\gamma)} P_j^{(\gamma)}}{\Phi_j^{(\gamma)} P_i^{(\gamma)}} \right) = \frac{1}{\gamma} \log \left( \frac{\pi_j(t) p_i(t)}{\pi_i(t) p_j(t)} \right) = \frac{1}{\gamma} \log \left( \frac{w_{ji}(t) p_i(t)}{w_{ij}(t) p_j(t)} \right) \tag{A9}$$

and

$$\log \left( \frac{\Pi_i^{(\gamma)} P_j^{(\gamma)}}{\Pi_j^{(\gamma)} P_i^{(\gamma)}} \right) = \frac{1}{\gamma} \log \left( \frac{\pi_j p_i(t)}{\pi_i p_j(t)} \right) = \frac{1}{\gamma} \log \left( \frac{w_{ji} p_i(t)}{w_{ij} p_j(t)} \right) \tag{A10}$$

respectively, and express  $\mathcal{Y}$  as

$$\mathcal{Y} = \frac{1}{\gamma} \sum_{ij} w_{ji} p_i \log \left( \frac{w_{ji} p_i}{w_{ij} p_j} \right) (P_i^{(\gamma)} - \Upsilon_i^{(\gamma)}). \tag{A11}$$

Then, by using  $\log(1/x) \geq 1 - x$ , one can obtain the following inequality:

$$\begin{aligned} \mathcal{Y} &\geq \frac{1}{\gamma} \sum_{ij} w_{ji} p_i (P_i^{(\gamma)} - \Upsilon_i^{(\gamma)}) - \frac{1}{\gamma} \sum_{ij} w_{ij} p_j (P_i^{(\gamma)} - \Upsilon_i^{(\gamma)}) \\ &= -\frac{1}{\gamma} \sum_{ij} (w_{ij} p_j - w_{ji} p_i) (P_i^{(\gamma)} - \Upsilon_i^{(\gamma)}), \end{aligned} \tag{A12}$$

which is equal to  $-\dot{Y}_\gamma$ . Thus, one finally finds that  $\dot{Y}_\gamma \geq -\dot{Y}_\gamma$ , which in turn leads to  $\dot{Y}_\gamma \geq 0$ . This proves both the second law ( $\dot{\Sigma}_\gamma \geq 0$ ) and the H-theorem ( $\dot{H}_\gamma \leq 0$ ).

### ORCID iDs

Pablo A Morales  <https://orcid.org/0000-0002-0965-5199>

Jan Korbel  <https://orcid.org/0000-0002-5371-5320>

Fernando E Rosas  <https://orcid.org/0000-0001-7790-6183>

### References

- [1] Kolmogorov A N 1930 Sur la notion de la moyenne *Accad. Naz. Lincei Mem. Cl. Sci. Fis. Mat. Natur. Sez.* **12** 388–91
- [2] Nagumo M 1930 Über eine klasse der mittelwerte *Jpn. J. Math.* **7** 71–79
- [3] De Finetti B 1931 *Sul Concetto di Media* (Istituto italiano degli attuari)
- [4] Jessen B 1931 Über die verallgemeinerung des arithmetischen mittels *Acta Sci. Math.* **4** 108–16
- [5] Kitagawa T 1934 On some class of weighted means *Proc. Phys.-Math. Soc. Japan III* **16** 117–26
- [6] Aczél J 1948 On mean values *Bull. Am. Math. Soc.* **54** 392–400
- [7] Fodor J and Roubens M 1995 On meaningfulness of means *J. Comput. Appl. Math.* **64** 103–15
- [8] Nielsen F and Sun K 2016 Guaranteed bounds on information-theoretic measures of univariate mixtures using piecewise log-sum-exp inequalities *Entropy* **18** 442
- [9] Porcu E, Mateu J and Christakos G 2009 Quasi-arithmetic means of covariance functions with potential applications to space–time data *J. Multivariate Anal.* **100** 1830–44
- [10] Klement E P, Mesiar R and Pap E 1999 Quasi- and pseudo-inverses of monotone functions and the construction of t-norms *Fuzzy Sets Syst.* **104** 3–13

- [11] Rényi A 1976 *Selected Papers of Alfréd Rényi (Selected Papers of Alfréd Rényi vol 2)* (Akadémiai Kiadó)
- [12] Stéphan J-M, Inglis S, Fendley P and Melko R G 2014 Geometric mutual information at classical critical points *Phys. Rev. Lett.* **112** 127204
- [13] Stéphan J-M 2014 Shannon and Rényi mutual information in quantum critical spin chains *Phys. Rev. B* **90** 045424
- [14] Jizba P, Ma Y, Hayes A and Dunningham J A 2016 One-parameter class of uncertainty relations based on entropy power *Phys. Rev. E* **93** 060104
- [15] Dong X 2016 The gravity dual of Rényi entropy *Nat. Commun.* **7** 12472
- [16] Barrella T, Dong X, Hartnoll S A and Martin V L 2013 Holographic entanglement beyond classical gravity *J. High Energy Phys.* **JHE09(2013)109**
- [17] Jizba P and Korbel J 2019 Maximum entropy principle in statistical inference: case for non-Shannonian entropies *Phys. Rev. Lett.* **122** 120601
- [18] Iaconis J, Inglis S, Kallin A B and Melko R G 2013 Detecting classical phase transitions with Rényi mutual information *Phys. Rev. B* **87** 195134
- [19] Zaletel M P, Bardarson J H and Moore J E 2011 Logarithmic terms in entanglement entropies of 2D quantum critical points and Shannon entropies of spin chains *Phys. Rev. Lett.* **107** 020402
- [20] Jizba P and Arimitsu T 2004 The world according to Rényi: thermodynamics of multifractal systems *Ann. Phys., NY* **312** 17–59
- [21] Jizba P and Arimitsu T 2004 Observability of Rényi's entropy *Phys. Rev. E* **69** 026128
- [22] Jizba P, Kleinert H and Shefaat M 2012 Rényi's information transfer between financial time series *Physica A* **391** 2971–89
- [23] Jizba P and Korbel J 2014 Multifractal diffusion entropy analysis: optimal bin width of probability histograms *Physica A* **413** 438–58
- [24] Thurner S, Hanel R and Klimek P 2018 *Introduction to the Theory of Complex Systems* (Oxford University Press)
- [25] Borges E P 2004 A possible deformed algebra and calculus inspired in nonextensive thermostatics *Physica A* **340** 95–101
- [26] Czachor M 2020 Unifying aspects of generalized calculus *Entropy* **22** 1180
- [27] Beck C and Schögl F 1993 *Thermodynamics of Chaotic Systems: An Introduction (Cambridge Nonlinear Science Series)* (Cambridge University Press)
- [28] Abe S 2003 Geometry of escort distributions *Phys. Rev. E* **68** 031101
- [29] Bercher J-F 2011 On escort distributions, q-gaussians and fisher information *AIP Conf. Proc.* **1305** 208–15
- [30] Frank T D 2004 On a nonlinear master equation and the Haken–Kelso–Bunz model *J. Biol. Phys.* **30** 139–59
- [31] Korbel J and Wolpert D H 2021 Stochastic thermodynamics and fluctuation theorems for non-linear systems *New J. Phys.* **23** 033049
- [32] Czachor M and Naudts J 2002 Thermostatistics based on Kolmogorov–Nagumo averages: unifying framework for extensive and nonextensive generalizations *Phys. Lett. A* **298** 369–74
- [33] Morales P A and Rosas F E 2021 Generalization of the maximum entropy principle for curved statistical manifolds *Phys. Rev. Res.* **3** 033216
- [34] Tsallis C 1988 Possible generalization of Boltzmann–Gibbs statistics *J. Stat. Phys.* **52** 479–87
- [35] Bağcı G B and Tirnaklı U 2009 On the way towards a generalized entropy maximization procedure *Phys. Lett. A* **373** 3230–4
- [36] Baez J C 2022 Rényi entropy and free energy *Entropy* **24** 706
- [37] Plastino A and Plastino A R 1997 On the universality of thermodynamics' legendre transform structure *Phys. Lett. A* **226** 257–63
- [38] Scarfone A M, Matsuzoe H and Wada T 2016 Consistency of the structure of legendre transform in thermodynamics with the Kolmogorov–Nagumo average *Phys. Lett. A* **380** 3022–8
- [39] Wong T-K L 2018 Logarithmic divergences from optimal transport and Rényi geometry *Inf. Geom.* **1** 39–78
- [40] Wong T-K L and Yang J 2019 Logarithmic divergences: geometry and interpretation of curvature *Int. Conf. on Geometric Science of Information* (Springer) pp 413–22
- [41] Peng L, Qian H and Hong L 2020 Thermodynamics of Markov processes with nonextensive entropy and free energy *Phys. Rev. E* **101** 022114
- [42] Enciso L, Gun M, Ruiz M S and Razzitte A C 2019 Entropy in multifractal non equilibrium structures of dielectric breakdown *J. Stat. Mech.* **094011**
- [43] Amari S 2016 *Information Geometry and its Applications* vol 194 (Springer)
- [44] Naudts J 2008 Generalised exponential families and associated entropy functions *Entropy* **10** 131–49
- [45] Korbel J, Hanel R and Thurner S 2019 Information geometric duality of  $\phi$ -deformed exponential families *Entropy* **21** 112
- [46] Korbel J, Hanel R and Thurner S 2020 Information geometry of scaling expansions of non-exponentially growing configuration spaces *Eur. Phys. J. Spec. Top.* **229** 787–807
- [47] Eguchi S, Komori O and Ohara A 2016 Information geometry associated with generalized means *Information Geometry and its Applications IV* (Springer) pp 279–95
- [48] Aczél J, Aczél J D, Aczél J and Dhombres J 1989 *Functional Equations in Several Variables* vol 31 (Cambridge University Press)
- [49] Hardy G H, Littlewood J E and Pólya G 1952 *Inequalities* (Cambridge University Press)
- [50] Touchette H 2009 The large deviation approach to statistical mechanics *Phys. Rep.* **478** 1–69
- [51] Hartley R V L 1928 Transmission of information 1 *Bell Syst. Tech. J.* **7** 535–63
- [52] Ince R A A 2017 Measuring multivariate redundant information with pointwise common change in surprisal *Entropy* **19** 318
- [53] Campbell L L 1965 A coding theorem and Rényi's entropy *Inf. Control* **8** 423–9
- [54] Burg J P 1972 The relationship between maximum entropy spectra and maximum likelihood spectra *Geophysics* **37** 375–6
- [55] Mandelbrot B B 1989 Multifractal measures, especially for the geophysicist *Fractals in Geophysics* (Birkhäuser) pp 5–42
- [56] Lovejoy S and Schertzer D 2018 *The Weather and Climate: Emergent Laws and Multifractal Cascades* (Cambridge University Press)
- [57] Calvet L E and Fisher A J 2008 *Multifractal Volatility: Theory, Forecasting and Pricing* (Academic)
- [58] Harte D 2001 *Multifractals: Theory and Applications* (CRC Press)
- [59] Halsey T C, Jensen M H, Kadanoff L P, Procaccia I and Shraiman B I 1986 Fractal measures and their singularities: the characterization of strange sets *Phys. Rev. A* **33** 1141
- [60] Schertzer D, Lovejoy S, Schmitt F, Chigirinskaya Y and Marsan D 1997 Multifractal cascade dynamics and turbulent intermittency *Fractals* **5** 427–71
- [61] Morales P A, Korbel J and Rosas F E 2023 Geometric structures induced by deformations of the legendre transform *Entropy* **25** 678
- [62] Seifert U 2012 Stochastic thermodynamics, fluctuation theorems and molecular machines *Rep. Prog. Phys.* **75** 126001
- [63] Esposito M and Van den Broeck C 2010 Three detailed fluctuation theorems *Phys. Rev. Lett.* **104** 090601

- [64] Salamon P and Berry R S 1983 Thermodynamic length and dissipated availability *Phys. Rev. Lett.* **51** 1127
- [65] Nulton J, Salamon P, Andresen B and Anmin Q 1985 Quasistatic processes as step equilibrations *J. Chem. Phys.* **83** 334–8
- [66] Bennett C H 1976 Efficient estimation of free energy differences from Monte Carlo data *J. Comput. Phys.* **22** 245–68
- [67] Crooks G E 2007 Measuring thermodynamic length *Phys. Rev. Lett.* **99** 100602
- [68] Hanel R, Thurner S and Gell-Mann M 2012 Generalized entropies and logarithms and their duality relations *Proc. Natl Acad. Sci.* **109** 19151–4
- [69] Ducuara A F, Skrzypczyk P, Buscemi F, Sidajaya P and Scarani V 2023 Maxwell's demon walks into wall street: stochastic thermodynamics meets expected utility theory (arXiv:2306.00449 [[cond-mat.stat-mech](#)])
- [70] Korbel J 2021 Calibration invariance of the maxent distribution in the maximum entropy principle *Entropy* **23** 96