Stochastic thermodynamics and fluctuation theorems for non-linear systems

Jan Korbel and David H Wolpert

1 Section for Science of Complex Systems, CeMSIIS, Medical University of Vienna, Spitalgasse 23, 1090 Vienna, Austria
2 Complexity Science Hub Vienna, Josefstädter Strasse 39, 1080 Vienna, Austria
3 Santa Fe Institute, Santa Fe, NM, United States of America
4 Arizona State University, Tempe, AZ, United States of America
* Author to whom any correspondence should be addressed.
E-mail: jan.korbel@meduniwien.ac.at and david.h.wolpert@gmail.com

Keywords: stochastic thermodynamics, non-linear systems, fluctuation theorem, generalized entropies

Abstract
We extend stochastic thermodynamics by relaxing the two assumptions that the Markovian dynamics must be linear and that the equilibrium distribution must be a Boltzmann distribution. We show that if we require the second law to hold when those assumptions are relaxed, then it cannot be formulated in terms of Shannon entropy. However, thermodynamic consistency is salvaged if we reformulate the second law in terms of generalized entropy; our first result is an equation relating the precise form of the non-linear master equation to the precise associated generalized entropy which results in thermodynamic consistency. We then build on this result to extend the usual trajectory-level definitions of thermodynamic quantities that are appropriate even when the two assumptions are relaxed. We end by using these trajectory-level definitions to derive extended versions of the Crooks fluctuation theorem and Jarzynski equality which apply when the two assumptions are relaxed.

Stochastic thermodynamics has uncovered many important results concerning non-equilibrium systems, including various fluctuation theorems [1–5], thermodynamic uncertainty relations [6–9], and speed limits [10, 11]. These results have been applied in a broad range of disciplines, including physics [12], chemistry [13], biophysics [14], active matter [15], quantum thermodynamics [16] and computation [17, 18].

However, there are several assumptions commonly made in stochastic thermodynamics, which restricts its applicability. In this paper we consider relaxing two of those assumptions, to expand the range of scenarios that can be analyzed using stochastic thermodynamics. However, as we prove, in those scenarios the second law would be violated—if the second law were formulated in terms of conventional Shannon entropy. As we show, thermodynamic consistency can be maintained, but only if we use a suitable generalized entropy. In addition, we show that appropriately extended forms of the Jarzynski equality and Crook’s theorem still hold in this broader setting if use such a generalized entropy.

Most of the research on stochastic thermodynamics to date has assumed that the system of interest is connected to one or more heat baths that are much larger than that system of interest (e.g. ‘infinite’ heat baths) [19]. Microreversibility, i.e. ‘local detailed balance’, then requires that the equilibrium distributions associated with each of those baths is a Boltzmann distribution [20]. However, previous research motivates the choice of non-Boltzmann local equilibrium distributions when considering systems at equilibrium, e.g. in connection with finite heat bath thermodynamics [21–23], long-range interactions [24–26] or generalized thermostatistics [27–30]. To date, none of this research has been extended to open, non-equilibrium systems, the domain of interest in this paper. (See [22] for a partial exception, where the starting distribution is defined by a finite heat bath but the subsequent dynamics is closed.) Accordingly, the first assumption we relax is that local detailed balance is defined in terms of a Boltzmann distribution, either due to the infinite size of the baths or (if they are finite) for some other reason.

Another typical assumption in stochastic thermodynamics has been that the probability distribution evolves according to a linear master equation. However, this assumption is not always warranted.
Accordingly, several authors have explored non-linear [31–33] or non-Markovian [34–36] thermodynamics. There has also been some very restricted earlier research investigating the non-linear master equation [37]. In addition, [38] analyzed stochastic thermodynamics with non-linear internal energy and non-additive Tsallis entropy [39], obtaining the $g$-deformed exponential equilibrium distribution. [40] considered a master equation with probability-dependent transition rates and showed that this is equivalent to a certain class of non-linear processes. These special cases notwithstanding, there has been very little research that systematically investigating stochastic thermodynamics of non-linear systems. Accordingly, the second assumption we relax is to allow the Markovian dynamics to be non-linear.

Given these two relaxations of conventional assumptions, in order to impose the second law in a thermodynamically consistent way, we need to consider other entropies besides Shannon entropy. There are many different possible choices for such a generalized class of entropies [39, 41–54]. However, it was recently shown that the class of entropies that fulfills the first three Shore–Johnson axioms [55, 56] (equivalent to a generalization of Shannon–Khinchin axioms [57]) can be written in the form $S = f(\sum g(p_i))$, which is known as the sum-class of entropies [26]. For these reasons, in order to ensure that the second law is satisfied despite our two relaxed assumptions, we generalize beyond Shannon entropy to arbitrary sum-class entropies. We briefly discuss some main aspects of generalized entropies related to this study in appendix A.

The rest of this letter is organized as follows. First, we formulate the three requirements that a thermodynamic system should fulfill, i.e. we introduce our generalized definitions of Markovian dynamics, local detailed balance, and the second law. Second, we derive the relation between the form of entropic functional and the form of non-linear master equation. Finally, we introduce stochastic versions of thermodynamic quantities and derive fluctuation theorems for non-linear systems.

1. Thermodynamic requirements

We consider a system with a finite number of states, labeled $m$, $n$, etc. All quantities are implicitly functions of time $t$, and for any variable $z$ which varies with $t$, we use $\dot{z}$ to indicate $\frac{dz(t)}{dt}$. Before formulating our requirements, we define several terms, all based on an arbitrary time-dependent scalar-valued internal energy function of the states, $\epsilon_m$. At this point, despite the terminology, there is no physical meaning associated with $\epsilon_m$; such meaning will arise through requirements 2 and 3 below.

We write the expected internal energy as

$$U := \sum_m p_m \epsilon_m$$

and we define $\dot{Q} := \sum_m p_m \epsilon_m$ and $\dot{W} := \sum_m p_m \dot{\epsilon}_m$ as the rate of work and as the heat flow, respectively. So the 'first law of thermodynamics' holds $\dot{U} = \dot{Q} + \dot{W}$.

Next, we define the (generalized) information entropy as

$$S := f\left(\sum_m g(p_m)\right),$$

which reduces to ordinary Shannon information entropy $H = -\sum_m p_m \log p_m$, for $f(x) = x$ and $g(x) = -x \log x$. The form of entropy used in equation (2) is known as sum-class [63], a generalization of trace-class [51], which is the special case where $f(x) = x$. Maximizing the generalized entropy subject to a given value of the expected internal energy gives us the equilibrium distribution:

$$\pi_m = (g')^{-1} \left(\frac{\alpha + \beta \epsilon_m}{C_f(\pi)}\right),$$

where $C_f(\pi) = f^\prime(\sum_m g(\pi_m))$. In order to obtain a unique equilibrium distribution, we also impose the following requirements on functions $f$ and $g$:

- $f$ is a strictly increasing function, i.e. $f^\prime > 0$,
- $g$ is a concave function, i.e. $g'' \leq 0$.

Next, we introduce requirements concerning the thermodynamics of a system coupled to one or more infinite heat baths. By infinite heat bath we mean the standard approach of infinite heat capacity of the

---

5 Some authors have considered different forms for the internal energy, e.g. by using escort distributions [58, 59], generalized means [60, 61] or tail constraints [62]. In our analysis, we stick to the common definition of internal energy because of its clear physical interpretation.
bath. Thus, the bath is statistically independent of the system, which leads directly to $S_{\text{TOT}} = S_{\text{Bath}} + S_{\text{Sys}}$. Furthermore, the bath equilibrates very fast. Therefore, we can consider $\Delta S_{\text{Bath}} = \beta \Delta Q$.

These requirements provide us with the promised physical interpretation of the ‘internal energy’, ‘work’, and ‘heat flow’.

**Requirement 1.** Markovian evolution. We assume that the dynamics of our system evolves according to a first-order differential equation, as in conventional stochastic thermodynamics, but allow that equation to be nonlinear:

$$\dot{p}_m = \frac{1}{C(p)} \sum_n \left[ w_{mn} \Omega(p_n) - w_{nm} \Omega(p_m) \right],$$

where $\Omega(p_n)$ is an arbitrary function while $w_{mn}$ is a set of real-valued parameters. In appendix B, we show how this non-linear master equation can be derived from Markov chains with probability-dependent transition probabilities. $C(p)$ is an arbitrary non-zero function depending on the probability distribution. An overview of physical and life-science applications of Markov chains with probability-dependent transition rates can be found e.g. in review [64].

Note that the antisymmetric form of the summand in equation (4) ensures probability conservation: regardless of the precise form of $\Omega$, $\sum_m \dot{p}_m = 0$. We refer to each $w_{nm}$ as a (generalized) transition rate. In general $w_{mn}$ will not have the form of a conventional transition rate. (The precise connection between the matrix $w_{mn}$ and conventional transition rates is discussed in appendix B.)

**Requirement 2.** Local detailed balance. We assume that the system is coupled to a finite set of infinite heat reservoirs, $\mathcal{V}$, with associated temperatures $T^\nu$, $\nu \in \mathcal{V}$. We formalize this assumption in several steps. First, we assume the transition rates are given by a sum over rates for each reservoir, i.e. $w_{mn} = \sum_\nu w_{\nu mn}$. Next, define the local equilibrium of the system for each such reservoir $\nu$ as the distribution that would be the equilibrium of the system if it were only connected to that reservoir, i.e. as the distribution $p_\nu$ with maximal entropy (as defined in equation (2)), subject to the constraint that the expected energy defined in equation (1) equals $\beta\nu = 1/T^\nu$ (up to a reservoir-independent proportionality constant).

In the usual way we can solve for that maximizing distribution, getting $\pi^\nu_m = \langle \nu \rangle \left( (\nu + \beta\nu \epsilon_m) / (\mathcal{C}_{\nu}^\nu(\pi^\nu)) \right)$ where $\mathcal{C}_{\nu}^\nu(\pi^\nu) = \int \nu(\pi^\nu)$. The requirement of local detailed balance means that the local probability flow $J^\nu_{mn} = w^\nu_{mn} \Omega(p_n) - w^\nu_{nm} \Omega(p_m)$, vanishes for the local equilibrium distribution, i.e.

$$\frac{w^\nu_{mn}}{w^\nu_{nm}} = \frac{\Omega(\pi^\nu_m)}{\Omega(\pi^\nu_n)}.$$  

**Requirement 3.** Second law of thermodynamics. We introduce a few more definitions in order to formalize the second law. We define the heat flow from a reservoir $\nu$ as

$$\dot{Q}^\nu := \frac{1}{C(p)} \sum_{mn} \left[ w^\nu_{mn} \Omega(p_n) - w^\nu_{nm} \Omega(p_m) \right] \epsilon_m$$

so that the total heat flow is $\dot{Q} = \sum_\nu \dot{Q}^\nu$. Similarly, we define the entropy flow from reservoir $\nu$ into the system as $\dot{S}_e := \dot{Q}^\nu / T^\nu$ and the overall entropy flow as $\dot{S}_e = \sum_\nu \dot{S}_e^\nu = \sum_\nu \dot{Q}^\nu / T^\nu$. We then define the entropy production as

$$\dot{S}_e := \frac{d\tilde{S}}{dt} - \dot{S}_e,$$

where the entropy has the form given in equation (2). Using these definitions, the generalized second law is the requirement that

$$\dot{S}_e \geq 0,$$

$$\dot{S}_e = 0 \iff J^\nu_{mn} = 0 \quad \forall \nu, n, m.$$

**2. Necessary relationship between nonlinear master equation and entropy**

Our first main result is a relation between the function $\Omega$ specifying the nonlinear Markovian dynamics and the precise generalized entropy which must hold if the three requirements introduced above all hold:

**Theorem 1.** Requirements 1–3 imply that

$$\Omega(p_m) = \exp(-g'(p_m))$$

(9)
Theorem 1 means that, in general, a given nonlinear Markovian dynamics will require that we not use Shannon entropy, if we wish to maintain thermodynamic consistent. Theorem 1 is also the basis for our other results below.

The complete proof of theorem 1 is in appendix C. Here we outline its main steps. First, we calculate the time derivative of generalized information entropy. Second, we decompose that derivative into two terms, which as we show equal the (non-negative) entropy production rate and the entropy flow rate. (The requirement of local detailed balance is invoked to identify the entropy flow rate with the heat rate over the temperature.) From this we deduce the form of the function \( \Omega \) presented in equation (4).

A corollary of theorem 1 arises if we plug it into equation (4), to derive the general form that the master equation must have in order to meet the three requirements:

\[
\dot{p}_m = \frac{1}{C(p)} \sum_{n,m} \left[ \frac{w_{nm}'}{w_{nm}} \exp(-g'(p_n)) - \frac{w_{mn}'}{w_{mn}} \exp(-g'(p_m)) \right].
\]

There are multiple examples in the literature of systems that can be described by a master equation (11) or corresponding Fokker–Planck equation. (See appendix D for derivation of that Fokker–Planck equation) These include chemical reaction networks [68], morphogen gradient formation [69], financial markets and option pricing [70], (\( g(p_n) \propto p_n \)); heavy-ion collisions [71], turbulence and vortex formation [72–74], \( g(p_n) = p_n \log(p_n) - (1 - p_n) \log(1 - p_n) \); or negative-feedback systems [75], \( g(p_n) = p_n \log(p_n) + (1 - p_n) \log(1 - p_n) \).

Another corollary of theorem 1 is that if we know the form of \( \Omega \), then the entropic functional satisfying requirements 1–3 can be written in the following form:

\[
\mathcal{S} = f \left( - \sum_m \int_0^{p_m} dz \log \Omega(z) \right).
\]

Note also by plugging theorem 1 into the requirement of local detailed balance that local detailed balance has the same form as in conventional stochastic thermodynamics:

\[
\frac{w_{nm}'}{w_{mn}'} = \exp \left( - \frac{\epsilon_m - \epsilon_n}{T} \right).
\]

In addition, we can express the entropy production rate as

\[
\dot{S}_i = \frac{1}{2C(p)} \sum_{mn} \left( \frac{w_{nm}'}{w_{nm}} \Omega(p_m) - \frac{w_{mn}'}{w_{mn}} \Omega(p_n) \right) \log \left( \frac{\frac{w_{nm}'}{w_{mn}} \Omega(p_m)}{\frac{w_{mn}'}{w_{nm}} \Omega(p_n)} \right)
\]

and the entropy flow rate as

\[
\dot{S}_e = \frac{1}{2C(p)} \sum_{mn} \left( \frac{w_{nm}'}{w_{mn}} \Omega(p_m) - \frac{w_{mn}'}{w_{nm}} \Omega(p_n) \right) \log \left( \frac{\frac{w_{nm}'}{w_{mn}}}{\frac{w_{mn}'}{w_{nm}}} \right).
\]

### 3. Stochastic thermodynamics for nonlinear master equations

We now introduce the stochastic (trajectory-level) versions of the ensemble thermodynamic quantities investigated above. For the rest of the paper, we assume \( f(x) = x \) (i.e. the trace-class entropy), for simplicity.

To begin, we define the **stochastic entropy** of a stochastic trajectory \( n(\cdot) \) at time \( \tau \in [0,T] \) as

\[
\mathcal{S}(n(\tau)) := \log \left( \frac{1}{\Omega(p_{n(\tau)}(\tau))} \right) = g'(p_{n(\tau)}(\tau))
\]

and stochastic energy as \( \epsilon(n(\tau)) := \epsilon(n) \). The stochastic heat flow and work flow can be obtained from the first law on the trajectory level \( \dot{\epsilon}(\tau) = \dot{q}(\tau) + \dot{w}(\tau) \) where \( \dot{q}(\tau) = \sum_m \delta_{n(\tau)}(\epsilon_m) \) and \( \dot{w}(\tau) = \sum_m \delta_{n(\tau)}(\epsilon_m) \), respectively. Thus, the ensemble first law of thermodynamics can be written as \( \langle \dot{\epsilon} \rangle = \langle \dot{q} \rangle + \langle \dot{w} \rangle \) where \( \langle \cdot \rangle \) denotes the ensemble average, i.e. multiplication of \( p_{n(\tau)} \) and summing over all trajectories \( n(\tau) \).

The average energy can be obtained as the ensemble average of the stochastic energy, i.e. \( U(\tau) = \langle \epsilon(\tau) \rangle \). However, in general such equivalence of averages does not hold for entropy, once one goes beyond the special case of Shannon information entropy, i.e. in general \( \mathcal{S}(\tau) \neq \langle \mathcal{S}(n(\tau)) \rangle \). Nevertheless, the second law of thermodynamics is still valid on the trajectory level. In appendix E, we show that
\( \dot{s}(n(\tau)) = \dot{s}_l(n(\tau)) + \dot{s}_r(n(\tau)) \) where \( \dot{s}_l(n(\tau)) \) is the entropy production rate and \( \dot{s}_r(n(\tau)) \). Due to the local detailed balance, we obtain that \( \dot{s}_r(n(\tau)) = \sum \beta \dot{s}'_l(n(\tau)) \). The ensemble second law (7) can be obtained by averaging over all possible trajectories.

Now consider an external protocol \( \lambda(\tau) \) which drives the energy spectrum, i.e. \( \epsilon_m(\tau) \equiv \epsilon_m(\lambda(\tau)) \). In the usual way define a time-reversed trajectory \( \tilde{n}(\tau) = n(T - \tau) \) and the associated time-reversed driving protocol \( \tilde{\lambda}(\tau) = \lambda(T - \tau) \). Suppose that local detailed balance is fulfilled. Then as we show in appendix F, \begin{equation}
\log \frac{P(n)}{P(\tilde{n})} = \Delta s_l,
\end{equation}
where \( \tilde{P} \) denotes probability under time-reversed protocol \( \tilde{\lambda}(\tau) \). This means that the standard form of the detailed fluctuation theorem [76] must hold, just with our modified definition of entropy production:
\begin{equation}
\frac{P(\Delta s_l)}{P(-\Delta s_l)} = e^{\Delta \psi_l}.
\end{equation}

Next, write the trajectory version of the Massieu function (also called free entropy [77]) as \( \psi = s - \beta e \). The entropy production can be written as
\begin{equation}
\Delta s_l = \beta w + \Delta \psi_l.
\end{equation}

Consider the situation where the initial distribution is equilibrium one. For equilibrium state the stochastic Massieu function is independent of the state since it is equal the Lagrange parameter \( \alpha \), i.e. \( \psi(\pi_m) = s(\pi_m) - \beta \epsilon_m = \alpha = \langle \psi \rangle \). Combining this with equations (18) and (19) gives (a generalized form of) the Crooks fluctuation theorem,
\begin{equation}
\frac{P(w)}{P(-w)} = \exp(\beta w + \Delta \alpha).
\end{equation}
As usual this in turn gives (a generalized form of) the Jarzynski equality
\begin{equation}
\langle \exp(-\beta w) \rangle = \exp(\Delta \alpha).
\end{equation}

Jensen’s inequality then gives
\begin{equation}
W = \langle w \rangle \geq -T \Delta \alpha.
\end{equation}
These results provide a clear physical interpretation of the Lagrange parameter \( \alpha \): for a process that starts and ends in an equilibrium, \( -T \Delta \alpha \) is the minimal amount of work necessary for transition between two equilibrium states.

4. Thermodynamic uncertainty relations and speed limits for non-linear systems

We finally mention that the validity of DFT (18) has further consequences. Consider the situation when initial and final distribution agree \( p_s(0) = p_s(T) \) and when the external protocol is time-symmetric, i.e. \( l(\tau) = l(\tau) \). Let us consider a trajectory quantity that is antisymmetric under time-reversal, i.e. \( \phi(n(\tau)) = -\phi(\tilde{n}(\tau)) \). Then the fluctuation theorem uncertainty relation holds:
\begin{equation}
\frac{\text{Var} (\phi)}{\langle \phi \rangle^2} \geq \frac{2}{e^{\Delta \alpha} - 1},
\end{equation}
as shown in [78].

Next, we mention the time-dissipation uncertainty relation [80]. By defining survival probabilities \( p^s(t) \) as in [80], we can derive an integral fluctuation theorem for \( p^s(t) \) as a consequence of the detailed fluctuation theorem in the main text:
\begin{equation}
p^s(t) = \tilde{p}^s(t) \langle \exp -\Delta s_l \rangle_{\tilde{\psi}},
\end{equation}
where \( \langle X \rangle_{\tilde{\psi}} = \sum_{n(\tau)} X(n(\tau)) \tilde{P}(n(\tau)) / \tilde{p}(t) \) denotes the normalized average of the generic observable \( X \) on the set of survived trajectories. By defining the instantaneous rate \( r(t) := -\frac{1}{p(t)} \frac{dp^s(t)}{dt} \), we recover the main result of [80]
\begin{equation}
\langle \dot{s} \rangle_{\tilde{\psi}}(t) \geq r(t) - \bar{r}(t),
\end{equation}
where \( \dot{s} \) is the entropy production rate for the case of non-linear systems.
Finally, we show that the results in [10] can be straightforwardly generalized to the case of systems governed by the non-linear master equation. In order to do so, we introduce the dynamical activity of the system as

$$A(t) := \sum_{n \neq m} w_{nm} \Omega(p_n)$$  \hspace{1cm} (26)$$

and define the total entropy production of the system as

$$S_{\text{tot}}(\tau) = \int_0^\tau dt \dot{S}(t)$$ where $\dot{S}(t)$ is defined in equation (14). Let us also define $\langle A \rangle_{\tau} = \frac{1}{\tau} \int_0^\tau dt A(t)$ and the variation distance as $
abla(p, p') := \sum_m |p_m - p'_m|$. Then by a straightforward calculation analogous to the proof in [10] we obtain the speed limit in the following form:

$$\frac{L(p(0), p(\tau))^2}{2S_{\text{tot}}(\langle A \rangle_{\tau})} \leq \tau.$$ \hspace{1cm} (27)

5. Discussion

In this paper we considered a broad type of nonlinear Markovian dynamics. We also define a generalized 'local equilibrium distribution', as the distribution that maximizes the 'information entropy', which we allow to be any specified sum-class entropy functional. This allowed us to impose a generalized form of local detailed balance, by requiring that the currents for each reservoir vanish if the system is in its local equilibrium distribution for the specified information entropy. We required that the generalized, nonlinear Markovian dynamics obey this generalized form of local detailed balance. Next, we define 'entropy flow' in the usual way in stochastic thermodynamics, in terms of heat flow into the reservoirs. Given these definitions, we define the (generalized) second law by requiring that under the nonlinear Markovian dynamics, the time-derivative of the difference between the specified information entropy and the entropy flow is non-negative.

Our first result was to show that for this generalized form of the second law to hold, there must be a certain simple relationship between the choices of which particular sum-class entropy to use to define the information entropy, and of the particular nonlinear form of the Markovian dynamics. We then used this result to derive trajectory-level definitions of (generalized) thermodynamic quantities. Our second result is that those generalized stochastic thermodynamic quantities obey both the Jarzynski equality and Crooks’ theorem.

It is important to emphasize that although we broadened the range of scenarios that stochastic thermodynamics can apply to, we still restricted those scenarios in several ways. For example equation (4) is not the most general form of a non-linear master equation. Other types of non-linear equations are obtained from mean-field approximations of many-body systems, phase transitions and self-organization or approximations of quantum master equations, see [67] for an overview. The question at stake is whether the results can be extended to more general types of non-linear Markovian and non-Markovian systems. This might require to consider entropic functionals of more generalized form, or even generalizations of the maximum entropy principle to e.g. principle of maximum caliber [81].

Acknowledgments

DHW thanks the Santa Fe Institute for support and acknowledges support from the US National Science Foundation, Grant No. CHE-1648973 and from the FQXi foundation, Grant No. FQXi-RFP-1622. JK was supported by the Austrian Science Fund (FWF) under projects I 3073 and P 33751 and the Austrian Research Agency (FFG) project number 882184. The opinions expressed in this paper are those of the authors and do not necessarily reflect the view of the National Science Foundation.

Data availability statement

No new data were created or analysed in this study.

Appendix A. Generalized entropies in statistical physics

Originally, the connection between thermodynamics and statistical physics was provided by Shannon entropy. However, in information theory and latter in statistical physics, there arose various generalizations of entropies [39, 41, 42, 44, 45, 50] (see e.g. [52] for an overview). There were also several attempts to
classify generalized entropies based on asymptotic scaling [51, 53, 82], group laws [54, 63] or generalized information axiomatics [57, 83, 84]. From the point of view of statistical inference, the maximum entropy principle is an inference method. Shore and Johnson formulated consistency requirements that the maximum entropy principle should fulfill [55, 56]. After a recent debate on validity of axioms [85–89], it has been shown in [26] that the class of entropies that fulfills first three Shore–Johnson axioms can be written in the form \( S(P) = f(\sum_i g(p_i)) \), which is known and sum class of entropies. As recently shown, Shore–Johnson axioms are also equivalent to a generalization of Shannon–Khinchin axioms [57] where the sum-class of entropies fulfills the so-called decomposition principle. For this reason, we use in our analysis the sum-class of entropies since it provides us with enough generality, yet it has a solid physical background.

Furthermore, let us discuss several aspects of generalized entropies relevant to the analysis in the main text.

*Non-additivity and non-extensivity:* probably the most popular example of generalized entropies is the entropy introduced by C Tsallis which can be defined as:

\[
S_t(P) = \frac{1}{1-q} \left( \sum_i p_i^q - 1 \right)
\]  

(A1)

The main property of Tsallis entropy is the non-additivity. Let us consider a system coupled to a heat bath. We consider two disjoint subsystems that are statistically independent. In other words, the joint system is the Cartesian product of the marginal systems \([A, B]\), which also holds for the respective probabilities. The entropy of the joint system can be expressed in terms of entropies for the subsystems as

\[
S_t(A, B) = S_t(A) + S_t(B) + (1-q)S_t(A)S_t(B).
\]  

(A2)

In general, the entropy of the joint system \(A \otimes B\) can be expressed as

\[
S_t(A \otimes B) = S_t(A) + S_t(B;A) + (1-q)S_t(A)S_t(B;A),
\]  

(A3)

where \(S_t(B;A)\) is the conditional entropy. Closely related, but generally different is the concept of extensivity. Let us consider a system with sample space \(W(N)\) where \(N\) denotes a characteristic size of the system. An entropy is called extensive if \(S(W(N)) \sim N\) in thermodynamic limit, i.e. \(N \rightarrow \infty\). While additivity is a property purely depending on the form of entropy, extensivity combines properties of entropy and the system itself. Therefore, for systems with exponentially growing sample space \(W(N) \sim N^\gamma\), Tsallis entropy is non-extensive. On the other hand, for systems with polynomially growing sample space \(W(N) \sim N^\gamma\), Tsallis entropy is extensive for \(\rho = 1/(1-q)\). Note that there exists some invariance in the sense that several entropies can lead to the same Max-Ent distribution [90]. For example, Tsallis entropy and Rényi entropy

\[
R_q(P) = \frac{1}{1-q} \ln \sum_i p_i^q
\]  

(A4)

are both maximized by the q-exponential distribution. While Tsallis entropy is extensive for systems with polynomial sample-space growth, Rényi entropy is extensive for the exponential sample-space growth. Thus both lead to different thermodynamics.

*Non-linear systems coupled to an infinite heat reservoir:* let us consider the case when a non-linear system is coupled to an infinite heat bath [67]. In this case, the system is driven by a non-linear master equation (e.g. equation (1) from the main text). By considering local detailed balance, we connect the stationary distribution obtained from the infinite relaxation limit to the equilibrium distribution obtained by the maximization of entropy. By considering, e.g. Tsallis entropy, we obtain, similarly to the case of finite heat bath, the q-exponential equilibrium distribution. However, this distribution arises not due statistical correlations between the system and the bath but due to the system’s non-linear nature. Since the bath is infinite and therefore uncorrelated with the system, the total entropy is simply the sum of entropy for the system of interest and the bath (note that the bath is always in equilibrium and therefore changes only due to absorbed heat). For more details, see [67] and the references therein. On the other hand, if we divide the system into subsystems, the entropy of the whole system may not be a sum of entropies of subsystems, due to long-range correlations among the subsystems. This is the case, e.g. for Tsallis entropy.

*Extension of the second law of thermodynamics:* on of the original motivations for introducing generalized entropies was to obtain non-exponential distributions from the maximum entropy principle. However, to obtain thermodynamically consistent entropy à la Clausius that can be used to calculate thermodynamic quantities, it is desirable to extend the validity of the laws of thermodynamics to generalized entropies as well. An extensive discussion about the validity of the second law for the case of generalized entropies has been heated up. A work by Gorban [91] discusses the second law in connection with H-theorem. Its
conclusions are surprisingly similar to the results presented in this paper: for a given dynamics described by a master equation, one needs to find a proper information measure such that H-theorem is fulfilled. This is also in agreement with [67]. Specifically, for the case of Tsallis entropy, several works are showing the validity of the second law for particular classes of systems (see, e.g. [92]). In our study, we do not consider a specific system but rather assume validity of the second law. We believe that this assumption is plausible and fulfilled for a wide class of non-linear systems.

**Extensivity issue for systems out of equilibrium:** let us finally make some note about the extension of the extensivity issue to the case of non-equilibrium systems driven by a master equation. We follow the setup from [93] and consider a system weakly coupled to an infinite heat bath. The system can be decomposed into two disjoint subsystems \( X \otimes Y \). Let us consider that the total system evolves according to the linear master equation for this moment. It can be expressed as

\[
\dot{p}_{xy}(t) = \sum_{x' y'} w_{xy,x'y'} p_{x'y'}(t).
\]

The evolution of the subsystem \( X \) is obtained by coarse-graining over \( y \), i.e.

\[
\dot{p}_x(t) = \sum_{y' y} w_{xy,x'y'} p_{x'y'}(t) = \sum_{x'} W_{xx'} p_x(t),
\]

where \( W_{xx'} := \sum_{y} w_{xy,x'y'} p_{y'x'} \) which is Markovian for the case of so-called 'time-scale separation' (see [93] for details).

The situation becomes even more complicated for the case of non-linear equation (1) from the main text since even in the case of time-scale separation, it might not be possible to find a unique decomposition \( \Omega(p_{xy}) = \Psi(p_{xy})\Omega(p_x) \) for some functions \( \Psi \) and \( \Omega \). This means that while our analysis might be applicable to the whole system, it might not apply to its subsystems. This might not be surprising since the non-linear master/Fokker–Planck equations typically arise from mean-field approximations of many-body systems in the thermodynamic limit. Thus, the rescaling of such systems might not make physical sense or can change the form of the master equation.

**Appendix B. Non-linear Markov chains and non-linear master equation**

In this appendix we show how to derive a non-linear master equation from a Markov chain with probability-dependent transition rates. More details about non-linear Markov chains and their expression in terms of stochastic processes with probability-dependent transition rates can be found e.g. in [64]. Consider a discrete-time stochastic process whose \( n \)-point distribution can be written in a form of non-linear Markov process, i.e.

\[
p(x_{n}; t_n; x_{n-1}, t_{n-1}; \ldots; x_0, t_0) = \prod_{j=1}^n W(x_j, t_j|x_{j-1}, t_{j-1})\Omega(p(x_0, t_0))
\]  

(B1)

Such a process is called 'non-linear Markov chain'. Note that the function \( W(x_j, t_j|x_{j-1}, t_{j-1}) \) cannot be interpreted as transition probability. This can be easily shown e.g. for \( n = 2 \) and summing (B1) over \( x_1 \). We get that

\[
\sum_{x_2} W(x_1, t_1|x_0, t_0) = \frac{p(x_0, t_0)}{\Omega(p(x_0, t_0))}. 
\]  

(B2)

Now consider a two-point distribution

\[
p(x_1, t_1; x_0, t_0) = W(x_1, t_1|x_0, t_0)\Omega(p(x_0, t_0))
\]  

(B3)

Then, the transition probability \( T \) that is defined in the usual way can be expressed as follows:

\[
T(x_1, t_1|x_0, t_0) = \frac{p(x_1, t_1; x_0, t_0)}{p(x_0, t_0)} = \frac{W(x_1, t_1|x_0, t_0)\omega(p(x_0, t_0))}{p(x_0, t_0)},
\]  

(B4)

where \( \omega(z) = \Omega(z)/z \). Thus, the transition probability explicitly depends on the probability distribution \( p(x_0, t_0) \), i.e.

\[
T(x_1, t_1|x_0, t_0) \equiv T(x_1, t_1|x_0, t_0, p(x_0, t_0)).
\]  

(B5)

Next consider a three point distribution

\[
p(x_2, t_2; x_1, t_1; x_0, t_0) = W(x_2, t_2|x_1, t_1)W(x_1, t_1|x_0, t_0)\Omega(p(x_0, t_0)).
\]  

(B6)
This can be rewritten in terms of transition probabilities as
\[
p(x_2, t_2; x_1; t_1; x_0, t_0) = \frac{T(x_2, t_2|x_1, t_1, p(x_1, t_1))T(x_1, t_1|x_0, t_0, p(x_0, t_0))}{\omega(p(x_1; t_1))} p(x_0, t_0). \tag{B7}
\]

By summation over \(x_1\), we obtain the Chapman–Kolmogorov equation for non-linear Markov chains:
\[
T(x_2, t_2|x_0, t_0, p(x_0, t_0)) = \sum_{x_1} T(x_2, t_2|x_1, t_1, p(x_1, t_1))T(x_1, t_1|x_0, t_0, p(x_0, t_0)) \frac{\omega(p(x_1; t_1))}{\omega(p(x_2; t_2))} \tag{B8}
\]

From the self-consistency requirement that
\[
p(x_1, t_1) = \lim_{t_2 \to t_1} T(x_2, t_2|x_1, t_1, p(x_1, t_1))p(x_1, t_1) \tag{B9}
\]
we obtain that \(\lim_{t_2 \to t_1} T(x_2, t_2|x_1, t_1, p(x_1, t_1)) = \delta(x_2 - x_1)\).

For small times, we can approximate the transition probability by the first-order expansion in terms of \(t_2\) around \(t_1\):
\[
T(x_2, t_2|x_1, t_1, p(x_1, t_1)) \approx (t_2 - t_1) \tilde{T}(x_2|x_1, t_1, p(x_1, t_1)) + O((t_2 - t_1)^2). \tag{B10}
\]

By adding a zero term to the right-hand side, we get
\[
T(x_2, t_2|x_1, t_1, p(x_1, t_1)) \approx (t_2 - t_1) \tilde{T}(x_2|x_1, t_1, p(x_1, t_1)) + \left(1 - (t_2 - t_1) \sum_{x'} \tilde{T}(x'|x_1, t_1, p(x_1, t_1))\right) \delta(x_2 - x_1). \tag{B11}
\]

By plugging this into the Chapman–Kolmogorov equation, we obtain
\[
T(x_2, t_2|x_0, t_0, p(x_0, t_0)) \approx (t_2 - t_1) \sum_{x_1} \left( \frac{\tilde{T}(x_2|x_1, t_1, p(x_1, t_1))T(x_1, t_1|x_0, t_0, p(x_0, t_0))}{\omega(p(x_1, t_1))} - \frac{\tilde{T}(x_1|x_2, t_2, p(x_2, t_2))T(x_2, t_2|x_0, t_0, p(x_0, t_0))}{\omega(p(x_2, t_2))}\right) + T(x_2, t_1|x_0, t_0, p(x_0, t_0)) \tag{B12}
\]

Reorder the terms in this equation and take the limit, to establish that
\[
\lim_{t_2 \to t_1} \tilde{T}(x_2, t_2|x_0, t_0, p(x_0, t_0)) - T(x_2, t_1|x_0, t_0, p(x_0, t_0)) = \sum_{x_1} \left( \frac{T(x_2|x_1, t_1, p(x_1, t_1))T(x_1, t_1|x_0, t_0, p(x_0, t_0))}{\omega(p(x_1, t_1))} - \frac{\tilde{T}(x_1|x_2, t_2, p(x_2, t_2))T(x_2, t_2|x_0, t_0, p(x_0, t_0))}{\omega(p(x_2, t_2))}\right). \tag{B13}
\]

As shorthand write \(w(x_2|x_1; t_1) = \tilde{T}(x_2|x_1, p(x_1, t_1))/\omega(p(x_1, t_1))\) multiply both sides of Section B by \(p(x_0, t_0)\), and integrate over \(x_0\). Then we get the final equation
\[
\frac{dp(x; t)}{dr} = \sum_y \left[ w(x|y, t)\Omega(p(y, t)) - w(y|x, t)\Omega(p(x, t)) \right] \tag{B14}
\]
which is the non-linear master equation from the main text. Note that \(w(x|y, t)\) cannot be obtained as the infinitesimal transition probability, but it is rather obtained as
\[
w(x|y, t) = \lim_{\Delta t \to 0} \frac{W(x; t + \Delta t|y, t) - W(x; t|y, t)}{\Delta t}. \tag{B15}
\]

Nevertheless, we still call \(w\) as transition rate since it plays similar role to the case of linear continuous-time Markov chain.
Appendix C. Proof of theorem 1

**Proof.** It will be useful to introduce the notation \( w_{mn}\Omega(p_m) = J_{mn}' \) and \( C_f = f'(\sum_m g(p_m)) \). The time derivative of the entropy is

\[
\dot{S}(t) = C_f \sum_m g'(p_m) \dot{p}_m = C_f \sum_{m,n,\nu} g'(p_m) \frac{1}{C(p)} (J_{mn}' - J_{mn}).
\]  

(C1)

We now divide this expression into a sum of two terms:

\[
\dot{S}(t) = C_f \sum_{m,n,\nu} \frac{1}{C(p)} (J_{mn}' - J_{mn}) \Phi_{mn}^\nu + C_f \sum_{m,n,\nu} \frac{1}{C(p)} (J_{mn}' - J_{mn}) (g'(p_m) - \Phi_{mn}^\nu)
\]

(C2)

where \( \Phi_{mn}^\nu \) is yet undetermined.

Next, we focus on the first term, identifying it as the entropy production rate. First, we rewrite it in the more suggestive form:

\[
\dot{S}_1 = C_f \sum_{m,n,\nu} \frac{1}{C(p)} (J_{mn}' - J_{mn}) (\Phi_{mn}^\nu - \Phi_{mn}^\nu).
\]  

(C3)

According to requirement 3, \( \dot{S}_1 \) is non-negative. Moreover, it is equal to zero only if all currents vanish, i.e. in equilibrium. First, the constant \( C_f \) is positive because \( f \) is increasing. Second, the non-negativity of the sum can be assured by considering non-negativity of each term. From this, we have that \( \Phi_{mn}^\nu < \Phi_{mn}^\nu \) if \( J_{mn}' < J_{mn} \) and vice versa. Thus, we can conclude that \( \Phi_{mn}^\nu \) is a function of \( J_{mn}' \), i.e. \( \Phi_{mn}^\nu = \phi(J_{mn}') \). Again, because we can relabel the probabilities, the function \( \phi \) does not explicitly depend on \( n, m \) or \( \nu \).

Now we focus on the second term, rewriting it in the more suggestive form

\[
\frac{C_f}{2} \sum_{m,n,\nu} \frac{1}{C(p)} (J_{mn}' - J_{mn}) [(g'(p_m) - \Phi_{mn}^\nu) - (g'(p_n) - \Phi_{mn}^\nu)].
\]  

(C4)

If the first term is the entropy production, then the second term is equal to entropy flow rate. Entropy flow is equal to

\[
\dot{S}_e = \frac{1}{2} \sum_{m,n,\nu} (J_{mn}' - J_{mn}) \frac{\epsilon_m - \epsilon_n}{T_{\nu}}.
\]  

(C5)

Write the difference of \( \dot{S}_1 \) and the second term as

\[
\frac{1}{2} \sum_{m,n,\nu} (J_{mn}' - J_{mn}) \left\{ \frac{C_f(p)}{C(p)} \left[ (g'(p_m) - g'(p_n)) - (\phi(J_{mn}') - \phi(J_{mn})) \right] - \frac{\epsilon_n - \epsilon_m}{T_{\nu}} \right\}.
\]  

(C6)

This sum can be positive or negative, depending on \( p_m \). Since we require that \( \dot{S}_1 \) must be non-negative, and vanishing only if the probability currents vanish, the difference must be zero, regardless of \( p_m \). Since \( (J_{mn}' - J_{mn}) \) can be positive or negative depending on \( p \), it means that the arguments in the curly brackets must be identically zero. Therefore we get that

\[
\frac{C_f(p)}{C(p)} \left[ (g'(p_m) - g'(p_n)) - (\phi(J_{mn}') - \phi(J_{mn})) \right] = \frac{\epsilon_n - \epsilon_m}{T_{\nu}}.
\]  

(C7)

The left-hand side is now depending on \( p_m \), which should not be the case. First, it means that \( C(p) = C_f(p) \). Second, it means that relation in square brackets does not depend on \( p_m \), from which we immediately get

\[
\phi(J_{mn}') = j'(w_{mn}') - g'(p_m)
\]  

(C8)

or more suggestively

\[
J_{mn}' = \psi(j'(w_{mn}') - g'(p_m))
\]  

(C9)

where \( j' \) is an arbitrary function and \( \psi = \phi^{-1} \). Since \( J_{mn}' = w_{mn}\Omega(p_m) \) we need to solve the following functional equation

\[
\psi(j'(w_{mn}') - g'(p_m)) = w_{mn}\Omega(p_m)
\]  

(C10)

for unknown functions \( \psi \) and \( j' \). This is a Cauchy multiplicative functional equation which determines \( \psi \) and \( j' \) as \( \psi(z) = \exp(z) \) and \( j'(z) = \log(z) \) (see [65]). By plugging into the equation, we immediately obtain that

\[
\Omega(p_m) = \exp(-g'(p_m)).
\]  

(C11)
By plugging the form of $\Omega$ into the local detailed balance, we immediately obtain that $\beta^r = \frac{1}{\tau^r}$. This result holds for any $g$. This concludes the proof. □

**Appendix D. Fokker Planck thermodynamics for nonlinear master equations**

Here we make a connection between the non-linear master and the corresponding Fokker–Planck equation on continuous spaces. Consider a general master equation in the form equation (20). Consider a continuous state space, and rewrite the master equation (20) in the following form:

\[
\dot{p}(x, t) = \int dr \left[ w(x - r|r)\Omega(p(x - r; t)) - w(|x - r)\Omega(p(x; t)) \right].
\]

(D1)

Proceeding in the usual way, expand the first term in terms of $(x - r)$ around $r = 0$. Zeroth term cancels with the second term in the previous equation and by truncation after the term quadratic in $r$, we obtain that

\[
\dot{p}(x, t) = -\frac{\partial}{\partial x} \int dr r \left[ w(x|r)\Omega(p(x; t)) \right]
+ \frac{1}{2} \int dr r^2 \frac{\partial^2}{\partial x^2} \left[ w(x|r)\Omega(p(x; t)) \right]
= -\frac{\partial}{\partial x} \left[ a_1(x)\Omega(p(x; t)) \right]
+ \frac{1}{2} \frac{\partial^2}{\partial x^2} \left[ a_2(x)\Omega(p(x; t)) \right],
\]

(D2)

where $a_n(x) = \int dr r^n w(x|r)$. Thus, we end with the non-linear Fokker–Planck equation

\[
\dot{p}(x, t) = -\frac{\partial}{\partial x} J(p(x, t), x, t),
\]

(D3)

where

\[
J = u(x, t)\Omega(p(x, t)) + D(x, t)\Omega(p(x, t)) \frac{\partial}{\partial x} g'(p(x, t)).
\]

(D4)

This equation is called the **free energy nonlinear Fokker–Planck equation** [67] due to its connection with thermodynamic quantities, as shown below. For simplicity, consider the case where $f(x) = x$, i.e. where the entropy has the trace-class form

\[
\mathcal{S}(t) = \int dx g(p(x, t)).
\]

(D5)

Evaluating, the time derivative of the entropy

\[
\dot{\mathcal{S}}(t) = \int dx g'(p(x, t))\dot{p}(x, t)
= \int dx J(p(x, t), x, t) \frac{\partial}{\partial x} g'(p(x, t)).
\]

(D6)

The last expression was obtained from integration by parts. By plugging in equation (D4), we get

\[
\dot{\mathcal{S}}(t) = \int dx J(p, x, t) \left( \frac{J(p, x, t)}{D(x, t)\Omega(p(x, t))} - \frac{u(x, t)}{D(x, t)} \right).
\]

(D7)

Here we can identify, similarly to the case of Shannon entropy [66], the entropy flow rate as

\[
\dot{\mathcal{S}}_f(t) = -\int dx \frac{J(p(x, t), x, t)u(x, t)}{D(x, t)}
\]

(D8)

and irreversible entropy production rate as

\[
\dot{\mathcal{S}}_r(t) = \int dx \frac{J^2(p(x, t), x, t)}{D(x, t)\Omega(p(x, t))} \geq 0.
\]

(D9)
Appendix E. Derivation of the second law of thermodynamics on the trajectory level

Let us consider a stochastic trajectory \(n(\tau)\) with jumps in \(t_j\) from \(n_j^-\) to \(n_j^+\). Let us calculate the time derivative of the stochastic entropy as

\[
\dot{s}(n(\tau)) = -\frac{\Omega_1(p_{n(\tau)}(\tau))}{\Omega(p_{n(\tau)}(\tau))} \frac{\partial}{\partial \tau} p_{n(\tau)}(\tau) - \sum_j \delta(\tau - t_j) \ln \frac{\Omega(p_{n_j^+}(\tau))}{\Omega(p_{n_j^-}(\tau))}
\]

which can be decomposed into entropy production rate

\[
\dot{s}_l(n(\tau)) = -\frac{\Omega(p_{n(\tau)}(\tau))}{\Omega(p_{n(\tau)}(\tau))} \frac{\partial}{\partial \tau} p_{n(\tau)}(\tau) - \sum_j \delta(\tau - t_j) \ln \frac{w_{n_j^+} \Omega(p_{n_j^+}(\tau))}{w_{n_j^-} \Omega(p_{n_j^-}(\tau))}
\]

and entropy flow rate

\[
\dot{s}_f(n(\tau)) = -\sum_j \delta(\tau - t_j) \frac{w_{n_j^+} \Omega(p_{n_j^+}(\tau))}{w_{n_j^-} \Omega(p_{n_j^-}(\tau))}.
\]

Thus, we get that \(\dot{s}(n(\tau)) = \dot{s}_l(n(\tau)) + \dot{s}_f(n(\tau))\). From the local detailed balance, we directly obtain that \(\dot{s}_f(n(\tau)) = \sum_j \beta_j \gamma_j(n(\tau))\). The ensemble second law can be obtained by averaging equation (E3) over all trajectories. For time derivatives of the stochastic quantities it means to multiply the quantities by \(p_n(\tau)\) and sum over all \(n\).

Appendix F. Derivation of the detailed fluctuation theorem

Here we focus on the log-ratio of probability distributions from the main text (equation (17)). The stochastic trajectory \(n(\tau)\) can be taken as the limit of a discrete trajectory \((n_k, t_k; n_{k-1}, t_{k-1}; \ldots, n_0, t_0)\). Express the log-ratio of multi-point probability distribution with help of equation (B1):

\[
\log \frac{P(n_k, t_k; \ldots, n_0, t_0)}{P(n_0, t_k; \ldots, n_k, t_0)} = \log \prod_{j=0}^{k} W(n_j, t_j | n_{j-1}, t_{j-1}) \Omega(p_{n_j}(t_j)) - \log \prod_{j=0}^{k} W(n_{j-1}, t_j | n_j, t_{j-1}) \Omega(p_n(t_j))
\]

\[
= \log \Omega(p_0) - \log \Omega(p_k) + \sum_{j=0}^{k} \log \frac{W(n_j, t_j | n_{j-1}, t_{j-1}) \Omega(p_{n_j}(t_j))}{W(n_{j-1}, t_j | n_j, t_{j-1}) \Omega(p_n(t_j))}.
\]

The first term corresponds to the change of the trajectory entropy from the initial to the final state and the second term corresponds to change of entropy flow along the trajectory. Thus, the log-ratio is equal to the change of entropy production \(\Delta s\) along the stochastic trajectory.

ORCID iDs

Jan Korbel https://orcid.org/0000-0002-5371-5320
David H Wolpert https://orcid.org/0000-0003-3105-2869

References

[1] Evans D J and Searles D J 1994 Equilibrium microstates which generate second law violating steady states Phys. Rev. E 50 1645–8
[2] Crooks G E 1999 Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences Phys. Rev. E 60 2721
[3] Jarzynski C 2000 Hamiltonian derivation of a detailed fluctuation theorem J. Stat. Phys. 98 77–102
[4] Evans D J and Searles D J 2002 The fluctuation theorem Adv. Phys. 51 1529–85
[5] Seifert U 2005 Entropy production along a stochastic trajectory and an integral fluctuation theorem Phys. Rev. Lett. 95 040602
[6] Barato A C and Seifert U 2015 Thermodynamic uncertainty relation for biomolecular processes Phys. Rev. Lett. 114 158101
[7] Horowitz J M and Gingrich T R 2020 Thermodynamic uncertainty relations constrain non-equilibrium fluctuations Nat. Phys. 16 15–20
[8] Falasco G, Esposito M and Delvenne J C 2020 Unifying thermodynamic uncertainty relations New J. Phys. 22 053046
[9] Ito S and Dechant A 2020 Stochastic time evolution, information geometry, and the cramér-rao bound Phys. Rev. X 10 021056
[10] Shiraishi N, Funo K and Saito K 2018 Speed limit for classical stochastic processes Phys. Rev. Lett. 121 070601
[11] Shiraishi N and Saito K 2019 Information-theoretical bound of the irreversibility in thermal relaxation processes Phys. Rev. Lett. 123 110603
[12] Ciliberto S 2017 Experiments in stochastic thermodynamics: short history and perspectives Phys. Rev. X 7 021051
Van den Broeck C and Esposito M 2010 Three faces of the second law. II. Fokker–Planck formulation Phys. Rev. E 82 011144
Frank T D 2004 Nonlinear Fokker–Planck Equations (Berlin: Springer)
Zhou Y and Yin C 2016 Power-law Fokker–Planck equation of unimolecular reaction based on the approximation to master equation Phys. A 463 445–51
Boon J P, Lutsko J F and Lutsko C 2012 Microscopic approach to nonlinear reaction-diffusion: the case of morphogen gradient formation Phys. Rev. E 85 021126
Borland I 1998 Microscopic dynamics of the nonlinear Fokker–Planck equation: a phenomenological model Phys. Rev. E 57 6634–42
Bertsch G F and Gupta S D 1988 A guide to microscopic models for intermediate energy heavy ion collisions Phys. Rep. 160 189–233
Chavanis P H 2003 Generalized thermodynamics and Fokker–Planck equations: applications to stellar dynamics in two-dimensional turbulence Phys. Rev. E 68 036108
Jin D Z and Dubin D H E 1998 Regional maximum entropy theory of vortex crystal formation Phys. Rev. Lett. 80 4434–7
Miller J 1990 Statistical mechanics of Euler equations in two dimensions Phys. Rev. Lett. 65 2137–40
Frank T D, Daffertshofer A and Beek P J 2002 Impacts of statistical feedback on the flexibility-accuracy trade-off of biological systems J. Biol. Phys. 28 39–54
Esposito M and Van den Broeck C 2010 Three detailed fluctuation theorems Phys. Rev. Lett. 104 090601
Naudts J 2011 Generalised Thermostatistics (Berlin: Springer Science & Business Media)
Hasegawa Y and Van Vu T 2019 Fluctuation theorem uncertainty relation Phys. Rev. Lett. 123 110602
Campisi M 2020 Dissipation asymmetry relation (arXiv:2011.01076)
Falasco G and Esposito M 2020 Dissipation-time uncertainty relation Phys. Rev. Lett. 125 120604
Dixit P D, Wagener J, Weistuch C, Presé S, Ghosh K and Dill K A 2018 Perspective: maximum caliber is a general variational principle for dynamical systems J. Chem. Phys. 148 010901
Hanel R and Thurner S 2011 When do generalized entropies apply? How phase space volume determines entropy Europhys. Lett. 96 50003
Illic V M and Stanković M S 2014 Generalized Shannon–Khinchin axioms and uniqueness theorem for pseudo-additive entropies Phys. A 411 138–45
Illic V M, Scarfone A M and Wada T 2019 Equivalence between four versions of thermostatistics based on strongly pseudoadditive entropies Phys. Rev. E 100 062135
Uffink J 1995 Can the maximum entropy principle be explained as a consistency requirement? Stud. Hist. Philos. Sci. B Stud. Hist. Philos. Mod. Phys. 26 223
Pressé S, Ghosh K, Lee J and Dill K A 2013 Nonadditive entropies yield probability distributions with biases not warranted by the data Phys. Rev. Lett. 111 180604
Tsallis C 2015 Conceptual inadequacy of the Shore and Johnson axioms for wide classes of complex systems Entropy 17 2853
Pressé S, Ghosh K, Lee J and Dill K A 2015 Reply to C. Tsallis’ ‘conceptual inadequacy of the Shore and Johnson axioms for wide classes of complex systems’ Entropy 17 5043
Ökonomou T and Bagci G B 2019 Rényi entropy yields artificial biases not in the data and incorrect updating due to the finite-size data Phys. Rev. E 99 032134
Korbel J 2021 Calibration invariance of the MaxEnt distribution in the maximum entropy principle Entropy 23 96
Gorban A N 2014 General H-theorem and entropies that violate the second law Entropy 16 2408–32
Abe S and Rajagopal A K 2003 Validity of the second law in nonextensive quantum thermodynamics Phys. Rev. Lett. 91 120601
Strasberg P and Esposito M 2017 Stochastic thermodynamics in the strong coupling regime: an unambiguous approach based on coarse graining Phys. Rev. E 95 062101