



Effective Fluctuation and Response Theory

Matteo Polettini¹ · Massimiliano Esposito¹

Received: 16 July 2018 / Accepted: 9 April 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

The so-called fluctuation theorems pushed the study of systems far beyond equilibrium, whose response to thermodynamic forces (affinities) is characterized by the reciprocal and the fluctuation-dissipation relations. All these results rely on the assumption that the observer has complete information about the system: no hidden leakage to the environment, exact evaluation of the thermodynamic cost of processes. Will an observer who has *marginal* information be able to perform an *effective* thermodynamic analysis? Given that such observer will only be able to establish local equilibrium, by perturbing the *stalling* currents will he/she observe equilibrium-like fluctuations? Within the formalism of Markov jump processes on finite networks, we provide a broad theory of the statistical behavior of some out of many currents that flow across a system. In particular (1) There exist effective affinities for which an integral fluctuation relation holds; (2) At stalling, i.e. where the marginal currents vanish, a symmetrized fluctuation-dissipation relation holds; (3) Under reasonable assumptions on the parametrization of the rates, effective affinities can be operationally defined by a procedure of tuning to stalling; (4) There exists a notion of *marginally time-reversed* process which restores the full-fledged fluctuation relation and reciprocity; (5) There exist fluctuation relations across different levels in the hierarchy of more and more “complete” theories. The above results apply to configuration-space currents, and to their *phenomenological* linear combinations provided certain symmetries of the effective affinities are respected—a condition whose range of validity we deem the most interesting question left open to future inquiry.

Keywords Fluctuation relations · Fluctuation–dissipation relations · Continuous-time Markov chains

Notation and Abbreviations

Acronyms

FR Fluctuation relation
IFR Integral fluctuation relation
MJPG Markov jump-process generator

✉ Matteo Polettini
matteo.polettini@uni.lu

¹ Complex Systems and Statistical Mechanics, University of Luxembourg, 162a Avenue de la Faïencerie, 1511 Luxembourg, Luxembourg

NESM	Non-equilibrium statistical mechanics
RR	Reciprocal relations
SCGF	Scaled-cumulant generating function
SFDR	Symmetrized fluctuation–dissipation relation
TR	Time reversed, time reversal
p.d.f.	Probability density function

Graphs

$ \cdot $	Cardinality of a set or range of an index
\mathcal{G}	Graph
\mathcal{E}	Edge set of a graph
\mathcal{I}	Site (vertex) set of a graph
\mathcal{C}	Simple oriented cycle
\mathcal{T}	Spanning tree
i, j, \dots	Sites
ij, ji, \dots	Oriented edges
∂	Incidence matrix

Linear Algebra

\vec{v}	Vector in $\mathbb{R}^{ \mathcal{I} }$
$\vec{1}$	Vector with all unit entries
A	Matrix $\mathbb{R}^{ \mathcal{I} } \rightarrow \mathbb{R}^{ \mathcal{I} }$
v	All other vectors

Observables

\cdot_{ij}	Edge observable
\cdot_{α}	Phenomenological observable
ψ, ϕ	Mean flux, mean current
Ψ^t, Φ^t	Time-integrated stochastic flux, and current
\mathcal{F}_{ij}	Thermodynamic force of a transition
\mathcal{A}	“Real” affinity
Q	Effective affinity
σ	Mean entropy production rate
Σ^t	Stochastic entropy production
\vec{p}	Steady-state of master equation
$\cdot^{\text{st}}, \cdot^{\text{eq}}$	At stalling, at equilibrium
$w_{ij}(\mathbf{x})$	Parametrized rates of the master equation

Stochastic Tools

ω^t	Stochastic trajectory
$\mathcal{D}\omega^t$	Path measure
P	Path p.d.f. and its marginals
$\langle \cdot \rangle$	Expected value w.r.t. $P(\omega^t)\mathcal{D}\omega^t$
$\tilde{\cdot}$	Hidden time-reversal
λ/Λ	SCGF of edge/phenomenological currents
ζ^t	Cumulant generating function at time t
$M(\{q_{\alpha}\})$	Tilted operator

1 Prologue

Consider an experiment where the flows of certain quantities are measured, due to certain applied forces. All it takes to properly address the thermodynamics of such setup is to be able to draw a net demarcation between the open changing system,¹ wherethrough physical quantities flow, and the decorrelated frozen environment wherefrom they come and go. Ideally for a proper thermodynamic analysis it is crucial to account for all of the currents flowing through the system, and to assign them their due thermodynamic cost.

However, things do not quite work that way, neither in practice nor in theory. The system's boundary, drawn for example on criteria of time-scale separation, of spatial localization, and of coarse-graining of irrelevant degrees of freedom, might not be crystal-clear. As a consequence, the measurement apparatus might not resolve important sources of dissipation due to parasitic currents. Furthermore, unless a microscopic theory is available that explains what exact causes produce which precise consequences, currents might not be assigned their proper thermodynamic cost. In other words, the observer might only have *marginal* information about the setup, due to technological and theoretical limitations. Nevertheless, the vast majority of physicists develop a sort of thermodynamic craftsmanship,² a learned sense for what is relevant in the controlled laboratory of their experiment (be it factual or thought). This leads them to identify *effective* thermodynamic forces that do not dispense with the laws of thermodynamics. In this process the logic of thermodynamic reasoning is preserved, but the nature of most of its measurable quantities needs to be renegotiated.

While according to Einstein thermodynamics is “the only physical theory of universal content [...] that will never be overthrown” [2], unlike other theories like Quantum Mechanics or General Relativity thermodynamics has for long been an intrinsically phenomenological science, a patchwork of profound laws and contingent principles that do not fit into a coherent mathematical framework. On these premises, it is common practice to invoke textbook thermodynamics in a literal way, deploying its jargon and formulas with little reference to the mental and physical processes through which such concepts were formulated. This is the fertile soil that foment the never-ending stream of pseudoscientific claims of violations the second law of thermodynamics, whose common root is a fundamental misunderstanding of which marginal currents and effective forces are actually at play.

Today a rigorous framework for the logical deduction of thermodynamic instances far from equilibrium is available [3] and in the course of experimental validation [4], and it claims to become *the* way thermodynamics is thought of and taught. Modern nonequilibrium statistical mechanics is based on the assumption that the system's configurations are explored by a Markovian dynamics, with transition probabilities biased according to thermodynamic incentives coming from the environmental reservoirs. This framework allows to characterize fluctuations of observables, and therefore it applies to small systems, not necessarily in the so-called “thermodynamic limit,” and in principle it applies arbitrarily far from equilibrium as long as the Markov assumption remains valid.

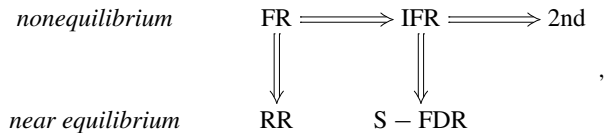
At the heart of equilibrium statistical mechanics lies the identification of *static* physical properties of a system (e.g. temperature, pressure etc.) with the average behavior of micro-

¹ In textbooks of thermodynamics and systems theory a distinction is made between open, closed, and isolated systems. In our perspective there is no substantial difference between flows of matter and energy. Hence closed systems are open; isolated systems can be thought of as idealizations of open systems where external influences are extremely feeble.

² The role of craftsmanship in the preparation and interpretation of a scientific experiment has been discussed by sociologist of science Collins [1]. In particular, his analysis of the trimmings with Bayesian inference (*ivi*, Chapter 5) struck us as particularly relevant for the foundations of statistical mechanics.

scopic degrees of freedom that fluctuate (e.g. average kinetic energy, velocity etc.). The first step out of equilibrium consists of slight perturbations of such observables, whose response can be characterized in terms of their spontaneous fluctuations at equilibrium, according to the so-called fluctuation–dissipation relation (FDR) and of the reciprocal relations (RR) that take the names of some of the heroes of 20th century physics [5–10]. For nonequilibrium systems, the picture is varied in a *dynamical* way: here the observables of interest quantify motility and directionality within a system. The connection between physical and statistical laws is encoded in the fluctuation relation (FR)—whose precise formulation is embodied in a plethora of so-called Fluctuation Theorems [11–15]. The FR states that the rate at which a system delivers entropy to the environment is a measure of the arrow of time, viz. of the asymmetry between the probability of microscopic paths and their time-reversed. To use a metaphor, the probability of “getting the toothpaste back into the tube” is exponentially suppressed with respect to that of “getting the toothpaste out of the tube,” using Woody Allen’s characterization of irreversibly in *Whatever works* [16].

Most often the paste spreads out according to the second law of thermodynamics (2nd), an inequality that in this setup easily follows from a more general identity, the integral fluctuation relation (IFR). All such relations can be resumed in the following implication diagram:



where by S-FDR we intend a symmetrized version of the usual Green–Kubo relation. The role of the First Law and other conservation laws is more subtle: it can be seen as a requirement on the form of the rates, which allows to identify the abstract Markovian jumps with physical currents. We will not consider the other laws of thermodynamics.

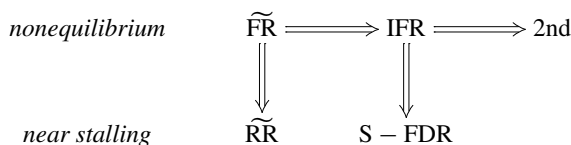
Crucially, establishing the above scheme requires that the observer has complete information about the system’s currents and forces. The question is then open as about how many of these results still apply to marginal observables of experimental interest, and what effective adjustments need eventually to be made. In particular, if the Markov process ventures into some sector of the configuration space that is hidden to the observer, how should we quantify the thermodynamic incentives over there?

The purpose of this paper is to present a general theory of the thermodynamics of a marginal set of currents and of the effective forces that drive them, under the assumption that somewhere in the belly of these coarser observables there lurk fundamental currents and forces that abide by the principles of Markovian stochastic thermodynamics. We first show that only the right-hand side of the above implication diagram stands:



Notice that in the marginal theory the analog of an equilibrium state is a *stalling* state in which the marginal currents vanish, while the hidden currents might still be flowing, as if the observer was in the eye of a hurricane. Importantly, on the condition that the hidden rates are not affected by the controllable parameters, the effective forces can be determined operationally by a simple tuning procedure, thus opening the way to experimental implementations of our theory.

Furthermore, from a more mathematical perspective, the left part of this diagram can be reinstated upon an appropriate redefinition of the underlying dynamics of the Markovian walker. Under this new “hidden time-reversed” dynamics, denoted by a wobble, we obtain the inference scheme



Finally, as the observer adds more and more currents and forces to his accounting, a “hierarchy” of marginal theories is explored: a cross-hierarchical FR holds, that is the ultimate core result of the whole construction.



The validity of FRs under coarse-graining and of FDRs far from equilibrium, in particular at stalling, are two questions that have been frequently addressed. Let us attempt an overview—itsself marginal.

Gallavotti [17] produced a convincing argument for why it is necessary to address FRs of local observables: As most systems of thermodynamic interest are large in volume, global observables are subjected to two extensive limits—one with respect to size and one with respect to time—hence rare events are even rarer. A special license is then needed to focus on localized non-extensive observables. In the formalism of chaotic dynamical systems, Gallavotti heuristically defined a local entropy production rate associated to a microscopic region of space that satisfies a FR.

The validity of FRs for coarse-grained observables has been considered in Refs. [18] and [19], where the IFR is studied when the observer has incomplete information. In particular, in the latter work measurement errors are introduced via a kernel that smoothens the sharp values of the “real” degrees of freedom into a distribution of coarser observed values. In one specific model it is found that the IFR can be preserved given a notion of effective work. However, differing from our setup, this quantity is not stochastic. The coarse-graining of the statistics of the currents for biochemical systems has been considered in Ref. [20]. The partial fluctuation theorem in systems weakly coupled to the environment has been studied in Ref. [21], where it is argued that a violation of the FR can persist even in the limit of vanishing interaction. Uhl et al. [22] have considered the fluctuations of an apparent entropy production in bipartite systems, finding many cases where an effective affinity restores the FR. For chemical networks where only some molecular species can be monitored experimentally, Bravi and Sollich [23] derived systematic models for subsystem dynamics that can help with the inference problem of estimating properties of the environment from observed sub-network dynamics. Another situation where the observer does not have access to all of the thermodynamic currents are stochastic models of so-called “Maxwell demons”, systems composed of an engine and a memory that operates a feedback control on the engine. To the total dissipation rate contribute fluxes of energy and of information, and an observer that does not duly keep into account the demon observes controversial behavior [24–26]. The FR in such models was investigated in Ref. [27], where the problem was solved by defining suitable observables that reinstate the FR, but which differ in nature from currents. Similar in spirit are the FRs for conditional and marginal probabilities discussed in Ref. [28], where appropriate terms are added to the entropy production rate in bipartite systems where one degree of freedom is neglected, and the hidden Markov models considered in Ref. [29]. All these approaches differ from ours as

we assess properties of marginal observables without resorting to *ad hoc* redefinitions of the stochastic observable under consideration (the currents).

Because of hidden heat flows, system-bath correlations in either classical or quantum systems, if not taken into proper account, might lead to violations of the laws of thermodynamics [30]. The authors of Ref. [31] comment that “in order to re-establish the laws of thermodynamics, one not only has to look at the local marginal systems, but also [at] the correlations between them”, and this can be achieved by some effective description. Effective thermodynamic potentials also play a role in systems strongly coupled to their surroundings [32]. Estimation of incomplete entropy production may also occur in systems with phenomenological forces accounting for fast interchange with the bath [33].

Another procedure that naturally leads to questions about marginal currents is the separation of fast vs. slow degrees of freedom. An effective affinity has been proposed to analyze experiments where a slow degree of freedom has been observed while the fast ones were integrated away [34]. The effect of time-scale separation in thermodynamics has been studied in Ref. [35] and recently in Refs. [36–38]. The former highlights that effective dynamics only preserve certain thermodynamic properties if internal detailed balance is obeyed, a feat that will play some role in our analysis of phenomenological currents. The latter show that in general the blanket is too narrow, and either dynamics or thermodynamics need to be sacrificed: while in their case it is thermodynamics, in ours it will be dynamics—viz. we are *not* presenting a theory of an effective dynamics in the observable configuration space.

Stalling steady states have been considered before by Qian [39], who dubbed the effective affinity that we will later introduce “isometric force”, and they play a role in the analysis of molecular motors [40]. Stalling currents also appear to play an important role in efficiency optimization, as e.g. in so-called Büttiker probes [41–43]. An effective two-terminal thermoelectric nanomachine, obtained starting from a more complete three-terminal machine with one stalling current, has been considered in Ref. [44] to study the effect of asymmetric Onsager coefficients on efficiency.

Response far from equilibrium is a broadly studied subject. In general, it is well-understood that the FDR has to be modified by including the correlation of the current with a quantity that is symmetric under time-reversal, alongside with the current’s self-correlation. This can give rise to interesting behavior such as negative differential mobility, i.e. the fact that one can “get less by pushing more,” as is well illustrated in the driven lattice Lorentz gas described in Ref. [45,46]. Experimental verifications of modified FDRs are also available [47]. We will only briefly address the response of systems arbitrarily far from equilibrium, and mostly focus on response at stalling. Far from equilibrium, the notion of an effective temperature has been investigated in weakly ergodic ageing systems [48,49].

Part of the material covered in this manuscript has been anticipated by the Authors in Ref. [50] for the case where the currents count a single transition in configuration space. Response out of stalling was presented to some extent in Ref. [51], which was stimulated by the specific analysis found in Refs. [52,53]. One of the Authors considered FRs for a marginal current in the case of electron transport in a double quantum dot in Ref. [54]. A construction analogous to ours that allows to prove IFRs for appropriate functionals was advanced by Shiraishi and other authors [55–57]. A comparison of our proposal and Shiraishi’s was provided in Ref. [58].



Outlying the theory in full requires to deploy a broad spectrum of techniques, ranging from Markov processes, to algebraic graph theory [59], to the theory of large deviations of stochastic processes [60], etc. We can only introduce them in a very compact form in Sect. 3.2

and give numerous references. We take the chance to shortly review in some detail some of the mathematical techniques that support the logical development of the theory, casting them in our own language. For some of these results, we provide novel derivations. A disclaimer about the mathematics: All our new results will be framed as “Propositions” in order pinpoint the logical structure of the discourse. Propositions are statements that, to the best of our understanding, are outlined and argued in a sufficiently self-consistent way, but which might fall short in meeting the standards that mathematicians intend. In particular, we make no distinction between Theorems, Lemmas, Corollaries, Remarks etc. We give no complete statement of the assumptions for each proposition because of an objective lack of expertise in the more advanced issues of probability theory and Markov processes. Nevertheless we trust the overall coherency of our argumentation, and we encourage improvement on rigor. The ornament on p.45 marks the point where most of the results are either new, or they are reinterpreted in a novel way.

Inspired by the pedagogical principle by Albert V. Baez in his unconventional physics textbook Ref.[61], we use a spiralling approach to the presentation of the material. The title, the abstract and this prologue represent the first three spirals of five more and more in-depth variations on the theme. The paper is structured as a Greek tragedy, with the main material exposed in a technical way in the *episodes*, preceded by the *prologue* that the reader is just reading, and most importantly by the *parode*, the first song sung by the chorus, which anticipates the main themes in a self-contained manner. Throughout the play the chorus stays on stage as a constant interlocutor, so the reader should always keep in mind the voice of the *parode*, which is structured into a *strophê* and a *antistrophê* with the same meter, where we present the older material and the newer one, in parallel ways. The *parode* ends with an *epode* on future perspectives related to our results, while more technical conclusions are drawn in the closing *exode*. The preceding *stasimon*, the final song sung by the choir, is in a diminished locrian mode. The reason why this story should be a tragedy, rather than a comedy, is not clear to the authors.

2 Parode: Enunciation of the Main Results

Before dwelling into our theory in full detail, in this Parode we present a less technical, yet self-contained discussion of the main results, which can be considered as an independent letter on its own. This “entrance ode” is meant to provide the reader knowledgeable in the field with enough details to reproduce most of the results on his own, and the neophyte with an overview on the main lines of reasoning. Quoting (with minor adjustments.) A. V. Baez [61], «the reader may go through this section rapidly, as it takes him through a round of the spiral; the treatment may strike him as light, even inadequate. But he/she should rest assured, however, that we are laying a good foundation for a more concise and mathematical treatment in the following chapters ».

We will first introduce the known facts regarding a “complete” set of currents to make contact with established knowledge and lingo in the field. We then introduce our new results about marginal sets of currents, paralleling them to the older results. Finally we explain some of the main technical ingredients underlying our results, and draw conclusions.

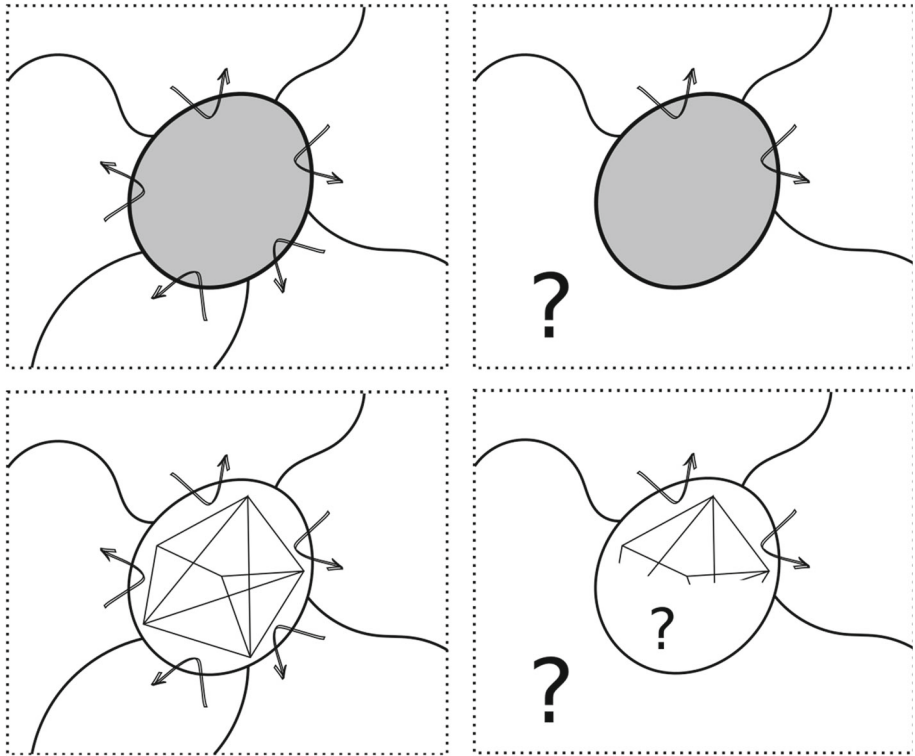


Fig. 1 First row: on the left, pictorial representation of a system through which currents flow; on the right, the standpoint of a marginal observer that only measures certain currents. In the second, the system's state space that mediates the passage of currents is resolved: it is a discrete network

2.1 Strophê: “Complete” Fluctuations and Response

Macroscopic thermodynamics describes systems through which a certain number³ $|\alpha|$ of (steady) currents ϕ_α flow, powered by conjugate thermodynamic forces or *affinities* \mathcal{A}_α . The system can be seen as an interface between several reservoirs, with the currents flowing through the system, to and from reservoirs. We take currents and affinities to be a “complete” set of core, irreducible observables, assuming that all conservation laws, e.g. of energy (First Law of thermodynamics), number of particles, etc. have already been taken care of by gauging out certain reference reservoirs. We shall explain what it exactly means to be “complete” later in this Parode. Then the affinities usually are *differences* of inverse temperatures, chemical potentials, etc., and currents are “conserved on their own”. For this reason we drew them as in-out “reservoir arrows” in the illustration Fig. 1. The macroscopic entropy production rate (EPR) is defined as the bilinear form

$$\sigma := \sum_{\alpha} \phi_{\alpha} \mathcal{A}_{\alpha}. \quad (1)$$

³ Symbol $|\cdot|$ denotes both the cardinality of sets and the range of indexes. Wherever possible, we will omit to specify the range of indexes.

When we go microscopic, because of thermal noise we need to allow for fluctuations. We thus make currents into random variables. We consider a single realization, or *path* or *trajectory* ω^t of a hypothetical experiment in a time window $[0, t]$. The stochastic time-integrated currents $\Phi_\alpha^t := \Phi_\alpha[\omega^t]$ are functionals of such trajectory, typically time-extensive. Therefore

$$\phi_\alpha := \lim_{t \rightarrow \infty} \frac{\langle \Phi_\alpha^t \rangle}{t} \quad (2)$$

converges and yields the mean steady currents, where the average is taken with respect to a probability measure over trajectories $P(\omega^t) \mathcal{D}\omega^t$, that we will describe in detail in Sect. 3.5.1.

Along a single realization of the process, we define the *entropy production* as

$$\sum \Phi_\alpha^t \mathcal{A}_\alpha + O(1). \quad (3)$$

Here $O(1)$ stands for contributions that do not add-up in time, which are due to the transient adjustment of the system's internal entropy. With a slight stretch of imagination, the entropy production can be considered as the amount of entropy delivered to the environment during the process; however, to be slightly pedantic, we must emphasize that in our approach there is no such thing like a state function “entropy of the environment”.

The entropy production is the major actor in the so-called fluctuation relation (FR) [13–15, 62–64]

$$\frac{P(\{\Phi_\alpha\})}{P(\{-\Phi_\alpha\})} = \exp \sum \Phi_\alpha \mathcal{A}_\alpha \quad (4)$$

where $P(\{\Phi_\alpha\})$ is the probability density function (p.d.f.) that the time-integrated currents Φ_α^t take values in a neighborhood of Φ_α . In the rest of the paper we will adopt a strategy discussed in Ref. [62] by which we can deal with finite-time FRs on the same footing as with asymptotic ones, whereby the above relation is exact equality at all times, provided the initial configuration from which the trajectory departs is selected with an appropriate probability distribution. In all those cases where results only hold in the long-time limit, we will use the asymptotic equality \asymp . The reader not interested in these subtleties might just view all the FRs as asymptotic at $t \rightarrow +\infty$.

An immediate corollary of the FR is the integral fluctuation relation (IFR)

$$\langle \exp - \sum \Phi_\alpha \mathcal{A}_\alpha \rangle = 1. \quad (5)$$

The IFR embodies and refines the Second Law of thermodynamics, which states that on average the EPR is non-negative,

$$\sigma \geq 0, \quad (6)$$

an immediate consequence of Eq. (5), via Jensen's inequality for convex functions.

A system is said to be *detailed balanced* when all of the affinities vanish; in this case the steady state is an equilibrium, that is, all mean steady currents vanish:

$$\text{equilibrium} : \mathcal{A}_\alpha = 0, \forall \alpha \iff \phi_\alpha = 0, \forall \alpha. \quad (7)$$

It is well known that the FR actually gives rise to a cornucopia of IFRs [65]. We notice here in passing, as a new result, that in the case where only two processes contribute to the total entropy production, $|\alpha| = 2$, from the FR also follows

$$\langle \exp - \Phi_1 \mathcal{A}_1 \rangle = \langle \exp - \Phi_2 \mathcal{A}_2 \rangle. \quad (8)$$

We dub this the *reciprocal IFR*. The interesting feature of this relation is that it resolves and relates the statistics of two currents that can be quite different in physical nature.

The FR allows to derive all known results of response theory close to equilibrium. To attack this problem, we introduce an explicit dependency of the structural properties of the system (viz. the transition rates of the underlying stochastic dynamics) on certain parameters $\mathbf{x} = \{x_\kappa\}_{\kappa=1}^{|\kappa|}$, $|\kappa| \geq |\alpha|$, with the following requirement:

- A0 The first $|\alpha|$ of these parameters are *thermodynamic*, meaning that there exist constants x_α^{eq} such that

$$\mathcal{A}_\alpha(\mathbf{x}) = x_\alpha - x_\alpha^{\text{eq}}. \quad (9)$$

All other parameters $\{x_\kappa\}_{\kappa>|\alpha|}$, on which the affinities do not depend, are *kinetic*.

Notice that while this is just a contrived way to say that we either perturb the affinities, or some property that does not alter the affinity, this subtlety will play an important role below. At $\mathbf{x} = \mathbf{x}^{\text{eq}}$ the system satisfies detailed balance whereby all of the forces $\mathcal{A}_\alpha(\mathbf{x}^{\text{eq}}) = 0$ vanish and so do the currents, $\phi_\alpha^{\text{eq}} := \phi_\alpha(\mathbf{x}^{\text{eq}}) = 0$, where the superscript “eq” means “evaluated at $\mathbf{x} = \mathbf{x}^{\text{eq}}$ ”.

For systems that are slightly perturbed out of equilibrium, two major results hold: the (symmetrized) fluctuation–dissipation relation (S-FDR) and the reciprocal relations (RR) [7,8,10]. Defining the response coefficients as

$$\phi_{\alpha;\kappa} := \frac{\partial \phi_\alpha}{\partial x_\kappa}, \quad (10)$$

these two near-equilibrium relations state respectively that the response to a variation of a thermodynamic parameter at $\mathbf{x} = \mathbf{x}^{\text{eq}}$ satisfies

$$\phi_{\alpha;\alpha'}^{\text{eq}} + \phi_{\alpha';\alpha}^{\text{eq}} = \phi_{\alpha\alpha'}^{\text{eq}} \quad (11a)$$

$$\phi_{\alpha;\alpha'}^{\text{eq}} - \phi_{\alpha';\alpha}^{\text{eq}} = 0 \quad (11b)$$

where

$$\phi_{\alpha\alpha'} := \lim_{t \rightarrow \infty} \frac{1}{t} \langle (\Phi_\alpha^t - \langle \Phi_\alpha^t \rangle) (\Phi_{\alpha'}^t - \langle \Phi_{\alpha'}^t \rangle) \rangle \quad (12)$$

is the steady-state variance of the currents, properly scaled with time. The S-FDR and the RR can be proven quite straightforwardly as corollaries of the FR. More importantly for this paper, the first result follows as a corollary of the IFR Eq. (5), and the second as a corollary of the reciprocal IFR Eq. (8). At equilibrium, currents do not respond to a variation of the kinetic parameters:

$$\phi_{\alpha;\kappa}^{\text{eq}} = 0, \quad \kappa > |\alpha|. \quad (13)$$

Furthermore, the FR can be employed to produce higher-order response relations [66–68], which constitute the most promising testing ground for our theory. For example, at third order, focusing on one single current, near equilibrium one obtains

$$\phi_{\alpha\alpha\alpha}^{\text{eq}} = 0 \quad (14a)$$

$$\phi_{\alpha\alpha;\alpha}^{\text{eq}} - \phi_{\alpha;\alpha\alpha}^{\text{eq}} = 0, \quad (14b)$$

where $\phi_{\alpha\alpha\alpha}$ is the scaled third-order cumulant. The first relation is due to the fact that at equilibrium the current p.d.f. is symmetric, $P^{\text{eq}}(\{\Phi_\alpha\}) = P^{\text{eq}}(\{-\Phi_\alpha\})$, and the second

expresses the second-order response of the average current in terms of the first-order response of its variance.



Let us now sketch the mathematical framework and assumptions based on which the above results can be derived (full details in Sect. 3.2). We consider a continuous-time, discrete configuration-space Markov “jump” process occurring on a finite network with configurations (*sites* in graph-theoretic language) i, j, \dots connected by transitions (*oriented edges*) ij, \dots , where the i is the final site and j is the starting site, following the right-to-left physicists’ convention. The dynamics can be described by an evolution equation for the probability $p_i(t)$ of being in configuration i at time t , governed by the master equation

$$\frac{d}{dt} \vec{p}(t) = W \vec{p}(t). \quad (15)$$

Here, $\vec{p} = (p_i(t))_i$ and W is a Markov-jump process generator (MJPG) with entries

$$W_{ij} = \begin{cases} w_{ij}, & i \neq j \\ -w_i & i = j. \end{cases}, \quad (16)$$

where $w_i = \sum_j w_{ji}$ is the exit rate out of a configuration. We call W the *forward* generator. The dependence on the external parameters is encoded in the rates $w_{ij} = w_{ij}(\mathbf{x})$. Given the steady state of the dynamics \vec{p} , satisfying $W \vec{p} = 0$, one can construct the time-reversed generator $\bar{W} = P W^T P^{-1}$, or, simply, time reversal (TR), where $P = \text{diag}(p_i)_i$. In a sense, time reversal “runs steady states back in time,” with detailed-balanced (equilibrium) systems obeying time-reversal symmetry

$$W(\mathbf{x}^{\text{eq}}) = \bar{W}(\mathbf{x}^{\text{eq}}). \quad (17)$$

Time-integrated *edge currents* Φ_{ij}^t count the net number of times a certain transition is performed along a single realization of the process; all possible current-like observables Φ_α^t are linear combinations of edge currents. It is usually more practical to study the currents’ statistics via their cumulants, properly scaled in time. An important result in the theory of the large deviations of Markov processes allows to obtain the currents’ scaled cumulant generating function (SCGF)⁴ $\lambda(\{q_\alpha\})$ as the dominant eigenvalue of a suitably defined “tilted” operator, see Sect. 3.5.3. Then the FR Eq. (4) translates into the following fluctuation symmetry [15]

$$\lambda(\{q_\alpha\}) = \lambda(\{\mathcal{A}_\alpha - q_\alpha\}), \quad (18)$$

while the IFR Eq. (5) reads

$$\lambda(\{\mathcal{A}_\alpha\}) = 0 \quad (19)$$

and the reciprocal IFR Eq. (8) reads

$$\lambda(\{\mathcal{A}_1, 0\}) = \lambda(\{0, \mathcal{A}_2\}). \quad (20)$$

All of the above results hold in the following “complete” setups:

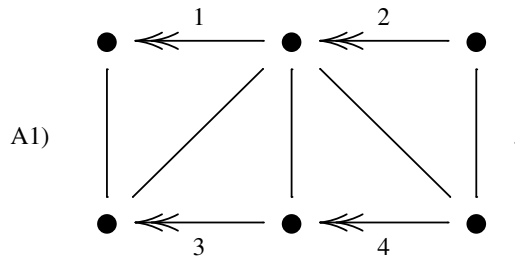
- A1 Index α spans through independent *cycles* in the network. The affinities are computed as the sum of the log-ratio of the rates $\log w_{ij}/w_{ji}$ along such cycles, and their conjugate currents Φ_α^t are edge currents associated to certain preferred transitions in the network,

⁴ We actually adopt the sign convention of Touchette [60] rather than that of Lebowitz and Spohn [15] on the definition of the tilted operator, and thus on the SCGF—which in our case is actually a *signed* SCGF.

in the light of Schnakenberg's cycle analysis of steady states, which is the analog of Kirchhoff's mesh analysis of electrical circuits applied to Markov processes [64,69,70].

- A2 Index α ranges through a smaller number of *phenomenological* currents, which in realistic physical models are associated to several transitions in the configuration network of a system. Provided such transitions cover at least a basis of cycles, correspondingly, for the FR to hold, the cycle affinities must enjoy certain symmetries, a condition called *local detailed balance* or *thermodynamic consistency*—or simply consistency, systematically analyzed in Refs. [71,73].

For example, if measured independently, the set of currents denoted by double arrows in the following network form a minimal “complete” set according to setup A1,

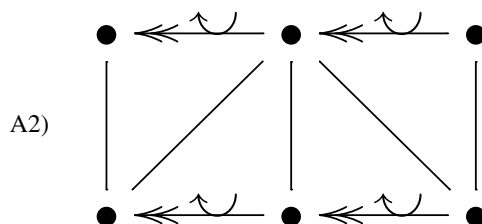


As detailed in Sect. 3.4, the corresponding cycle affinities are calculated as the log-ratio of the product of the rates on a basis of fundamental cycles. We can represent them diagrammatically as

$$\begin{aligned}
 \mathcal{A}_1 &= \log \frac{\text{Diagram 1}}{\text{Diagram 2}}, & \mathcal{A}_2 &= \log \frac{\text{Diagram 3}}{\text{Diagram 4}}, \\
 \mathcal{A}_3 &= \log \frac{\text{Diagram 5}}{\text{Diagram 6}}, & \mathcal{A}_4 &= \log \frac{\text{Diagram 7}}{\text{Diagram 8}},
 \end{aligned}
 \tag{21}$$

where the arrows in the diagrams imply multiplication of the corresponding rates.

Instead, if the observer is only capable of measuring a linear combination of the above currents, then we fall within setup A2. This might be the case when several transitions in configuration space correspond to the exchange of the same physical quantity with one particular reservoir, a situation that we illustrate by a curly “reservoir arrow,” borrowing from the chemistry literature. For example, the system



corresponds to the situation where each of the four transitions contributes one unit to the phenomenological current, but the observer would not be able to tell which one of the transitions happened. Then, to grant consistency, for this particular example the affinities along the cycles depicted above must all take the same value (otherwise, by a calorimetric experiment the observer *would* be able to tell the difference!).

Let us give an insight on the physical interpretation of transition rates and on their thermodynamically consistent parametrization assumed in A0. For detailed balanced systems subject to conservative forces, the most general form that the rate of hopping from site j to site i can take is

$$w_{ij}^{\text{eq}} = v_{ij} \exp -u_j \quad (22)$$

where $v_{ij} = v_{ji} > 0$ is symmetric. From a physical standpoint, in view of e.g. the Arrhenius law [72], one can portray the configuration space of a system as a landscape with the sharpest minima at the configuration sites, separated by activation barriers. The configuration function u_i and the symmetric term v_{ij} fully describe such an *internal landscape*. For systems that do not satisfy detailed balance an asymmetric term $a_{ij} = -a_{ji}$ appears and we can generally write [73]

$$w_{ij} = w_{ij}^{\text{eq}} \exp a_{ij}/2. \quad (23)$$

The intuition is that the non-conservative term a_{ij} is a relic of the interaction of the system with the degrees of freedom of an external reservoir that influences the transition. For example, in the procedure of obtaining an open irreversible chemical network from a closed one by chemostating chemical species described in Ref.[74], the internal landscape is fully encoded in the reaction rates, while the terms a_{ij} correspond to the concentrations of the external chemostats. Importantly, thermodynamic affinities only depend on the latter: the transformation of one particular affinity $d\mathcal{A}_\alpha$, at fixed values of all other affinities, only involves: (A0.i) a *local* transformation of the external “relic” terms along the network’s edges that are peculiar to that particular mechanism; (A0.ii) a *global* transformation of the internal energy landscape. We will detail this issue in Sect.5.8.

We will call a transformation of the form

$$w_{ij} \rightarrow w'_{ij} = w_{ij} e^{-a_j} \quad (24)$$

a *gauge* transformation. Under such a transformation, the log-ratio of the rates $\log w_{ij}/w_{ji}$ transforms like an inhomogeneous gauge connection, but the affinities are invariant. In a sense, they are the Wilson loops of the theory. This nomenclature is, in fact, more than an analogy. Gauge invariance of nonequilibrium thermodynamics is a concept put forward by one of the authors in Refs. [75,76]. There, it is argued that it is a necessary property if one wants to make sense of thermodynamics as a science of information and ignorance [77], as the corresponding continuous symmetry corresponds to a modification of prior probabilities. Therefore gauge invariance allows to deal with biases encoded in prior information, often

perceived as a threat to the “objectivity” of the theory. A more physical motivation of gauge invariance in terms of thermodynamics potentials can be found in Ref. [78].

2.2 Antistrophê: Marginal Fluctuations and Response

We now focus on a subset of $|\mu| < |\alpha|$ currents $\{\Phi_\mu^t\}$ and consider their marginal p.d.f.

$$P(\{\Phi_\mu\}) := \int \prod_{\alpha > |\mu|} d\Phi_\alpha P(\{\Phi_\alpha\}). \quad (25)$$

The questions we address are: which of the above relations survive, what new results emerge, and under which (presumably stricter) conditions?

The central result in this paper (Propositions 3, 19) is that there exist *effective affinities* Q_μ such that a *marginal IFR* holds

$$\langle \exp - \sum \Phi_\mu Q_\mu \rangle = 1, \quad (26)$$

despite the fact that the full-fledged FR does not,

$$\frac{P(\{\Phi_\mu\})}{P(\{-\Phi_\mu\})} \neq \exp \sum Q_\mu \Phi_\mu, \quad (27)$$

and, provided there is at least one additional unobserved current, neither does the reciprocal IFR,

$$\langle \exp - \Phi_1 Q_1 \rangle \neq \langle \exp - \Phi_2 Q_2 \rangle, \quad (28)$$

Here and below \neq loosely means “generally not,” keeping into consideration that one can always fabricate systems whose marginal currents do obey the marginal FR (e.g. systems with statistically independent currents because of a special topology of the network). That the FR does not generally hold can already be deduced by the analysis of specific examples, see e.g. Ref. [53]. Again, our marginal IFR holds asymptotically, for any given initial ensemble, or at all times provided the trajectory’s initial configuration is sampled from a special state \vec{p}^{st} that we will describe shortly.

However, a moment of reflection leads to the conclusion that, *per se*, the existence of values of the $\{Q_\mu\}$ that make Eq. (26) true should be no surprise. If we are allowed to tune such values at will, the average of the exponential can definitely range anywhere from 0 to $+\infty$. As we will discuss later in this Parode, for $|\mu| > 1$ there actually is a continuum of candidate effective affinities fulfilling the marginal IFR. Thus, what is important is not that there exist such values, but that they can be given an operational interpretation.⁵ We reserve the expression “effective affinities” and the notation $\{Q_\mu\}$ to one particular choice of those values, identified by a constructive procedure that we will soon detail, and that most importantly grants that they are *marginally thermodynamic* in the sense that

$$\frac{\partial}{\partial x_{\mu'}} Q_\mu(\mathbf{x}) = \delta_{\mu, \mu'} = \frac{\partial}{\partial x_{\mu'}} \mathcal{A}_\mu(\mathbf{x}). \quad (29)$$

This is crucial if we want to produce a response theory. However, this latter fact requires to slightly reduce the scope of assumption A0:

⁵ Interestingly, the same emphasis on this operational aspect is found in the already mentioned textbook by Baez: “An understanding of concepts requires, however, much more than the ability to recite the associated words and their dictionary definitions. It is necessary to study, and preferably to experience, the *operations* that give meaning to the words.”

B0 Thermodynamic parameters only affect the rates of the networks' edges that support the current of observational interest.

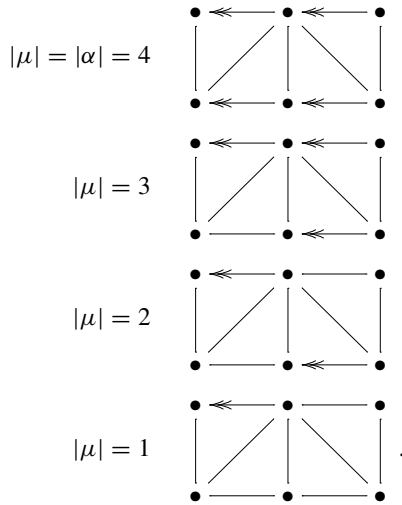
We will investigate at length the difference between A0 and B0 in Sect. 5.8. Let us already give a piece of good news, in the light of the physical parametrization of the transition rates discussed in the previous section. The only difference with respect to the parametrization of the “real” affinities of the “complete” theory is that modifications of the internal landscape might affect effective affinities. Therefore we can only afford (A0.i) the same *local* transformation of the external antisymmetric terms along the network's edges that are peculiar to that particular mechanism. Instead, we need to replace (A0.ii) with (B0.ii) a *local* transformation of the internal energy landscape. This is not a dramatic restriction. As a matter of fact, the workings of Proposition 28 basically show that there is not much more to “thermodynamic parametrization” than there is in “local parametrization,” so that this whole discussion can be safely dismissed: the whole point of this discourse is to show that the parametrization does not really affect the theory, unless one plays devil's advocate by picking a very nonlocal and contrived parametrization. As far as we only modify reservoir properties (e.g. temperatures, chemical potentials), we are on safe grounds.

From the marginal IFR follows that the marginal EPR $\sum Q_\mu \phi_\mu$ is positive, while notice that in general the “piece” of EPR $\sum \mathcal{A}_\mu \phi_\mu$ might be not, due to the phenomenon of transduction by which some currents can be made to run against their conjugate thermodynamic forces by a conjure of the other currents and forces [79]. More interestingly, we prove in Proposition 21 that the marginal EPR is always smaller than the “complete” EPR

$$0 \leq \sum Q_\mu \phi_\mu \leq \sum \mathcal{A}_\alpha \phi_\alpha, \quad (30)$$

This generalizes the results of Ref. [58], which deals with the case $|\mu| = 1$. These considerations open up the question in what sense $\sum_\mu Q_\mu \phi_\mu$ can actually be interpreted as EPR from a marginal point of view. As shown in Ref. [50], and recapitulated in Sect. 4.5, in the case of a single current supported on one edge, indeed this quantity represents the putative EPR evaluated by a local observer that can only access information about a specific transition of the system, and who formulates a minimal steady-state model of the hidden sector of the system. We lack the generalization of this latter argument to $|\mu| > 1$.

In fact, there exists entire *hierarchies* of marginal theories, according to whether one measures $|\mu| = 1, 2, \dots, |\alpha|$ currents, up to a “complete” set. For example, the above case study admits $|\alpha|! = 24$ hierarchies, among which



Within any one such hierarchy, we will be able to show that the mean EPR estimated at each level is smaller than that estimated at the subsequent level,

$$0 \leq \dots \leq \sum_{\mu=1}^{|\mu|} Q_{\mu}^{1, \dots, |\mu|} \phi_{\mu} \leq \dots \leq \sum_{\alpha=1}^{|\alpha|} Q_{\alpha}^{1, \dots, |\alpha|} \phi_{\alpha}, \quad (31)$$

where we now added a superscript as a further specification of the effective affinities, to highlight the fact that they are associated with the $|\mu|$ -th theory in the hierarchy. That is because effective affinities associated to the same current, but referring to different levels in the hierarchy, are generally different. The “real” affinities $\mathcal{A}_{\alpha} = Q_{\alpha}^{1, \dots, |\alpha|}$ are the last in the hierarchy.

(We now go back to dropping the hierarchy specification superscript $1, \dots, |\mu|$.) A system for which all marginal currents vanish, $\phi_{\mu}^{\text{st}} = 0, \forall \mu$, is said to be at *stalling*, where it stalls. We will show (Propositions 11, 22) that one achieves stalling if and only if all of the effective affinities vanish:

$$\text{stalling} : Q_{\mu} = 0, \forall \mu \iff \phi_{\mu} = 0, \forall \mu. \quad (32)$$

While the marginal currents vanish at stalling, all other currents need not vanish. Hence stalling steady states are generally far from equilibrium, and can be interpreted as states of “local equilibrium” with respect to our hypothetical marginal observer. Clearly, the variety of stalling values \mathbf{x}^{st} includes that of equilibrium values \mathbf{x}^{eq} , and typically the latter is a set of zero measure in the former.

Let us now consider response to perturbations out of stalling. The IFR alone grants the validity of the S-FDR, but not of the RR:

$$\phi_{\mu; \mu'}^{\text{st}} + \phi_{\mu'; \mu}^{\text{st}} = \phi_{\mu \mu'}^{\text{st}} \quad (33a)$$

$$\phi_{\mu; \mu'}^{\text{st}} - \phi_{\mu'; \mu}^{\text{st}} \neq 0. \quad (33b)$$

This is a clear-cut experimental prediction of our theory: While the S-FDR relation is common to response out of equilibrium and out of stalling, the violation of the RR is a signature of stalling. Notice that Eq.(29) is a guarantee that perturbations with respect to the effective

affinities are the same as perturbations with respect to the “real” affinities, so that to test response at stalling no specific experimental protocol has to be devised that is inherently different than that at equilibrium, provided assumption B0 is met. This is crucially important: We want the experimental apparatuses of the “complete” and the marginal theories to be the same, because in principle there is no a priori assurance that the system we are going to measure is actually complete.

Let us now look at other signatures of stalling. Differing from the “real” affinities, the effective affinities might still depend on the rest of the parameters $\mathbf{z} := \{x_\kappa\}_{\kappa > |\mu|}$

$$Q_\mu(\mathbf{x}) = x_\mu - x_\mu^{\text{st}}(\mathbf{z}). \quad (34)$$

These include the kinetic ones. In other words, the effective affinities might be sensible to modifications of the internal landscape even in the hidden sector of the system. This gives rise, in spite of Eq. (13), to the response formula

$$\left(\sum_\mu \phi_{\mu;\kappa} \frac{dx_\mu^{\text{st}}}{dx_\kappa} + \frac{1}{2} \sum_{\mu, \mu'} \phi_{\mu\mu'} \frac{dx_\mu^{\text{st}}}{dx_\kappa} \frac{dx_{\mu'}^{\text{st}}}{dx_\kappa} \right)^{\text{st}} = 0, \quad (35)$$

equipped with orthogonality relation

$$\sum_\mu \phi_\mu \frac{dx_\mu^{\text{st}}}{dx_\kappa} = 0. \quad (36)$$

Hence, perturbations of kinetic parameters, even far from the observable configurations, might lead to a perturbation of the steady state out of stalling. Local equilibrium is more fragile than equilibrium, as intuitive.

Considering higher cumulants, at third order we find, in spite of the two equations in Eq. (14), that

$$\phi_{\mu;\mu\mu}^{\text{st}} - \phi_{\mu\mu;\mu}^{\text{st}} = \frac{1}{3} \phi_{\mu\mu\mu}^{\text{st}}, \quad (37)$$

while in general

$$\phi_{\mu\mu\mu}^{\text{st}} \neq 0. \quad (38)$$

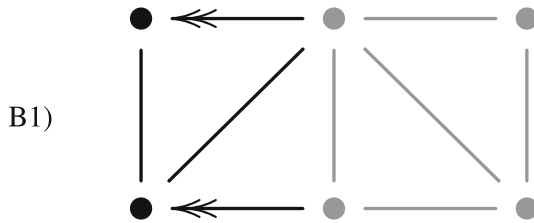
This is due to the fact that, at stalling, the marginal p.d.f. is not necessarily symmetric, $P^{\text{st}}(\{\Phi_\mu\}) \neq P^{\text{st}}(\{-\Phi_\mu\})$. Therefore, the skewness of the p.d.f. is a signature of a stalling steady state.



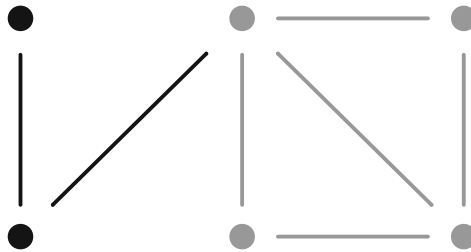
In the framework of Markov jump-processes on a network briefly described above, marginal currents and their conjugate effective affinities can either be

- B1 Currents flowing along single edges, but with some cycles left out from the accounting. Effective affinities are uniquely identified by the theory.
- B2 Phenomenological currents. In this case, our construction only holds provided that effective affinities satisfy a condition of *marginal consistency* that with simple examples can be shown to be stricter than “complete” consistency.

An example of setup B1 is the following network



where arrowed transitions are observable, and grey transitions are hidden. The remaining two transitions are kept black because, at a steady state, the current flowing through them is known. Let us denote the observable transitions $i_\mu j_\mu$. The effective affinities are identified according to the following recipe: Remove all the observable transitions (we may assume for simplicity that the network remains connected, but this is not mandatory):



On such a reduced configuration space, let us consider the dynamics described by the *hidden generator* W_{hid} obtained by setting the rates of the observable transitions to zero, $w_{i_\mu j_\mu} = w_{j_\mu i_\mu} = 0$. Let the system relax to the *stalling* steady state \vec{p}^{st} of the hidden dynamics, $W_{\text{hid}} \vec{p}^{\text{st}} = 0$. Then the effective affinities are given by

$$Q_\mu = \log \frac{w_{i_\mu j_\mu} p_{j_\mu}^{\text{st}}}{w_{j_\mu i_\mu} p_{i_\mu}^{\text{st}}}. \quad (39)$$

From an operational perspective, to find the effective affinities there is no need to know the actual rates, as one can just tune parameters to make the currents stall $\{\phi_\mu\} = 0$. In fact, notice that a local parametrization of the rates such as

$$\frac{w_{i_\mu j_\mu}(\mathbf{x})}{w_{j_\mu i_\mu}(\mathbf{x})} = \exp x_\mu \quad (40)$$

yields

$$Q_\mu(\mathbf{x}) = x_\mu - x_\mu^{\text{st}} \quad (41)$$

where $x_\mu^{\text{st}} = \log(p_{i_\mu}^{\text{st}}/p_{j_\mu}^{\text{st}})$. Furthermore, we can show that there is no difference between “tuning to stalling” and “removing” as far as the stalling steady state is concerned (see Proposition 23). This latter expression is the fundamental link between the mathematical and the operational definitions of the effective affinities, thus the cornerstone of the physical interpretation of our theory. The effective affinities have a twofold characterization. On the one hand, they can be interpreted as the forces exerted on the observable edges at a quench, that is, by preparing the system in steady state \vec{p}^{st} and then suddenly switching on the transition rates. On the other, they can be obtained by tuning the controllable parameters

to the stalling values that make currents stall. Furthermore, the effective affinities will be shown to be gauge invariant under transformation Eq. (24), exactly like their “complete” counterparts, thus granting the compatibility of the theory with foundational requirements.

Notice that, like we briefly mentioned above, when considering an observer who adds more and more currents to his/her observational basket, different marginal theories are generated. Now the reason is clear: the stalling steady state obtained by removing $|\mu|$ edges is different from that obtained by removing a subset $|\mu'| \subset |\mu|$. As observed, the effective affinities conjugate to one particular current differ among themselves at different levels of such hierarchy, which implies that the stalling values $x_{\mu}^{|\mu|,st}$ differ as well. From an operational point of view, this phenomenon has a simple interpretation: by virtue of Eq. (35), once the first $|\mu'|$ currents stall, tuning the other $|\mu| - |\mu'|$ to stalling will also perturb the first, thus disrupting the stalling steady state achieved before. This creates space for an interesting question, whether there exists a smart iterative procedure to tune to stalling.

Like with affinities, we can give a graphical representation of effective affinities. For example, for level $|\mu| = 3$ in the hierarchy illustrated above we have three effective affinities. Let's consider only the first, which reads:

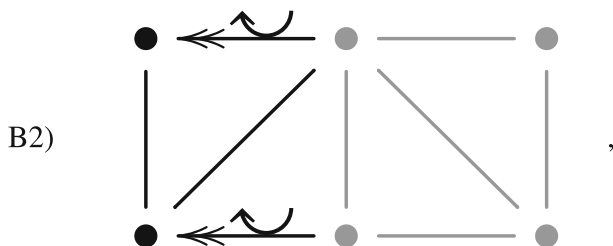
$$Q_1^{1,2,4} = \log \frac{\begin{array}{c} \bullet \leftarrow \bullet \\ \downarrow \nearrow \uparrow \searrow \\ \bullet \end{array} + \begin{array}{c} \bullet \leftarrow \bullet \\ \downarrow \nearrow \uparrow \searrow \\ \bullet \end{array} + \begin{array}{c} \bullet \leftarrow \bullet \\ \downarrow \nearrow \uparrow \searrow \\ \bullet \end{array}}{\begin{array}{c} \bullet \rightarrow \bullet \\ \uparrow \nwarrow \downarrow \swarrow \\ \bullet \end{array} + \begin{array}{c} \bullet \rightarrow \bullet \\ \uparrow \nwarrow \downarrow \swarrow \\ \bullet \end{array} + \begin{array}{c} \bullet \rightarrow \bullet \\ \uparrow \nwarrow \downarrow \swarrow \\ \bullet \end{array}}. \quad (42)$$

In a way, like each “real” affinity in Eq. (21) is defined along one fundamental cycle, cycles (more than one) still play a role in the definition of the effective affinity, though in a more involved way. The effective affinity includes all of the cycles that pass through the observable transition and that are not already “taken care for” by other observable transitions. This “dressing” of the affinity by resumming diagrams is somewhat reminiscent of the paradigm of the renormalization of particles’ masses and charges in Quantum Field Theory. It’s interesting to compare these effective affinities to the “real” affinities of such cycles $\mathcal{C} \ni \mu, \nexists \mu' \neq \mu$ that contain the observable edge under scrutiny but not all others. We can prove as a consequence of the log-sum inequality in information theory that

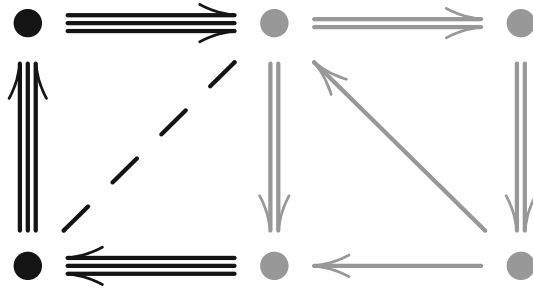
$$Q_{\mu} \leq \frac{\sum_{\mathcal{C}} \phi_{\mathcal{C}} \mathcal{A}_{\mathcal{C}}}{\sum_{\mathcal{C}} \phi_{\mathcal{C}}}, \quad (43)$$

where the $\phi_{\mathcal{C}}$ are the so-called Hill cycle currents [79,80], which provide a fundamental cycle decomposition of the stochastic process.

If instead we are in framework B2, and for example the two observable currents considered above are associated to the same reservoir so that the observer measures the sum of their values



then the theory stands on the major assumption of marginal consistency, which in this particular case requires the two effective affinities to be identical. Systems that fail to meet this condition exhibit a violation of the IFR and of the S-FDR. At stalling, where the phenomenological current vanishes (e.g. the sum of the two currents in the current example), the condition of marginal consistency has an intuitive physical interpretation: we will show in Proposition 31 that it is met if all of the microscopic edge currents contributing to a phenomenological current also vanish. For example, the following configuration of currents makes the phenomenological current vanish, but it is internally lively, which would lead the observer to estimate a vanishing effective EPR where, instead, there is effective dissipation:



Here, every arrow depicts a “quantum” of current; notice that Kirchhoff’s current law is satisfied at each site of the network. For such a system, our theory will not work. Proposition 32 makes the point that, if for all possible values of the microscopic currents there is no internal dissipation, then the theory is marginally consistent.



The main instrument we will employ is the SCGF of the marginal currents $\lambda(\{q_\mu\})$, which by the *contraction principle* in Large Deviation Theory can be obtained from that of the “complete” currents by setting $q_\alpha = 0$, for all unobserved currents $\alpha > |\mu|$. As briefly mentioned above, it is well known from Large Deviation Theory that the SCGF is the dominant Perron–Frobenius eigenvalue of the so-called *tilted operator* $M(\{q_\mu\})$, obtained from the MJPG W by augmenting the off-diagonal entries corresponding to the transitions of interest with exponential factors that depend on the *counting variables* q_μ . In general the tilted operator is not a MJPG. Nevertheless, the central result in our paper, stated in Proposition 1 (for one single edge current) and Proposition 17 (for several edge currents), shows that, letting P_{st} be the diagonal positive-definite matrix $P_{st} = \text{diag}\{p_i^{st}\}_i$, then the operator

$$\tilde{W} := P_{st} M(\{Q_\mu\})^T P_{st}^{-1} \quad (44)$$

is indeed a MJPG, which we call the *hidden time-reversal* (hidden TR) generator. Since \tilde{W} and $M(\{Q_\mu\})^T$ are similar, their common Perron eigenvalue vanishes and we obtain the *marginal IFR* for the SCGF

$$\lambda(\{Q_\mu\}) = 0. \quad (45)$$

For $|\mu| > 1$ there actually exists a continuum of values $q_\mu \equiv q_\mu^*$ for which $\lambda(\{q_\mu^*\}) = 0$. Consider for example the case $|\mu| = 2$. In this case, unless the system displays critical behavior (nonequilibrium phase transitions [81]), the SCGF is a paraboloid-like curve. Its locus of zeroes is a closed convex curve that includes $(0, 0)$, and the effective affinities $(Q_1^{1,2}, Q_2^{1,2})$. Also, where it meets with the two axis, it also includes the two effective affinities $(0, Q_2^2)$ and $(Q_1^1, 0)$ (some illustrative figures can be found on p. 74). All these values have a special

physical interpretation, and in particular they are well-behaved as comes to thermodynamically consistent parametrizations of the rates. All other values q_1^*, q_2^* are not representative of anything physical, to the best of our understanding.

A compelling question is what kind of process evolves by the hidden TR generator. We collect evidence that hidden TR “tends to preserve” the dynamics in the observable sector of the configuration space, while it “tends to invert” it in the hidden sector. We show that (Propositions 4, 17):

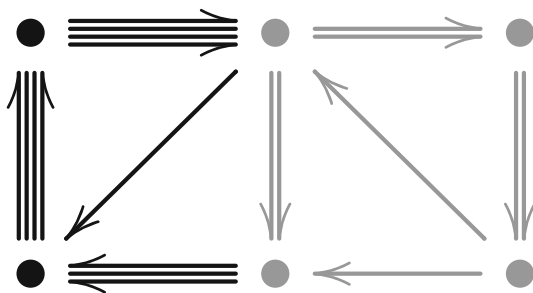
$$W =: W_{\text{mar}} + W_{\text{hid}} \quad (46a)$$

$$\tilde{W} = W_{\text{mar}} + \overline{W_{\text{hid}}} \quad (46b)$$

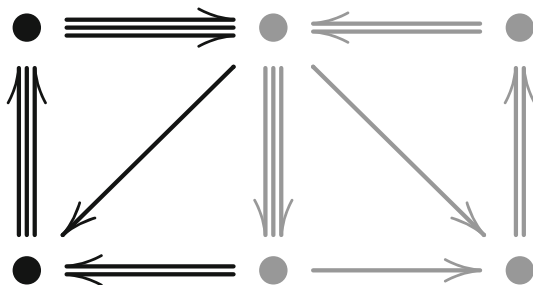
The first equation actually defines the *marginal generator* W_{mar} on the marginal edge set as that which has vanishing entries for all edges that do not belong to the observable configuration space. In the second expression there appears the time reversal of the hidden generator $\overline{W_{\text{hid}}} = P_{\text{st}} W_{\text{hid}}^T P_{\text{st}}^{-1}$. Notice that, differing from the time reversal of the full generator \overline{W} , here inversion needs to be taken with respect to the stalling steady state (which in fact is the steady state of W_{hid}). Then, the hidden TR \tilde{W} only reverses the dynamics in the hidden configuration space. Furthermore, the hidden TR construction is involutive:

$$\tilde{\tilde{W}} = W. \quad (47)$$

Marginal and hidden degrees of freedom are intertwined. In particular, as a simple consequence of Kirchhoff’s current law, one cannot modify hidden currents without affecting the observable currents. So, for example, if this is a steady configuration of currents in the forward dynamics,



then the corresponding steady configuration according to the hidden TR dynamics might look something like this (we emphasize that these are just pictorial illustrations):



Notice that the observable cycle currents maintain the same direction, while the hidden currents “tend to be reversed,” though such inversion cannot be exact otherwise Kirchhoff’s current law would be violated.

We can also consider the behavior of the other law of Kirchhoff, the loop (or cycle) law prescribing the values of the “real” affinities. We show (Proposition 7) that the hidden TR generator reverses all of the hidden “real” affinities, while it “tries to preserve” the marginal ones:

$$\tilde{\mathcal{A}}_\mu = 2Q_\mu - \mathcal{A}_\mu, \quad \mu \leq |\mu|, \quad (48a)$$

$$\tilde{\mathcal{A}}_\alpha = -\mathcal{A}_\alpha, \quad \alpha > |\mu|. \quad (48b)$$

Notice that at stalling all of the affinities are exactly reversed. In fact, at stalling the hidden TR generator coincides with the forward TR generator

$$\tilde{W}(\mathbf{x}^{\text{st}}) = \overline{W}(\mathbf{x}^{\text{st}}), \quad (49)$$

which is the analog of the detailed-balance condition Eq. (17).

Associated to the marginal dynamics is a marginal path measure \tilde{P} , in terms of which we can prove the *marginal FR*

$$\frac{P(\{\Phi_\mu\})}{\tilde{P}(\{-\Phi_\mu\})} = \exp \sum Q_\mu \Phi_\mu, \quad (50)$$

which can be equivalently stated as a marginal fluctuation symmetry as

$$\lambda(\{q_\mu\}) = \tilde{\lambda}(\{Q_\mu - q_\mu\}). \quad (51)$$

Notice the crucial difference with respect to its “complete” counterpart Eq. (27): in this case we are comparing different probability distributions, which opens up the question whether the hidden TR dynamics can be operationally defined, just like effective affinities were. From Eq. (50) we can restore the generalized RR at stalling

$$\phi_{\mu;\mu'}^{\text{st}} = \tilde{\phi}_{\mu';\mu}^{\text{st}}. \quad (52)$$

More in general, all of the higher-order response relations that characterize “complete” systems can be restored upon appropriate hidden TR. Surprisingly, we can even prove, only at long times, a inter-hierarchical FR

$$\frac{\tilde{P}^{|\mu|}(\{\Phi_\mu\}_{\mu=1}^{|\mu|})}{\tilde{P}^{|\mu'|}(\{\Phi_\mu\}_{\mu=1}^{|\mu|})} \asymp \exp \left(\sum_{\mu=1}^{|\mu|} Q_\mu^{1,\dots,|\mu|} \Phi_\mu - \sum_{\mu=1}^{|\mu'|} Q_\mu^{1,\dots,|\mu'|} \Phi_\mu \right), \quad (53)$$

where $|\mu'| > |\mu|$ and where now $\tilde{P}^{|\mu|}$ is the hidden TR associated to the $|\mu|$ -th marginal theory in the hierarchy.

Finally, we also stack one negative result to the pile. Recently an uncertainty relation connecting a current’s error and total dissipation has been proven [82–85]. However, the bound is not strict and it is only significant when the current quantifies the full dissipation. For marginal currents, it makes for a natural conjecture to speculate that the effective affinity would enter the bound in place of the “real” ones. We show in Sect. 7 that this is not the case.

2.3 Epode: Discussion and Perspectives

Let us draw some general conclusions. More technical perspectives will be discussed in Sect. 8.

In this paper we present a rather general theory of fluctuation relations and response formulas for an observer that only measures and controls a marginal subset of currents.

The context is that of the stochastic thermodynamic analysis of continuous-time, discrete configuration-space autonomous Markov “jump” processes. The theory makes some clear-cut experimental predictions, in particular the integral fluctuation relation with respect to the effective affinities, the violation of the reciprocal relations at stalling steady states, and the validity of the symmetrized fluctuation–dissipation relation. The theory is fairly complete as regards currents that account for single transitions in configuration space, and it also holds for phenomenological currents that are linear combinations of edge currents, provided the additional requirement of marginal consistency is met, which is analogous to local detailed balance, but stricter. Therefore, the most imminent open question left aside is what kind of systems satisfy marginal consistency, and if a system does not, how does the surplus of entropy production at the stalling states affect response.

Central objects in the theory are the effective affinities. While they are mathematically expressed in terms of the rates of the Markovian dynamics all over the configuration space, an operational procedure allows to evaluate them without full knowledge of the transition rates, provided the parameters that the observer controls are known to only affect the rates corresponding to the measurable degrees of freedom. If this is not the case, then one can turn the story the other way around, and use the predictions of our theory as a test of locality of the physical parameters.

The full fluctuation relation and the reciprocal relations can be reinstated upon the identification of a suitable hidden time-reversed Markovian dynamics. The question is open whether to obtain such dynamics one needs to be able to micro-engineer all rates, in which case the latter relations remain only formal, or else, as is the case for the effective affinities, whether there exists a phenomenological procedure to determine the dynamics. This would unlock a new set of experimental predictions of our theory. A test-bed for this possibility is that of a computational experiment: is it possible to program a Gillespie simulation of the hidden time-reversed dynamics without specifying all of the rates as an input, but rather by performing a smaller transformation of the known parameters with respect to the simulations of the forward dynamics? We are not yet in the position to give a definitive answer to this question.

An important consideration is that ours is not a kinetic theory, that is, it does not provide a procedure to coarse-grain the dynamics in the hidden sector of the configuration space in order to obtain an effective dynamics in the marginal configuration space. While the *gedanken*-observer described in Ref. [50] and Sect. 4.5 does cook up a marginal dynamics that explains his steady-state observations, in no way this dynamics is representative of the finite-time behavior, including such questions as the rate of convergence to the steady state, first exit times out of the hidden sector etc. Furthermore, our theory does not involve an exquisitely dynamical, but physically relevant limiting situation, that of time-scale separation between the marginal and the hidden degrees of freedom. The relationship between our theory and various other approaches, such as those described in Refs. [23,36], is an interesting territory to explore.

The results that we presented are amenable to several generalizations. To lattice gas models, where response theory is enriched by all aspects regarding the spatial disposition of particles [45]. To Markov jump processes on infinite configuration spaces, in particular population dynamics, chemical reaction kinetics, and reaction-diffusion theory. To diffusion processes on continuous configuration spaces. To time-periodic processes, rather than stationary, and more generally to time-dependent perturbations, to systems with resetting [86]. To finite-time response to a sudden perturbation $x_\mu(t) = \theta(t)(x_\mu - x_\mu^{\text{st}})$ (θ being Heaviside’s step-function), or to perturbations that are modulated in a finite-time interval. Periodicity calls for a study in Fourier space, where response relations incarnate into susceptibilities and

spectral response functions, and where it is already known that far from equilibrium several of the equilibrium results are violated [87,88].

Further questions on marginal and effective theories are genuinely thermodynamic. As a matter of fact, any question addressed in recent years in the field can be turned marginal: the study of efficiency and efficiency fluctuations, of the linear regime where (marginal) currents are approximately linear in the (effective) affinities, of transduction [79], of variational principles such as the minimum and maximum entropy production principles [89,90]. Recent models cope with strong system-environment interactions by envisaging the system as a subsystem of a larger system-environment complex, itself weakly interacting with its surrounding. Again, such system-environment complex could be analyzed in terms of our theory. Systems that have irreversible transitions, such as stochastic processes with resetting, always posed a challenge, because the thermodynamic force diverges along irreversible transitions; one way out, among others [91], could be to dump the irreversible transitions into the hidden trash bin.

The observables that we consider, the currents, are antisymmetric under time reversal. A new central paradigm is that response out of equilibrium depends in a crucial way on the *activity* of the system, i.e. some measure of the gross amount of stuff flowing, in opposition to the current that measures the net amount of stuff delivered. Many recent results regarding currents have been generalized to flows and other symmetric quantities, e.g. the uncertainty relation briefly mentioned above [92], fluctuation relations [93], and response formulas [94–96].

Violations of the reciprocal relations are often associated with broken time-reversal symmetry, e.g. the microscopic dynamics involves axial fields that are antisymmetric under time reversal, such as magnetic fields, Coriolis forces, the momentum variable in underdamped Brownian motion etc. It is well known that in these cases the Onsager symmetry can be restored upon inversion of the axial fields. With an eye on Eq. (52), it is tempting to speculate that the marginal dynamics might be the discrete analog of the axial-field inversion. The analysis of proper time reversal of continuous noisy systems with even and odd variables with respect to the FR has been broadly studied [97]. If our speculation is fruitful, it would allow to include even and odd variables within the formalism of Markov jump-processes without additional requirements.

Local observers are reminiscent of the theory of relativity. Some authors have considered [98,99] fluctuation relations in moving frames where a “local equilibrium” can be attained by a privileged observer. It would be interesting to inspect whether such transformations could be framed within our theory of a marginal observer.

One interesting aspect that is completely missing is that of *duality*, whereby one swaps the role of the marginal and the hidden state spaces. Is there any relation between the theories so obtained? We notice in passing that by first performing the hidden time reversal, and then the dual hidden time reversal, one does *not* obtain the “complete” time reversal of the forward dynamics. Thus, if a relation exists, it might be subtle.

Let us conclude with some more epistemological remarks. At all stages we insisted on its operational character. We also revived gauge invariance as a solution to the “dilemma of the observer”. This is because we strongly believe that physics is not about properties of some absolute “thing in itself,” but it rather deals with relations and processes, and about how an idealized observer interprets his/her observations. Furthermore, a priori there is no reason to presume that a system is “complete”. In our view, theories are always marginal to some extent—in particular statistical physics is intrinsically a theory of incomplete information. For this reason we always comment the words “complete” and “real”.

3 Episode 1: Setup

Here we provide a compendium of the thermodynamic analysis of “complete” systems evolving by a Markovian jump-process dynamics on a graph, to set the notation, introduce the basic techniques, and provide numerous references to more in-depth studies. While the knowledgeable reader might safely skip this section, the newbie should not be discouraged either, as in the following sections we will attempt to construct our theory in a pedagogical and self-contained manner. The ornament on p.45 divides the old material from the new one.

3.1 Algebraic Graph Theory in a Pistachio-Shell

The system’s configuration space is a finite oriented graph $\mathcal{G} = (\mathcal{I}, \mathcal{E}, \partial)$ with a number $|\mathcal{I}|$ of sites $i, j, \dots \in \mathcal{I}$ connected by $|\mathcal{E}|$ oriented edges $ij \in \mathcal{E}$, corresponding to the possible transitions between sites. We assume that the graph is connected, without loops nor multiple edges between two sites.⁶ We assign an orientation to the edges $ij = i \leftarrow j$ by prescribing an arbitrary order relation⁷ $i < j$.

The *incidence matrix* $\partial : \mathbb{R}^{|\mathcal{E}|} \rightarrow \mathbb{R}^{|\mathcal{I}|}$, prescribing which sites are boundaries of which edges, has entries $\partial_k^{ij} = \delta_k^i - \delta_k^j$. Square matrices defined on the configuration space \mathcal{I} of a system are denoted $A : \mathbb{R}^{|\mathcal{I}|} \rightarrow \mathbb{R}^{|\mathcal{I}|}$, and they act on vectors $\vec{v} \in \mathbb{R}^{|\mathcal{I}|}$. All other vectors, including those living in the linear space generated by edges $\mathbb{R}^{|\mathcal{E}|}$, are denoted in bold \mathbf{v} .

A spanning tree $\mathcal{T} \subseteq \mathcal{E}$ is a collection of $|\mathcal{I}| - 1$ (unoriented) edges that connect all sites. In a rooted spanning tree \mathcal{T} edges are oriented in such a way that there is a unique directed path leading from any site to i . An oriented cycle is a succession of oriented edges such that at every site there are as many incoming edges as outgoing ones. A cycle is *simple*, and it is denoted \mathcal{C} , when it has no crossings. A cycle can be algebraically identified as a right-null vector \mathbf{c} of the incidence matrix, $\partial \mathbf{c} = 0$.

3.2 Master Equation Dynamics

3.2.1 Master Equation

We assign to each edge time-independent positive transition rates w_{ij} and w_{ji} of jumping respectively from j to i and from i to j , and we let $w_i := \sum_j w_{ji}$ be the total escape rate out of site i . Let $\vec{p}(t) = (p_i(t))_i$ be the vector of probabilities of being at a site at a given time t , sometimes called *ensemble*. Given the initial ensemble $\vec{p}(0) = \vec{p}^0$, $\vec{p}(t)$ obeys the master equation Eq. (15). Entries along columns of W add up to zero, $W^T \vec{1} = 0$, where T is matrix transposition and $\vec{1}$ is the vector with all entries equal to unity. The master equation can be cast as a continuity equation

$$\frac{d}{dt} \vec{p}(t) + \partial \phi(t) = 0 \quad (54)$$

⁶ This assumption excludes the possibility of resolving multiple transitions, which is crucial in stochastic thermodynamics, especially in the light of the assumption of local detailed balance [100] whereby different reservoirs enhance transitions. The generalization of all of our results is straightforward, but it makes the notation overly baroque, to the detriment of clarity. We discuss it in Sect. 4.10.

⁷ This is just one way to introduce an arbitrary orientation of the edges. Not all orientations come from an order relation.

in terms of the vector of currents $\phi(t) = (\phi_{ij}(t))_{ij \in \mathcal{E}}$ with entries $\phi_{ij}(t) = \psi_{ij}(t) - \psi_{ji}(t)$, where

$$\psi_{ij}(t) := w_{ij} p_j(t) \quad (55)$$

is sometimes called the mean flux from site j to i .

3.2.2 Steady Ensemble

Assuming that the graph is connected and that rates are non-negative, then the system tends to a steady ensemble $\vec{p} = \lim_{t \rightarrow \infty} \vec{p}(t)$ that is the unique right-null vector of the generator, and that makes the steady currents $\phi = \lim_{t \rightarrow \infty} \phi(t)$ “divergenceless”:

$$0 = W\vec{p} = -\partial\phi. \quad (56)$$

The right-hand side of this equation is Kirchoff’s current law. It is well known that, up to a normalization factor, the steady ensemble can be found in terms of minors of the MJPG [69,101]

$$p_i \propto (-1)^{i+j} \det W(j|i) \quad (57)$$

where $W(j|i)$ is the matrix obtained by removing the j -th row and the i -th column; the above expression holds independently of j . A one-line proof of this fact is as follows: since the determinant of W vanishes (its null eigenvector being the steady state), expanding with the Laplace formula along the j -th column $0 = \det W = \sum_{i \neq j} w_{ij} (-1)^{i+j} W(j|i) - (-1)^{i+j} w_i W(i|i)$, and we conclude \square .

The steady ensemble can be expressed in terms of rooted oriented spanning trees as

$$p_i = \frac{\tau_i(\mathcal{G})}{\tau(\mathcal{G})} \quad (58)$$

where

$$\tau_i(\mathcal{G}) := \sum_{\mathcal{T} \subseteq \mathcal{E}} \prod_{ij \in \mathcal{T}} w_{ij}, \quad (59)$$

is the so-called spanning-tree polynomial, where \mathcal{T} ranges over oriented spanning trees with root in i , and $\tau(\mathcal{G}) = \sum_i \tau_i(\mathcal{G})$ is the normalization. Since rates w_{ij} have dimensions of an inverse time, we have a liberty in the choice of the time unit. We choose to spend this liberty by setting, unless otherwise stated,

$$\tau(\mathcal{G}) \stackrel{!}{=} 1. \quad (60)$$

The equivalence between Eqs.(57) and (58) is an instance of the matrix-tree theorem in algebraic graph theory, an important paradigm that will play a major role in the physical interpretation of our results, in particular when we will consider portions of the configuration space, a case that is covered by the crucial all-minors matrix-tree theorem for weighted oriented graphs [103].

A steady state is said to be an equilibrium if it satisfies the condition of detailed balance

$$\psi_{ij}^{\text{eq}} = w_{ij} p_j^{\text{eq}} = w_{ji} p_i^{\text{eq}} = \psi_{ji}^{\text{eq}}. \quad (61)$$

Hence at equilibrium the steady currents vanish. Equilibrium steady states admit a particularly simple expression in terms of rate ratios:

$$p_k^{\text{eq}} = \left[\sum_l \prod_{ij \in \gamma_{k \leftarrow l}} \frac{w_{ij}}{w_{ji}} \right]^{-1}, \quad (62)$$

where $\gamma_{k \leftarrow l}$ is an arbitrary connected path leading from state l to k .

3.2.3 Time Reversal

The generator of the time-reversed dynamics is defined as follows. Given a forward generator W , compute its steady state \vec{p} , construct the diagonal matrix P whose diagonal entries are the steady-state probabilities, $P := \text{diag} \{p_i\}_i$. Then the TR generator is

$$\overline{W} := P W^T P^{-1} \quad (63)$$

It can easily be shown that \overline{W} is indeed a MJPG, with, among others [104], the following properties: same spectrum as W , same exit frequencies out of configurations, same steady state, all inverted steady-state currents. A system satisfies detailed balance if and only if $W = \overline{W}$.

3.3 Master Equation Thermodynamics

We hereby consider mean currents. The stochastic counterpart is covered in the next subsection.

3.3.1 Observational Currents

Currents of observational interest are linear combinations of the edge currents,

$$\phi_\alpha = \sum_{ij} \varphi_\alpha^{ij} \phi_{ij}, \quad (64)$$

where \sum_{ij} sums over edges, while $\sum_{i,j}$ sums over couple of sites. By the handshaking lemma in graph theory, $\sum_{i,j} = 2 \sum_{ij}$ for any summand.

The antisymmetric weight factor $\varphi_\alpha^{ij} = -\varphi_\alpha^{ji}$ prescribes by what amount the α -th current increases [decreases] when transition $i \leftarrow j$ [$j \leftarrow i$] is performed. We will describe the conditions upon which such currents form a complete set later in this section. Observational currents might either be subsets of the edge currents, in which case there exists some particular edge ij_α such that $\varphi_\alpha^{ij} = \delta_{ij,ij_\alpha}$, or otherwise they are *phenomenological*, which means that at least one such current is supported on more than one edge. The first will be the subject study of Sects. 4 and 5, the second of Sect. 6. While we use a unified notation for all observational currents, the treatment of phenomenological currents poses specific problems. Of course, the observer might stipulate that *all* edge currents are of observational interest, in which case α is a multi-index $\alpha = i'j'$, with $\varphi_\alpha^{ij} = \delta_{i',i}^i \delta_{j',j}^j$.

3.3.2 Forces, Entropy Production Rate and Local Detailed Balance

Thermodynamic reasoning ensues when one complements the dynamical information contained in currents with conjugate *forces* that quantify the cost of transitions. The steady-state force associated to a particular transition is given by

$$\mathcal{F}_{ij} := \log \frac{w_{ij} p_j}{w_{ji} p_i} = \log \frac{\psi_{ij}}{\psi_{ji}}. \quad (65)$$

Clearly, a system has all vanishing steady forces if and only if it satisfies the condition of detailed balance Eq. (61).

The mean steady-state EPR is defined as

$$\sigma := \sum \phi_{ij} \mathcal{F}_{ij} = \sum (\psi_{ij} - \psi_{ji}) \log \frac{\psi_{ij}}{\psi_{ji}} \geq 0. \quad (66)$$

When working with phenomenological currents, to achieve a thermodynamically consistent treatment the corresponding thermodynamic forces cannot be arbitrary, rather they need to enjoy certain symmetries, in such a way that ultimately the steady entropy production rate can be expressed in terms of the phenomenological currents and forces only. Thermodynamic consistency (or simply, consistency) is realized if the following condition of *local detailed balance* holds

$$\mathcal{F}_{ij} = \sum \varphi_{ij}^\alpha \mathcal{F}_\alpha + a_j - a_i \quad (67)$$

where the \mathcal{F}_α are observational thermodynamic forces. For the sake of generality we included an arbitrary function of the configuration a_i that takes into account the liberty offered by Kirchhoff's current law at steady states (see next section) and the steady state. Under this condition the steady EPR writes

$$\sigma = \sum \phi_\alpha \mathcal{F}_\alpha. \quad (68)$$

Notice that any dependence on a_i is lost. Eq.(67) is not just a convenient physically meaningful parametrization of the rates, it actually imposes constraints on the space of possible rates, that can be interpreted as *symmetries* that the edge forces must satisfy. The number of these symmetries eventually increases if phenomenological currents are not linearly independent, i.e. if $\sum \ell_\alpha \varphi_{ij}^\alpha = 0$ for some vector $(\ell_\alpha)_\alpha$. Symmetries and conservation laws arise from the interplay between the definition of the phenomenological currents, an information contained in φ_{ij}^α , and the structure of the network, an information contained in ∂ [71,73]. In this work we exclude the possibility of linearly dependent currents, though it would make for an interesting problem to investigate our results in those situations where a conserved current flows across the marginal/hidden configuration space.

3.4 Cycle Analysis

Cycles are ubiquitous in thermodynamics. For example, cycles solve Kirchhoff's current law, hence they are useful to describe steady states. While the focus is usually on the methods described in Schnakenberg's review [69], where the freedom in the choice of a cycle basis of $\ker \partial$ is broken to provide a compact expression for the EPR, there exists another cycle decomposition that is less compact but more general, and which will turn out to play an important

role. In the following we will (somewhat improperly) refer respectively to Schnakenberg's and Hill's analysis.

3.4.1 Schnakenberg Analysis

Consider an arbitrary spanning tree \mathcal{T} . There are $|\mathcal{E}| - |\mathcal{T}| + 1$ edges $ij_\alpha = i_\alpha \leftarrow j_\alpha \in \mathcal{E} \setminus \mathcal{T}$, called *chords*, that do not belong to the spanning tree. Adding a chord to \mathcal{T} generates a unique simple cycle \mathcal{C}_α , that can be oriented along the direction of ij_α . To such cycle we associate a vector \mathbf{c}_α with entries

$$c_\alpha^{ij} = \begin{cases} +1, & \text{if } ij \in \mathcal{C}_\alpha \\ -1, & \text{if } ji \in \mathcal{C}_\alpha \\ 0 & \text{otherwise} \end{cases} \quad (69)$$

The set of simple oriented cycles so generated forms a basis for the right null-space of the incidence matrix, $\partial \mathbf{c}_\alpha = 0$. Hence, in view of Eq. (56), cycles describe steady states. Transient states can be studied in terms of cocycles [70] (see below). In particular, we define *chord currents*, obtained by setting

$$\varphi_\alpha^{ij} = \begin{cases} +1, & \text{if } ij = ij_\alpha \\ -1, & \text{if } ji = ij_\alpha \\ 0 & \text{otherwise} \end{cases} \quad (70)$$

and their conjugate cycle affinities, defined as

$$\mathcal{A}_\alpha := \mathcal{A}(\mathcal{C}_\alpha) := \sum_{ij \in \mathcal{C}_\alpha} \mathcal{F}_{ij} = \log \prod_{ij \in \mathcal{C}_\alpha} \frac{w_{ij}}{w_{ji}}. \quad (71)$$

Notice that the steady state disappears from the final expression. Also, because $\sum_{ij} c_\alpha^{ij} \varphi_{\alpha'}^{ij} = \delta_{\alpha, \alpha'}$, consistently with the condition of local detailed balance Eq. (67) we obtain $\mathcal{A}_\alpha = \mathcal{F}_\alpha$ and thus the EPR only writes in terms of cycle observables $\sigma = \sum \phi_\alpha \mathcal{A}_\alpha$.

We conclude this section by providing an interesting expression for the cycle affinities. Consider the system obtained by removing all of the chords, *ossia* by setting their rates to zero. Because it has no cycles, it satisfies detailed balanced. Let \vec{p}^{eq} be its equilibrium steady state. Then, it is easy to show that

$$\mathcal{A}_\alpha = \log \frac{w_{ij_\alpha} p_{j_\alpha}^{\text{eq}}}{w_{ji_\alpha} p_{i_\alpha}^{\text{eq}}}. \quad (72)$$

It follows from the fact that, due to the property of detailed balance, by Eq. (62) the equilibrium state obeys

$$\frac{p_k^{\text{eq}}}{p_l^{\text{eq}}} = \prod_{ij \in \gamma_{k \leftarrow l}} \frac{w_{ij}}{w_{ji}}. \quad (73)$$

where $\gamma_{k \leftarrow l} \subseteq \mathcal{T}$ is now unique.

3.4.2 Hill Analysis

As a consequence of the spanning-tree expression for the steady state Eq. (58), the steady-state current along edge ij can be written as a sum over all simple oriented cycles that contain edge ij

$$\phi_{ij} = \sum_{\mathcal{C} \ni ij} (\psi_{+\mathcal{C}} - \psi_{-\mathcal{C}}), \quad (74)$$

where the *cycle fluxes* are given by

$$\psi_{\pm\mathcal{C}} = \theta_{\mathcal{C}} \prod_{ij \in \pm\mathcal{C}} w_{ij}. \quad (75)$$

The factor $\theta_{\mathcal{C}} = T_{\mathcal{C}}(\mathcal{G}/\mathcal{C})$ is a rooted oriented spanning tree polynomial over the graph obtained by contracting cycle \mathcal{C} to a unique vertex. Importantly, as observed in Ref. [105], it is symmetric under inversion of the cycle's orientation. Therefore the cycle affinity can be written as

$$\mathcal{A}(\mathcal{C}) = \log \frac{\psi_{+\mathcal{C}}}{\psi_{-\mathcal{C}}}. \quad (76)$$

With a few passages one obtains for the steady-state entropy production rate

$$\sigma = \sum_{\mathcal{C}} (\psi_{+\mathcal{C}} - \psi_{-\mathcal{C}}) \log \frac{\psi_{+\mathcal{C}}}{\psi_{-\mathcal{C}}} \quad (77)$$

where it is stipulated that each cycle is summed over only once, along one arbitrary choice of its orientation (otherwise a factor 1/2 should be included).

We call the quantity

$$\sigma_{ij}^{\text{Hill}} = \sum_{\mathcal{C} \ni ij} (\psi_{+\mathcal{C}} - \psi_{-\mathcal{C}}) \log \frac{\psi_{+\mathcal{C}}}{\psi_{-\mathcal{C}}} \quad (78)$$

the *local EPR* associated to Hill's cycle decomposition.

3.4.3 Cocycles

By the rank-nullity theorem, the edge vector space $R^{|\mathcal{E}|}$ of a graph can be decomposed in a basis of cycles \mathbf{c} , which span the null vectors of the incidence matrix ∂ , and of *cocycles* (also known in graph theory as *cuts* or *bonds*) \mathbf{c}_{α^*} , which span the image of ∂ . By construction cocycles are orthogonal to cycles

$$\sum_{ij} c_{\alpha^*}^{ij} c_{\alpha}^{ij} = 0. \quad (79)$$

Simple cocycles are minimal sets of edges whose removal disconnects the graph into two subgraphs. The algebra of cycles and cocycles and their relationship to thermodynamics has been studied to great extent by one of the Authors in Ref. [70]. Cocycles play a role in characterizing transient states. In this work, they will play a minor role related to gauge invariance, see Sect. 5.7.

3.5 Stochastic Thermodynamics

3.5.1 Trajectories and Their Measure

We consider a single realization of a jump-process,⁸ described by the trajectory

$$\omega^t := i_0 \xrightarrow{t_0} i_1 \xrightarrow{t_1} \dots \xrightarrow{t_{N-1}} i_N \xrightarrow{t_N}, \quad (80)$$

which performs N jumps in the time arc $[0, t]$. It is described by the succession of sites visited⁹ and by that of waiting times at sites, which add up to $\sum_{n=0}^N t_n = t$. The trajectory p.d.f. is given by¹⁰

$$P[\omega^t] = e^{-w_{i_N} t_N} \prod_{n=0}^{N-1} (w_{i_{n+1}, i_n} e^{-w_{i_n} t_n}) p_{i_0}^0 \quad (81)$$

with respect to the trajectory integration measure

$$\int \mathcal{D}\omega^t = \sum_N \sum_{i_0, \dots, i_N} \int_0^t \prod_{n=0}^N dt_n \delta\left(t - \sum_n t_n\right). \quad (82)$$

The remarkable review Ref. [108] proposes two derivations of the measure over realizations of Markov jump processes, one heuristic, based on the intuitive mechanisms of the Gillespie algorithm, and one exact, based on the Laplace transform of the propagator (see also [109] for analytical inversion formulas). The time-reversed trajectory is

$$\bar{\omega}^t := i_N \xrightarrow{t_N} i_{N-1} \xrightarrow{t_{N-1}} \dots \xrightarrow{t_1} i_0 \xrightarrow{t_0}. \quad (83)$$

The following fact, which is a fundamental (and often underestimated) passage, will play an important role in our discussion, because it is one of the few missing links that differentiates complete and marginal thermodynamics. Time reversal of the dynamics induces a time-reversed measure over trajectories \bar{P} , such that the time-reversed p.d.f. of a trajectory, conditioned on the initial configuration, coincides with the forward probability density of the time-reversed trajectory, conditioned on the final configuration:

$$\bar{P}(\omega^t | i_0) = P(\bar{\omega}^t | i_N). \quad (84)$$

3.5.2 Time-Integrated Currents

We now introduce the stochastic observables of interest, defined along a trajectory ω^t . Letting $\tau_n = \sum_{n' < n} t_{n'}$ be the time elapsed before the n -th jump, the time-integrated *edge fluxes*

$$\Psi_{ij}^t := \Psi_{ij}[\omega^t] := \int_0^t d\tau \sum_n \delta(\tau - \tau_n) \delta_{i_{n+1}, i} \delta_{i_n, j} \quad (85)$$

⁸ For an exact and explicit mathematical characterization of the process as a solution of a stochastic equation in terms of so-called random time-change Poisson processes with intensity depending on the history of the process, see [106] in the specific case of the chemical master equation and [107] for general Markov jump processes.

⁹ Notice that, had we resolved multiple edges, we should characterize a trajectory not by the sites visited but rather by the edges.

¹⁰ More precisely, what follows is a family of p.d.f.'s labeled by N , the number of arguments which depends on total number of jumps N , itself a stochastic variable.

count the net number of times transition $j \rightarrow i$ occurs. The integrand in Eq. (85) is a spiking function of time [64]. Edge currents are defined as

$$\Phi_{ij}^t := \Psi_{ij}^t - \Psi_{ji}^t \quad (86)$$

Currents are antisymmetric under time reversal $\Phi[\omega^t] = -\Phi[\omega]$. The corresponding symmetrized quantities are sometimes called *activities*. Activities will not play a role in our theory.

3.5.3 Currents' p.d.f. and Scaled Cumulants

We will denote by $P^t(\{\Phi_\alpha\})$ the density function of the probability that the currents take values $\Phi_\alpha^t \in [\Phi_\alpha, \Phi_\alpha + d\Phi_\alpha]$ (we will drop the superscript t from now on). Formally,

$$P(\{\Phi_\alpha\}) = \int \mathcal{D}\omega^t P(\omega^t) \prod_\alpha \delta(\Phi_\alpha[\omega^t] - \Phi_\alpha). \quad (87)$$

It can be shown that cumulants of the time-integrated currents grow linearly in time. Let us then introduce the (signed) *scaled-cumulant generating function* (SCGF),

$$\lambda(\{q_\alpha\}) := \lim_{t \rightarrow \infty} \frac{1}{t} \log \int e^{-\sum_\alpha q_\alpha \Phi_\alpha} P(\{\Phi_\alpha\}) \prod_{\alpha'} d\Phi_{\alpha'} \quad (88)$$

and the scaled cumulants

$$\phi_{\beta_1, \dots, \beta_{|\alpha|}} := (-1)^{\beta_1 + \dots + \beta_{|\alpha|}} \frac{\partial^{\beta_1}}{\partial q_1^{\beta_1}} \dots \frac{\partial^{\beta_{|\alpha|}}}{\partial q_{|\alpha|}^{\beta_{|\alpha|}}} \lambda(\{q_\alpha\}) \Big|_{\{q_\alpha\}=\{0\}}. \quad (89)$$

SCGFs can be calculated as the ‘‘Perron–Frobenius’’ eigenvalue (i.e. unique eigenvalue with largest real part) of the so-called *tilted operator*¹¹ $M(\{q_\alpha\})$. Defining the tilting matrix $T(\{q_\alpha\})$ with entries

$$T(\{q_\alpha\})_{ij} := e^{\sum_\alpha q_\alpha \varphi_{ij}^\alpha} \quad (90)$$

the tilted operator is defined as

$$M(\{q_\alpha\}) := W \circ T(\{q_\alpha\}). \quad (91)$$

where \circ is the entry-wise Hadamard product. For derivations, see [64] for cycle currents, [81] for one-dimensional lattice gases. For useful implicit function techniques for the calculation of the cumulants, we will make use of Refs. [110–112]. Importantly, the SCGF for the edge currents, and of any linear contraction of them, is a convex function. An accessible proof of this fact can be found in [113]).

Finally, the contraction principle in large deviation theory allows to derive the SCGF of a coarse-grained observable from that of more specialized observables. For definiteness, suppose we know the SCGF $\lambda_{\mathcal{E}}(\{q_{ij}\})$ of *all* the edge currents (which, as a matter of fact, is usually inaccessible). Then the rate function for the observable currents reads

$$\lambda(\{q_\alpha\}) = \lambda_{\mathcal{E}}\left(\left\{\sum_{ij} \varphi_{ij}^\alpha q_\alpha\right\}\right). \quad (92)$$

¹¹ We call it *tilted operator*, rather than *generator*, because it does not generate a Markov jump-process dynamics.

3.5.4 Fluctuation Relations for the Currents

The forward and backward p.d.f. obey the FR

$$\frac{P(\omega^t)}{P(\bar{\omega}^t)} = \frac{p_{i_0}^0}{p_{i_N}^0} \frac{p_{i_N}^{\text{eq}}}{p_{i_0}^{\text{eq}}} \exp \sum \mathcal{A}_\alpha \Phi_\alpha^t \quad (93)$$

where we $(\bar{p}_i^0)_i$ is the initial sstate from which the backward trajectory is sampled, which in general can be different from that from which the forward trajectory is sampled. This relation holds for cycle currents, and for phenomenological currents provided thermodynamic consistency is respected. The appearance of the equilibrium state in the above equation is far from trivial, and it is related to a cycle/cocycle decomposition of the EPR [62,114], or equivalently to the expression Eq. (72) for the affinities.

The multivariate FR for the cycle currents Eq. (4) is then obtained by taking the marginal for the currents, either asymptotically in time for any initial ensemble, where the boundary terms become subdominant with respect to the entropy production, or at all times on the assumption that the trajectories are sampled with the equilibrium ensemble $\bar{p}^0 = \bar{p}^t = \bar{p}^{\text{eq}}$. The IFR Eq. (5) is easily obtained by integrating the above FR.



In fact, a wider class of such IFRs can be generated as follows. We consider here subsets of currents, namely the first $|\mu|$ currents and of the last $|\alpha| - |\mu|$. We have from the FR Eq. (4)

$$P(\{\Phi_\alpha\}) \exp \sum_{\alpha \leq |\mu|} \Phi_\alpha \mathcal{A}_\alpha = P(\{-\Phi_\alpha\}) \exp \sum_{\alpha > |\mu|} \Phi_\alpha \mathcal{A}_\alpha. \quad (94)$$

Integrating over currents and employing antisymmetry, we obtain the *generalized IFR*

$$\left\langle \exp \sum_{\mu \leq |\mu|} \Phi_\mu^t \mathcal{A}_\mu \right\rangle = \left\langle \exp \sum_{\alpha > |\mu|} \Phi_\alpha^t \mathcal{A}_\alpha \right\rangle. \quad (95)$$

In particular, for two currents, for $|\mu| = 1, 2$ we obtain

$$\left\langle e^{\mathcal{A}_1 \Phi_1^t + \mathcal{A}_2 \Phi_2^t} \right\rangle = 1 \quad (96a)$$

$$\left\langle e^{\mathcal{A}_1 \Phi_1^t} \right\rangle = \left\langle e^{\mathcal{A}_2 \Phi_2^t} \right\rangle. \quad (96b)$$

The latter we call the reciprocal IFR. As a minor side argument, let us now provide a simple remark highlighting that the nonequilibrium reciprocal relations are a fundamental complement to the IFR. Let us rewrite Eq. (96a) by conditioning over the second current:

$$\left\langle \left\langle e^{-\mathcal{A}_1 \Phi_1^t} \middle| \Phi_2^t \right\rangle e^{-\mathcal{A}_2 \Phi_2^t} \right\rangle = 1. \quad (97)$$

Now suppose that the currents Φ_1^t and Φ_2^t are independent processes. Then

$$\left\langle e^{-\mathcal{A}_1 \Phi_1^t} \right\rangle \left\langle e^{-\mathcal{A}_2 \Phi_2^t} \right\rangle = 1 \quad (98)$$

Notice that this equation alone is not sufficient to grant that each individual current satisfies the IFR itself, unless the reciprocal IFR also holds true.

3.5.5 Response Relations

We now consider the response of a system at equilibrium to perturbations of the thermodynamic forces. We assume a thermodynamic parametrization as described in assumption A0 on p. 16. While usually only linear response coefficients are considered, higher-order response relations, which were derived by Andrieux and Gaspard [66,67] and recently reformulated in an alternative and pedagogical way in Ref. [68]. The coefficients describing response at order $m = \sum m_\alpha$ of cumulants of order $n = \sum n_\alpha$ are given by

$$\phi_{\{n_\alpha\};\{m_\alpha\}} := \frac{\partial^{m_1}}{\partial x_1^{m_1}} \cdots \frac{\partial^{m_{|\alpha|}}}{\partial x_{|\alpha|}^{m_{|\alpha|}}} \phi_{n_1, \dots, n_{|\alpha|}}^{\text{eq}} \quad (99)$$

$$= (-1)^n \prod_\alpha \frac{\partial^{m_\alpha}}{\partial x_\alpha^{m_\alpha}} \frac{\partial^{n_\alpha}}{\partial q_\alpha^{n_\alpha}} \lambda(\{0\}; \mathbf{x}). \quad (100)$$

To produce relations between response coefficients, we take $m + n$ total derivatives of the fluctuation symmetry Eq. (18)

$$\prod_\alpha \frac{d^{m_\alpha}}{dx_\alpha^{m_\alpha}} \frac{d^{n_\alpha}}{dq_\alpha^{n_\alpha}} [\lambda(\{q_\alpha\}; \mathbf{x}) - \lambda(\{x_\alpha - x_\alpha^{\text{eq}} - q_\alpha\}; \mathbf{x})] = 0 \quad (101)$$

and evaluate at $q_\alpha = 0$, $\mathbf{x} = \mathbf{x}^{\text{eq}}$. Notice that the m -th derivative with respect to x of a function $f(x) = g(x - x', x)$ is

$$\frac{d^m f}{dx^m} = \sum_{k=0}^m \binom{m}{k} \frac{\partial^k}{\partial y_1^k} \frac{\partial^{m-k}}{\partial y_2^{m-k}} g(y_1, y_2) \Big|_{\substack{y_1 = x - y \\ y_2 = x}}. \quad (102)$$

We then obtain

$$\begin{aligned} \phi_{\{n_\alpha\};\{m_\alpha\}}^{\text{eq.}} &= \sum_{\{k_\alpha\}}^{\{m_\alpha\}} \left[\prod_\alpha \binom{m_\alpha}{k_\alpha} \frac{\partial^{n_\alpha+k_\alpha}}{\partial q_\alpha^{n_\alpha+k_\alpha}} \frac{\partial^{m_\alpha-k_\alpha}}{\partial x_\alpha^{m_\alpha-k_\alpha}} \right] \lambda(\{0\}, \mathbf{x}) \\ &= \sum_{\{k_\alpha\}}^{\{m_\alpha\}} (-1)^{\sum (n_\alpha+k_\alpha)} \prod_\alpha \binom{m_\alpha}{k_\alpha} \phi_{\{n_\alpha+k_\alpha\};\{m_\alpha-k_\alpha\}}^{\text{eq.}} \end{aligned} \quad (103)$$

The same set of response coefficients obeys several different response relations for different values of m_α , n_α , at fixed order $m_\alpha + n_\alpha$, for all α . Then the complete set of response relations at order $\{m_\alpha + n_\alpha\}$, $\forall \alpha$ is given by

$$\begin{aligned} &\sum_{\{k_\alpha\} \neq \{0\}}^{\{m_\alpha\}} (-1)^{k_\alpha} \prod_\alpha \binom{m_\alpha}{k_\alpha} \phi_{\{n_\alpha+k_\alpha\};\{m_\alpha-k_\alpha\}}^{\text{eq.}} \\ &= \begin{cases} -2\phi_{\{n_\alpha\};\{m_\alpha\}}^{\text{eq.}}, & \text{if } \sum_\alpha n_\alpha \text{ odd} \\ 0, & \text{if } \sum_\alpha n_\alpha \text{ even} \end{cases}, \\ &\text{for } \{n_\alpha = 1 \dots, m_\alpha\}. \end{aligned} \quad (104)$$

In the above line we isolated the contribution that comes from setting all $k_\alpha = 0$. Notice that the term on the right-hand side is the lowest-order response coefficient. In general, not all of these response relations are independent.

Let us provide a specific example by compiling tables of coefficients for response relations of a system with two currents, at third order. The upper-left entry labeling the table is $(n_1 + m_1, n_2 + m_2)$, the other entries are self-explaining:

$(3, 0)$	ϕ_{111}	$\phi_{11;1}$	$\phi_{1;11}$		
dq_1^3	2	0	0		
$dq_1^2 dx_1$	0	2	-2		
$dq_1 dx_1^2$	0	0	0		
dx_1^3	1	3	-3		

$(2, 1)$	ϕ_{112}	$\phi_{11;2}$	$\phi_{12;1}$	$\phi_{1;12}$	$\phi_{2;11}$
$dq_1^2 dq_2$	2	0	0	0	0
$dq_1^2 dx_2$	1	0	0	0	0
$dq_1 dq_2 dx_1$	-1	0	0	0	0
$dq_1 dx_1 dx_2$	1	1	1	2	0
$dq_2 dx_1^2$	1	1	1	0	2
$dx_1^2 dx_2$	-1	-1	-2	-2	-1

Tables for $(n_1 + m_1, n_2 + m_2) = (1, 2), (0, 3)$ can be obtained from the upper two upon switching $1 \leftrightarrow 2$. Response relations are vectors in the image of the above matrices. The number of independent response relations at arbitrary order has been recently derived [102].

4 Episode 2: Single Edge Current

Focus is on one observational current supported on one edge. We first provide the central result of our construction, then characterize hidden time-reversal, describe properties of the effective affinity, analyze stalling steady states, and prove certain generalized FRs. Part of this material has been anticipated in Refs. [50, 58].

4.1 Main Result

We concentrate on the time-extensive current $\Phi^t := \Phi_{12}^t$ along edge 1–2, assuming it is not a bridge (an edge whose removal disconnects the network). We construct the tilted operator

$$M(q) := \begin{pmatrix} -w_1 & e^{-q} w_{12} & w_{13} & \dots \\ e^q w_{21} & -w_2 & w_{23} & \\ w_{31} & w_{32} & -w_3 & \\ \vdots & & & \ddots \end{pmatrix} \quad (105)$$

where q is the tilting parameter. The Perron–Frobenius eigenvalue $\lambda(q)$ of $M(q)$ is the SCGF of the current. Notice that $M(0) = W$ is the forward generator. The forthcoming propositions introduce and characterize the crucial objects of our study, the hidden time-reversal generator \tilde{W} , and the effective affinity Q .

Proposition 1 *There exists a unique nonvanishing value of the tilting parameter $q = Q$, which we call the effective affinity, and a unique positive-definite diagonal matrix P_{st} , with unit trace $\text{tr } P_{\text{st}} = 1$, such that the following matrix is a MJPG*

$$\tilde{W} := P_{\text{st}} M(Q)^T P_{\text{st}}^{-1}. \quad (106)$$

Proposition 2 *The vector \vec{p}^{st} of positive diagonal entries of P_{st} is the unique normalized steady-state ensemble of the system where edge 1–2 is removed. The effective affinity is given by*

$$Q = \log \frac{w_{12} p_2^{\text{st}}}{w_{21} p_1^{\text{st}}}. \quad (107)$$

Before attacking the proof of these propositions, we point out that this result yields the IFR as a corollary.

Proposition 3 *The integral fluctuation relation holds*

$$\lambda(Q) = 0. \quad (108)$$

Proof The tilted operator $M(q)$ has a unique dominant Perron–Frobenius eigenvalue $\lambda(q)$ that is the SCGF of the current under scrutiny. Since by Eq. (106) at $q = Q$ operators $M(Q)$ and \tilde{W} have the same spectrum and the Perron eigenvalue of \tilde{W} is zero, we conclude. \square

Proof of Propositions 1 and 2 We have

$$\tilde{W} = \begin{pmatrix} -w_1 & e^Q w_{21} \frac{p_1^{\text{st}}}{p_2^{\text{st}}} & w_{31} \frac{p_1^{\text{st}}}{p_3^{\text{st}}} & \dots \\ e^{-Q} w_{12} \frac{p_2^{\text{st}}}{p_1^{\text{st}}} & -w_2 & w_{32} \frac{p_2^{\text{st}}}{p_3^{\text{st}}} & \\ w_{13} \frac{p_3^{\text{st}}}{p_1^{\text{st}}} & w_{23} \frac{p_3^{\text{st}}}{p_2^{\text{st}}} & -w_3 & \\ \vdots & & & \ddots \end{pmatrix}. \quad (109)$$

Since off-diagonal entries are positive, and diagonal entries are negative, for \tilde{W} to be a MJPG we only need to impose that $\mathbf{1} := (1, 1, \dots, 1)^T$ is a left null vector. We obtain the system of equations

$$w_1 = e^{-Q} w_{12} \frac{p_2^{\text{st}}}{p_1^{\text{st}}} + \sum_{j>2} w_{1j} \frac{p_j^{\text{st}}}{p_1^{\text{st}}} \quad (110a)$$

$$w_2 = e^Q w_{21} \frac{p_1^{\text{st}}}{p_2^{\text{st}}} + \sum_{j>2} w_{2j} \frac{p_j^{\text{st}}}{p_2^{\text{st}}} \quad (110b)$$

$$w_i = \sum_{j \neq i} w_{ij} \frac{p_j^{\text{st}}}{p_i^{\text{st}}}, \quad i > 2. \quad (110c)$$

Noticing that $w_1 = w_{21} + \sum_{i>2} w_{i1}$ (and similarly for w_2), we can rearrange the latter system of equations into

$$\sum_{j>2} (w_{j1} p_1^{\text{st}} - w_{1j} p_j^{\text{st}}) = w_{21} p_1^{\text{st}} - e^{-Q} w_{12} p_2^{\text{st}} \quad (111a)$$

$$\sum_{j>2} (w_{j2} p_2^{\text{st}} - w_{2j} p_j^{\text{st}}) = w_{12} p_2^{\text{st}} - e^Q w_{21} p_1^{\text{st}} \quad (111b)$$

$$\sum_j (w_{ji} p_i^{\text{st}} - w_{ij} p_j^{\text{st}}) = 0, \quad i > 2. \quad (111c)$$

The crucial passage now is to recognize that the left-hand side of this system of equations can be written as $W_{\text{hid}} \vec{p}^{\text{st}}$, where W_{hid} is the generator for the MJPG of the network where edge 1–2 is *deleted*

$$W_{\text{hid}} := \begin{pmatrix} -w_1 + w_{21} & 0 & w_{13} & \dots \\ 0 & -w_2 + w_{12} & w_{23} & \\ w_{31} & w_{32} & -w_3 & \\ \vdots & & & \ddots \end{pmatrix}. \quad (112)$$

Since $\vec{1}$ must also be a left-null vector of W_{hid} , we necessarily have

$$w_{21}p_1^{\text{st}} - e^{-Q}w_{12}p_2^{\text{st}} + w_{12}p_2^{\text{st}} - e^Qw_{21}p_1^{\text{st}} = 0. \quad (113)$$

The only solutions are the trivial one $Q = 0$, in which case \vec{p}^{st} solves the steady-state equation $W\vec{p}^{\text{st}} = 0$, and therefore \tilde{W} coincides with the time-reversed generator, and the nontrivial solution anticipated in Eq. (107). In this case Eqs. (111) are steady-state equations in the network where the edge 1–2 is removed, which by assumption is connected, hence there exists a unique positive normalized solution \vec{p}^{st} , which we call the stalling steady state. \square

4.2 Dynamics: Marginal and Hidden

We hereby investigate the relationship between generators W and \tilde{W} . Given the explicit expression of the effective affinity Eq. (107), the hidden TR generator reads

$$\tilde{W} = \begin{pmatrix} -w_1 & w_{12} & w_{31} \frac{p_1^{\text{st}}}{p_3^{\text{st}}} & \dots \\ w_{21} & -w_2 & w_{32} \frac{p_2^{\text{st}}}{p_3^{\text{st}}} & \\ w_{13} \frac{p_3^{\text{st}}}{p_1^{\text{st}}} & w_{23} \frac{p_3^{\text{st}}}{p_2^{\text{st}}} & -w_3 & \\ \vdots & & & \ddots \end{pmatrix}. \quad (114)$$

Notice that the upper 2×2 block is identical to that of the forward generator, while the rest of the matrix resembles a time-reversed generator, see Eq. (63). Hence we are tempted to speculate that the hidden TR generator preserves the dynamics along edge 1–2, while it reverses the dynamics in the hidden sector of the network. Let us make this intuition more precise.

We already introduced in Eq. (112) what we call the hidden MJPG W_{hid} of the dynamics occurring on network where edge 1–2 is removed. Let us further consider the *marginal* MJPG that generates the two-configuration dynamics along the sole edge 1–2:

$$W_{\text{mar}} := W - W_{\text{hid}} = \begin{pmatrix} -w_{21} & w_{12} & 0 & \dots \\ w_{21} & -w_{12} & 0 & \\ 0 & 0 & 0 & \\ \vdots & & & \ddots \end{pmatrix}. \quad (115)$$

Proposition 4 *The forward and the hidden time-reversal generators can be expressed as*

$$W = W_{\text{mar}} + W_{\text{hid}} \quad (116a)$$

$$\tilde{W} = W_{\text{mar}} + \overline{W_{\text{hid}}}. \quad (116b)$$

Proof This obviously follows from the fact that \vec{p}^{st} is the steady state of the hidden dynamics, and by the definition of time-reversed generator Eq. (63). \square

This makes our intuition more precise: the hidden TR generator inverts the dynamics in the hidden sector, but not in the marginal sector.

4.3 Thermodynamics: The Effective Affinity

The effective affinity can be interpreted operationally as follows. Prepare the system at the stalling steady state where 1–2 is removed and suddenly quench the system by connecting edge 1–2. Then Q is the thermodynamic force at the instant of connecting. In this section we further characterize the effective affinity and put it in relation to other thermodynamic forces.

Let us first report the expressions for the effective affinity Eq. (107), for the steady-state thermodynamic force $\mathcal{F} := \mathcal{F}_{12}$ along edge 1–2 Eq. (65), and for the cycle affinity Eq. (72):

$$\mathcal{F} = \log \frac{w_{12} p_2}{w_{21} p_1}, \quad Q = \log \frac{w_{12} p_2^{\text{st}}}{w_{21} p_1^{\text{st}}}, \quad \mathcal{A} = \log \frac{w_{12} p_2^{\text{eq}}}{w_{21} p_1^{\text{eq}}}. \quad (117)$$

This sequence displays the insurgence of a sort of hierarchy of marginal theories: from the most complete one, on the right-end of the spectrum, through the marginal, to the “agnostic” one on the far left. More on the hierarchy in Sect. 5.4.

Notice that in general the steady state \vec{p} of the forward dynamics, and that $\vec{\tilde{p}}$ of the hidden TR dynamics, differ among themselves. Nevertheless, the following result holds.

Proposition 5 *The steady-state edge affinity $\tilde{\mathcal{F}}$ in the hidden time-reversal generator is identical to that of the forward dynamics,*

$$\tilde{\mathcal{F}} = \mathcal{F}. \quad (118)$$

Proof This amounts to prove that

$$\frac{p_1}{p_2} = \frac{\tilde{p}_1}{\tilde{p}_2}. \quad (119)$$

As we mentioned in Sect. 3.2, the steady-state ensemble can be calculated in terms of minors. In view of the explicit form of the generators Eqs. (105, 114), we consider the minors obtained by removing the first row and column and the second row and column, and notice that

$$\tilde{W}(1|1) = P_{\text{st}}(1|1) W(1|1) P_{\text{st}}(1|1)^{-1} \quad (120a)$$

$$\tilde{W}(2|2) = P_{\text{st}}(2|2) W(2|2) P_{\text{st}}(2|2)^{-1}. \quad (120b)$$

where we remind that matrix $A(i_1, \dots, i_n | j_1, \dots, j_m)$ is obtained by removing rows i_1, \dots, i_n and columns j_1, \dots, j_m . Since these are matrix similarities their determinants coincide. \square

While the ratio of the populations at sites 1 and 2 stays the same upon hidden TR, the total density might change due to normalization, which is a nonlocal factor; for this reason the steady-state currents of the forward and hidden TR current along edge 1–2 are not the same. However, the following result is an obvious corollary of the above proposition, corroborating the intuition expressed above that the hidden TR dynamics “tends to preserve” the dynamics in the marginal sector.

Proposition 6 *The steady-state mean currents ϕ_{12} and $\tilde{\phi}_{12}$ respectively of the forward and of the hidden TR dynamics have the same sign.*

We will now compare the “real” affinities of the forward and the marginal TR dynamics.

Proposition 7 For any simple cycle \mathcal{C} , let $\mathcal{A}(\mathcal{C})$ and $\tilde{\mathcal{A}}(\mathcal{C})$ denote the cycle affinities with respect to the forward and hidden TR dynamics, respectively. We have

$$\tilde{\mathcal{A}}(\mathcal{C}) = -\mathcal{A}(\mathcal{C}), \quad \text{if } \mathcal{C} \not\ni 1-2 \quad (121a)$$

$$\tilde{\mathcal{A}}(\mathcal{C}) = -\mathcal{A}(\mathcal{C}) + 2Q, \quad \text{if } \mathcal{C} \ni 1-2. \quad (121b)$$

Proof For all cycles that do not contain edge 1–2, one has

$$\tilde{\mathcal{A}}(\mathcal{C}) = \log \prod_{ij \in \mathcal{C}} \frac{\tilde{w}_{ij}}{\tilde{w}_{ji}} = \log \prod_{ij \in \mathcal{C}} \frac{w_{ji} p_i^{\text{st}} / p_j^{\text{st}}}{w_{ij} p_j^{\text{st}} / p_i^{\text{st}}}. \quad (122)$$

All terms p_i^{st} cancel out one with each other along cycles, so the last term can be identified with $-\mathcal{A}(\mathcal{C})$. As regards cycles containing edge 1–2, one has

$$\tilde{\mathcal{A}}(\mathcal{C}) = \log \prod_{\substack{ij \in \mathcal{C} \\ ij \neq 12}} \frac{w_{ji} p_i^{\text{st}} / p_j^{\text{st}}}{w_{ij} p_j^{\text{st}} / p_i^{\text{st}}} + \log \frac{w_{12}}{w_{21}} \quad (123)$$

$$= \log \prod_{ij \in \mathcal{C}} \frac{w_{ji} (p_i^{\text{st}})^2}{w_{ij} (p_j^{\text{st}})^2} - \log \frac{w_{21} (p_1^{\text{st}})^2}{w_{12} (p_2^{\text{st}})^2} + \log \frac{w_{12}}{w_{21}} \quad (124)$$

and we conclude. \square

We will now employ the expression of the steady state in terms of spanning trees to compare edge and effective affinities. In Sect. 4.4 we provide a graphical representation by an example. Let us remind the definition of fluxes in the full network introduced in Sect. 3.2.1 and introduce fluxes in the network where 1–2 is either deleted or contracted, that is, where it is shrunk to a unique site (here we set for simplicity the normalization $\tau(\mathcal{G}) \stackrel{!}{=} 1$).

$$\psi_{12} = w_{12} \tau_2(\mathcal{G}) = w_{12} \det W(2|1) \quad (125a)$$

$$\psi_{\setminus 12} := w_{12} \tau_2(\mathcal{G} \setminus 1-2) = w_{12} \det W_{\text{hid}}(2|1) \quad (125b)$$

$$\psi_{/12} := w_{12} w_{21} \tau_1(\mathcal{G}/1-2) = w_{12} w_{21} \det W(1, 2|1, 2). \quad (125c)$$

Similar definitions are obtained by interchanging 1 and 2. Notice that $\psi_{/12} = \psi_{/21}$. Then we can write:

$$Q = \log \frac{\psi_{\setminus 12}}{\psi_{\setminus 21}}. \quad (126)$$

We can also consider spanning trees in the graph $\mathcal{G}/1-2$ where edge 1–2 is contracted (sites 1 and 2 are identified). The following deletion-contraction formula holds

$$\tau_1(\mathcal{G}) = \tau_1(\mathcal{G} \setminus 1-2) + w_{12} \tau_2(\mathcal{G}/1-2). \quad (127)$$

which in terms of determinants reads

$$\det W(2|1) = \det W_{\text{hid}}(2|1) + w_{12} \det W(1, 2|1, 2). \quad (128)$$

Basically, this formula states that all spanning trees (on the left) either do not contain edge 1–2 (first term on the right) or else they do (second term on the right). If they do, they need to be oriented towards the contracted vertices. Notice that because in the contracted graph 1 is identified with 2, clearly $\mathcal{T}_2(\mathcal{G}/1-2) = \mathcal{T}_1(\mathcal{G}/1-2)$. Then we find that

$$0 \leq \psi_{/12} = \psi_{12} - \psi_{\setminus 12} = \psi_{21} - \psi_{\setminus 21}, \quad (129)$$

that is, the difference between the fluxes is independent of the direction of the edge. Before giving a graphical example, we are now in the position to prove the following result.

Proposition 8 *The effective affinity has larger modulus than the edge affinity,*

$$|Q| \geq |\mathcal{F}|. \quad (130)$$

Proof Suppose $\psi_{12} > \psi_{12}$, whence both $Q > 0$ and $\mathcal{F} > 0$. We have

$$\mathcal{F} = \log \frac{\psi_{12}}{\psi_{21}} = \log \frac{\psi_{12} + \psi_{12}}{\psi_{21} + \psi_{12}} \quad (131)$$

$$= Q + \log \frac{1 + \frac{\psi_{12}}{\psi_{12}}}{1 + \frac{\psi_{12}}{\psi_{21}}}. \quad (132)$$

Because $\psi_{12} \geq \psi_{21}$, then the latter term above is negative and thus we conclude. \square

4.4 Graphical Representation

Let us consider the graph:



In the following, each directed edge means “multiply by the corresponding rate”. The oriented spanning-tree polynomials with root in $i = 1, 2$ are given by

$$\begin{aligned} \tau_1(\mathcal{G}) = & \begin{array}{c} \leftarrow \uparrow \uparrow + \leftarrow \uparrow \downarrow + \leftarrow \uparrow \downarrow + \leftarrow \uparrow \downarrow + \leftarrow \uparrow \downarrow \\ + \leftarrow \uparrow \downarrow + \leftarrow \uparrow \downarrow + \leftarrow \uparrow \downarrow \end{array}, \end{aligned} \quad (134a)$$

$$\begin{aligned} \tau_2(\mathcal{G}) = & \begin{array}{c} \leftarrow \uparrow \uparrow + \leftarrow \uparrow \downarrow + \leftarrow \uparrow \downarrow + \leftarrow \uparrow \downarrow + \leftarrow \uparrow \downarrow \\ + \leftarrow \uparrow \downarrow + \leftarrow \uparrow \downarrow + \leftarrow \uparrow \downarrow \end{array}. \end{aligned} \quad (134b)$$

Let us now consider the graph with deleted edge $\mathcal{G} \setminus 1-2$:



We have for the unnormalized steady state

$$\tau_1(\mathcal{G} \setminus 1-2) = \begin{array}{c} \leftarrow \uparrow \uparrow + \leftarrow \uparrow \downarrow + \leftarrow \uparrow \downarrow \\ + \leftarrow \uparrow \downarrow + \leftarrow \uparrow \downarrow + \leftarrow \uparrow \downarrow \end{array} \quad (136a)$$

$$\tau_2(\mathcal{G} \setminus 1-2) = \begin{array}{c} \leftarrow \uparrow \uparrow + \leftarrow \uparrow \downarrow + \leftarrow \uparrow \downarrow \\ + \leftarrow \uparrow \downarrow + \leftarrow \uparrow \downarrow + \leftarrow \uparrow \downarrow \end{array} \quad (136b)$$

Multiplying respectively by w_{21} and w_{12} one obtains Hill's fluxes, yielding for the edge affinity

$$\mathcal{F} = \log \frac{\begin{array}{c} \uparrow \rightarrow + \rightarrow \rightarrow + \rightarrow \leftarrow + \nearrow + \nearrow \uparrow + \rightarrow \leftarrow + \nearrow \leftarrow \\ \uparrow \uparrow + \rightarrow \rightarrow + \rightarrow \leftarrow + \nearrow + \nearrow \uparrow + \rightarrow \leftarrow + \nearrow \leftarrow \end{array}}{\begin{array}{c} \uparrow \rightarrow + \rightarrow \rightarrow + \rightarrow \leftarrow + \nearrow + \nearrow \uparrow + \rightarrow \leftarrow + \nearrow \leftarrow \\ \uparrow \uparrow + \rightarrow \rightarrow + \rightarrow \leftarrow + \nearrow + \nearrow \uparrow + \rightarrow \leftarrow + \nearrow \leftarrow \end{array}} \quad (137)$$

and for the effective affinity

$$Q = \log \frac{\begin{array}{c} \rightarrow \rightarrow + \nearrow \uparrow + \rightarrow \leftarrow \\ \rightarrow \rightarrow + \nearrow \uparrow + \rightarrow \leftarrow \end{array}}{\begin{array}{c} \rightarrow \rightarrow + \nearrow \uparrow + \rightarrow \leftarrow \\ \rightarrow \rightarrow + \nearrow \uparrow + \rightarrow \leftarrow \end{array}} \quad (138)$$

Notice that a "real" cycle affinity with respect for example to spanning tree \square is

$$\mathcal{A} = \log \frac{\begin{array}{c} \rightarrow \rightarrow \\ \leftarrow \leftarrow \end{array}}{\begin{array}{c} \rightarrow \rightarrow \\ \leftarrow \leftarrow \end{array}} \quad (139)$$

which gives a representation of decreasing cycle contributions. Finally, consider the contracted graph $\mathcal{G}/1-2$

$$\begin{array}{ccc} 1 & \cdots & 2 \\ | & & | \\ 4 & \text{---} & 3 \end{array} \quad (140)$$

where the dotted line signifies identification of sites. We have

$$\tau_1(\mathcal{G}/1-2) = \begin{array}{c} \cdots \uparrow \\ \uparrow \end{array} + \begin{array}{c} \cdots \uparrow \\ \rightarrow \end{array} + \begin{array}{c} \cdots \uparrow \\ \leftarrow \end{array} + \begin{array}{c} \cdots \nearrow \\ \nearrow \end{array} + \begin{array}{c} \cdots \nearrow \\ \nearrow \end{array} \quad (141)$$

and we obtain

$$\det W(1, 2|1, 2) = \begin{array}{c} \uparrow \rightarrow \\ \rightarrow \end{array} + \begin{array}{c} \uparrow \rightarrow \\ \rightarrow \end{array} + \begin{array}{c} \uparrow \rightarrow \\ \leftarrow \end{array} + \begin{array}{c} \uparrow \rightarrow \\ \leftarrow \end{array} + \begin{array}{c} \uparrow \rightarrow \\ \nearrow \end{array}. \quad (142)$$

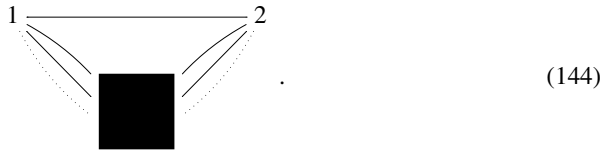
One can then verify the deletion-contraction relation and the validity of Proposition 8.

4.5 Marginal Entropy Production Rate

We call the quantity

$$\sigma_1 := \phi Q \geq 0 \quad (143)$$

the mean *marginal EPR*. The superscript marks the observable current $\phi = \phi_1$. We have shown in Ref. [50] that it can be interpreted as the mean EPR estimated by an observer who formulates a minimal Markovian model accounting for the observable steady-state mean current. Let us review the argument. We consider an observer who controls w_{12} , w_{21} , and the steady-state ratio p_1/p_2 . To him, the rest of the system is a black box:



The observer needs to formulate a minimal model that is compatible with the observation of a current along 1–2. The simplest possible setup is



In this minimal model the black box is responsible of returning an event at 1 or 2 at some effective rates $\tilde{w}_{12}, \tilde{w}_{21}$, which we require to be independent of w_{12}, w_{21} . These effective rates differ from the one considered e.g. by Uhl et al. in Ref. [22]. Notice that the effective model has only one effective cycle



Proposition 9 *The affinity of the effective cycle is equal to the effective affinity,*

$$\mathcal{A}(\mathcal{C}^{\text{eff}}) = Q. \quad (147)$$

Proof The topology of the minimal model is such that, lumping together the transitions w and \tilde{w} , global detailed balance must hold:

$$\frac{w_{21} + \tilde{w}_{21}}{w_{12} + \tilde{w}_{12}} = \frac{p_2}{p_1}. \quad (148)$$

If we require this condition to hold independently of w_{12}, w_{21} , using Eq. (128) we obtain for the effective rates

$$\tilde{w}_{12} = \frac{\det W_{\text{hid}}(2, 1)}{\det W(1, 2|1, 2)} \quad (149a)$$

$$\tilde{w}_{21} = \frac{\det W_{\text{hid}}(1, 2)}{\det W(1, 2|1, 2)}. \quad (149b)$$

The cycle affinity of the effective cycle \mathcal{C}^{eff} is given by

$$\mathcal{A}(\mathcal{C}^{\text{eff}}) = \log \frac{w_{12}\tilde{w}_{21}}{w_{21}\tilde{w}_{12}} \quad (150)$$

and by replacing the values of the effective rates we conclude. \square

It is interesting to notice that, in view of Eq. (119), the ideal observer is completely blind to hidden TR.

Let us now show that such observer always underestimates the mean EPR and, even more strictly, the local mean EPR associated to Hill's cycle decomposition Eq. (78). A different derivation was proposed in Ref. [58].

Proposition 10 *The mean marginal EPR is positive, and it is always smaller than the Hill mean marginal EPR and of the complete mean EPR,*

$$0 \leq \sigma_1 \leq \sigma_{12}^{\text{Hill}} \leq \sigma. \quad (151)$$

Proof Positivity follows from the explicit expression

$$\sigma_1 = (\psi_{\setminus 12} - \psi_{\setminus 21}) \log \frac{\psi_{\setminus 12}}{\psi_{\setminus 21}}, \quad (152)$$

whereby the linear and the logarithmic terms are either both positive or both negative.

To show that the marginal EPR underestimates the full EPR, we employ the cycle decomposition of the entropy production rate Eq. (77). On the one hand we have

$$\sigma = \left(\sum_{\mathcal{C} \ni 12} + \sum_{\mathcal{C} \not\ni 12} \right) (\psi_{+\mathcal{C}} - \psi_{-\mathcal{C}}) \log \frac{\psi_{+\mathcal{C}}}{\psi_{-\mathcal{C}}}, \quad (153)$$

where we remind that $-\mathcal{C}$ is the cycle with reversed orientation. Notice that each of the summands is positive. Hence we can focus on just the cycles that pass through 1–2:

$$\begin{aligned} \sigma &\geq \sum_{\mathcal{C} \ni 12} (\psi_{+\mathcal{C}} - \psi_{-\mathcal{C}}) \log \frac{\psi_{+\mathcal{C}}}{\psi_{-\mathcal{C}}} \\ &\geq \left[\sum_{\mathcal{C} \ni 12} (\psi_{+\mathcal{C}} - \psi_{-\mathcal{C}}) \right] \log \frac{\sum_{\mathcal{C} \ni 12} \psi_{+\mathcal{C}}}{\sum_{\mathcal{C} \ni 12} \psi_{-\mathcal{C}}}. \end{aligned} \quad (154)$$

The first expression is Hill's mean EPR along edge 1–2; we then used the log-sum inequality for the logarithm, an usual tool in matters of coarse-graining in information theory [35, 115]. In the latter expression we recognize the mean marginal EPR. \square

4.6 Stalling

The following two results connect the thermodynamic and the kinematic interpretation of stalling.

Proposition 11 *The steady-state mean current ϕ vanishes if and only if the effective affinity Q vanishes.*

Proof If we prepare the system in state \vec{p}^{st} and then transition 1–2 is turned on abruptly by a quench, initially we have

$$W \vec{p}^{\text{st}} = \tilde{W} \vec{p}^{\text{st}} = \begin{pmatrix} w_{12} p_2^{\text{st}} - w_{21} p_1^{\text{st}} \\ w_{21} p_1^{\text{st}} - w_{12} p_2^{\text{st}} \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \quad (155)$$

This follows from the fact that \vec{p}^{st} is the steady state of both W_{hid} and $\overline{W}_{\text{hid}}$. Then the current vanishes if and only if the effective affinity vanishes. \square

Proposition 12 *At stalling, the hidden TR generator coincides with the time-reversed generator*

$$\tilde{W} = \overline{W}. \quad (156)$$

Proof We notice that at stalling, because \vec{p}^{st} is the steady state and the effective affinity vanishes, then $W_{\text{mar}} = \overline{W}_{\text{mar}}$. The conclusion follows from Eqs. (120). \square

The latter can be interpreted as a condition of marginal detailed balance.

4.7 Response at Stalling

In this section we look at what happens in the vicinity of a stalling steady state. Let us reintroduce the index $\phi = \phi_1$ denoting that this is the first observable current of a theory with only $|\mu| = 1$ observable currents (the reader should not confuse this index with the site $i = 1$). We introduce a parametrization of the rates along the observable edge $w_{12}, w_{21} \rightarrow w_{12}(x), w_{21}(x)$ such that

$$\frac{w_{12}(x)}{w_{21}(x)} = \exp x \quad (157)$$

and no other rate depends on x . We will discuss in Sect. 5.8 how general such a parametrization is. Let us make explicit the dependence of the SCGF on x , $\lambda(q) \rightarrow \lambda(q; x)$. The effective affinity now takes the form

$$Q(x) = x - x^{\text{st}} \quad (158)$$

where

$$x^{\text{st}} := \log \frac{p_2^{\text{st}}}{p_1^{\text{st}}}. \quad (159)$$

The above expression for the effective affinity allows for simple phenomenological interpretation: in fact, as far as the parameter is thermodynamic in the sense discussed in the Parode, the effective affinity can be determined as the distance of the parameter from the value that makes the current stall, which can in principle be found by a simple tuning procedure.

The IFR Eq. (108) now reads:

$$\lambda(x - x^{\text{st}}; x) = 0. \quad (160)$$

We now focus to second-order expansion in the vicinity of stalling. Taking the first total derivative of the marginal IFR with respect to x we recover the fact that if the effective affinity vanishes, then the current vanishes:

$$0 = \frac{d}{dx} \lambda(x - x^{\text{st}}; x) \Big|_{x=x^{\text{st}}} = \left[\frac{\partial \lambda}{\partial q} + \frac{\partial \lambda}{\partial x} \right] (0; x^{\text{st}}) = \phi_1^{\text{st}}. \quad (161)$$

Similarly, taking the second derivative and evaluating at stalling we rederive the FDR at stalling found in Ref. [51]:

$$\frac{\partial \phi_1(x)}{\partial x} \Big|_{x=x^{\text{st}}} = \frac{1}{2} \frac{\partial^2 \lambda}{\partial q^2} (0; x^{\text{st}}) = \frac{1}{2} \phi_{11}^{\text{st}}. \quad (162)$$

We can also provide an explicit expression for the variance at stalling:

Proposition 13 *The current's variance at stalling is given by*

$$\phi_{11}^{\text{st}} = 2w_{12}(x^{\text{st}})p_2^{\text{st}} = 2w_{21}(x^{\text{st}})p_1^{\text{st}}. \quad (163)$$

Proof We notice that the determinant of the tilted operator has the functional form

$$\det M(q) = w_{12} \det W(1|2)e^{-q} + w_{21} \det W(2|1)e^q + \text{const.} \quad (164)$$

On the other hand, it is the product of the eigenvalues $\det M(q) = \prod_{i=1}^{|\mathcal{S}|} \lambda_i(q)$, including $\lambda_1 = \lambda$ the SCGF. Taking the first derivative

$$\frac{\partial}{\partial q} \det M = \frac{\partial \lambda}{\partial q} \prod_{i \neq 1} \lambda_i(q) + \lambda(q) \sum_{i \neq 1} \frac{\partial \lambda_i}{\partial q} \prod_{j \neq 1, i} \lambda_j(q). \quad (165)$$

Evaluating at $q = 0$ where $\lambda(0) = 0$, using the fact that $\phi = -\partial_q \lambda(0)$ and comparing with Eq. (164) we obtain

$$\phi_1 = \frac{w_{12} \tau_2(\mathcal{G}) - w_{21} \tau_1(\mathcal{G})}{\prod_{i \neq 1} \lambda_i(q)}. \quad (166)$$

This shows that $\prod_{i \neq 1} \lambda_i(q) = \tau(\mathcal{G})$, the product of the nonvanishing eigenvalues is the normalization factor of the steady state.

Now, taking the second derivatives and evaluating at $q = 0$ we obtain

$$w_{12} p_2 + w_{21} p_1 = \phi_{11} - 2\phi_1 \frac{\partial}{\partial q} \log \prod_{i \neq 1} \lambda_i(0) \quad (167)$$

and from this we easily conclude by evaluating at stalling. \square

We notice that a similar expression holds in the linear regime near equilibrium, and that this is the basis for the linear-regime analysis of Schnakenberg, see the latest paragraph in Ref. [69]. It might therefore be possible to attempt a similar marginal linear-regime analysis.

4.8 Fluctuation Relations for Marginal Currents

In this section and the following we prove the FR announced in Eq. (50) by a direct method involving path probabilities. In particular we consider the probability associated to the hidden TR process, which in the light of Eq. (81) reads

$$\tilde{P}(\omega^t) = \tilde{p}_{i_0}^0 e^{-w_{i_N} t_N} \prod_{n=0}^{N-1} \tilde{w}_{i_{n+1}, i_n} e^{-w_{i_n} t_n}. \quad (168)$$

where ω^t is the path described by Eq. (80), and $\vec{\tilde{p}}^0$ is the ensemble from which the initial configuration is sampled. We notice that all exit rates out of sites are the same as for the forward dynamics.

Proposition 14 *The marginal FR*

$$\frac{P(\Phi)}{\tilde{P}(-\Phi)} = \exp Q \Phi \quad (169)$$

holds at all times, provided the initial configuration of the process is sampled from the stalling state $\vec{\tilde{p}}^{\text{st}}$.

Proof Let us consider the ratio of the probability densities

$$\frac{P(\bar{\omega}^t)}{\tilde{P}(\omega^t)} = \frac{\bar{p}_{i_N}}{\tilde{p}_{i_0}^0} \prod_{n=0}^{N-1} \frac{w_{i_n, i_{n+1}}}{\tilde{w}_{i_{n+1}, i_n}} = \frac{\bar{p}_{i_N}}{\tilde{p}_{i_0}^0} \prod_{i,j} \left(\frac{w_{ji}}{\tilde{w}_{ij}} \right)^{\Psi_{ij}^t} \quad (170)$$

where we expressed the path probability in terms of the time-integrated fluxes. We now single out the rates corresponding to transition 1–2 and use the explicit expression of the rates:

$$\begin{aligned} \frac{P(\bar{\omega}^t)}{\tilde{P}(\omega^t)} &= \frac{\bar{p}_{i_N}}{\tilde{p}_{i_0}^0} \left(\frac{w_{12}}{\tilde{w}_{21}} \right)^{\Psi_{12}^t} \left(\frac{w_{21}}{\tilde{w}_{12}} \right)^{\Psi_{21}^t} \prod_{(i,j) \neq (1,2), (2,1)} \left(\frac{w_{ji}}{\tilde{w}_{ij}} \right)^{\Psi_{ij}^t} \\ &= \frac{\bar{p}_{i_N}}{\tilde{p}_{i_0}^0} \left(\frac{w_{21}}{w_{12}} \right)^{\Phi^t} \prod_{ij \neq 12} \left(\frac{p_j^{\text{st}}}{p_i^{\text{st}}} \right)^{\Phi_{ij}^t}, \end{aligned} \quad (171)$$

where in the latter passage we used $\Phi_{ij}^t = \Phi_{ij}^t - \Phi_{ji}^t$ and moved from fluxes to currents using antisymmetry. We now multiply and divide by $(p_2^{\text{st}}/p_1^{\text{st}})^{\Phi^t}$ to obtain

$$\begin{aligned} \frac{P(\bar{\omega}^t)}{\tilde{P}(\omega^t)} &= \frac{\bar{p}_{i_N}}{\tilde{p}_{i_0}^0} \left(\frac{w_{21} p_1^{\text{st}}}{w_{12} p_2^{\text{st}}} \right)^{\Phi^t} \prod_{ij} \left(\frac{p_j^{\text{st}}}{p_i^{\text{st}}} \right)^{\Phi_{ij}^t} \\ &= \frac{\bar{p}_{i_N}}{\tilde{p}_{i_0}^0} \left(\frac{w_{21} p_1^{\text{st}}}{w_{12} p_2^{\text{st}}} \right)^{\Phi^t} \exp \sum_{n=0}^{N-1} \log \frac{p_{i_n}^{\text{st}}}{p_{i_{n+1}}^{\text{st}}}. \end{aligned} \quad (172)$$

In the last passage we moved back from an expression in terms of the fluxes to one in terms of the jumping events. Now, being $\log p_i^{\text{st}}/p_j^{\text{st}}$ an exact discrete differential, terms cancel out and its sum along a path is just the difference between the initial and final values:

$$P(\bar{\omega}^t) = \frac{\bar{p}_{i_N} p_{i_0}^{\text{st}}}{\tilde{p}_{i_0}^0 p_{i_N}^{\text{st}}} e^{-Q \Phi^t} \tilde{P}(\omega^t). \quad (173)$$

Sampling the initial site for all processes from the stalling steady state, the finite-time prefactor cancels out. We can now integrate over all trajectories that give current $\Phi^t \equiv \Phi$, obtaining

$$P(-\Phi) = e^{-Q \Phi} \tilde{P}(\Phi), \quad (174)$$

which is the desired result (up to $\Phi \rightarrow -\Phi$). \square

Notice that integrating Eq. (174) with respect to $d\Phi$ immediately yields the marginal IFR

$$\left\langle e^{-Q \Phi^t} \right\rangle = 1. \quad (175)$$

Here, the probability density of the hidden TR dynamics is traced out, so that it does not play a role for the marginal IFR, a common trick that was employed e.g. in [93] to obtain IFRs for time-symmetric observables.

According to the lines of Ref. [58], we now consider the *hidden entropy production*,

$$\Sigma_1^t := \sum_{\alpha} \mathcal{A}_{\alpha} \Phi_{\alpha}^t - Q \Phi^t, \quad (176)$$

where again the subscript 1 refers to the fact that we are consider the marginal theory w.r.t. current $\Phi^t = \Phi_1^t$.

Proposition 15 *The following effective FR and the IFR for the hidden entropy production hold asymptotically at long times:*

$$\frac{P(\Sigma_1)}{\tilde{P}(\Sigma_1)} \asymp e^{-\Sigma_1}, \quad \left\langle e^{\Sigma_1^t} \right\rangle \asymp 1. \quad (177)$$

Proof Putting together Eqs. (173) and (93), we obtain

$$\frac{P(\omega^t)}{\tilde{P}(\omega^t)} = \frac{p_{i_0}^{\text{st}} p_{i_0}^0}{\tilde{p}_{i_0}^0 p_{i_N}^{\text{st}}} \frac{p_{i_N}^{\text{eq}}}{p_{i_0}^{\text{eq}}} \exp \Sigma_{\text{hid}}^t. \quad (178)$$

Notice that now we cannot make a physically relevant choice of initial ensemble to cancel the prefactor. Therefore we need to go to long times. We obtain the two relations by the usual manipulations. \square

Notice that the IFR for the hidden entropy production implies $\sigma_1 \geq 0$ and $\sigma \geq \sigma_1$ for their mean time-scaled values, which we proved in Proposition 10 by a different route.

4.9 Fluctuation Relations for Marginal Cycle Currents

As discussed in Sect. 3.2, the FR for the currents is intertwined with the notion of time-reversal of a trajectory, see Eqs. (84) and (83). Time-reversal of paths is involutive, one-to-one, it inverts all of the currents and maintains the time-symmetric properties, i.e. the waiting probabilities at sites and the activities $\Psi_{ij}^t + \Psi_{ji}^t$.

The above relation Eq. (169) cannot be considered as a proper FR for current Φ^t , the crucial difference with the FR Eq. (4) being that it does not compare the same p.d.f. evaluated along different paths. In this section we propose a similar notion of hidden TR of a trajectory with properties that are analogous to those of time-reversal, and provide a marginal FR for a set of cycle currents associated with edge 1–2; however, this FR cannot be further contracted to current Φ^t alone. The construction is based on a cycle analysis at the level of paths proposed by Jia et al. [113].

To accomplish this, we consider all *simple* oriented cycles \mathcal{C} , with $-\mathcal{C}$ denoting the cycle with the inverse orientation. Suppose that a Markovian Dedalus lost in the network carries Ariadne's thread with himself. Chased by the Minotaurus, as the trajectory unravels Dedalus might cross the previously laid thread, forming loops; whenever such a crossing occurs, to avoid wasting precious thread he/she accounts for the loop and its directionality, cuts the loop away and sews the leftover strands together. This procedure produces counters of simple cycles. Let $\Psi^t(\mathcal{C})$ be the winding number in one direction of the cycle. Then the cycle current is defined as

$$\Phi^t(\mathcal{C}) := \Psi^t(\mathcal{C}) - \Psi^t(-\mathcal{C}). \quad (179)$$

Importantly, the current along edge 1–2 is the sum of all cycle currents whose cycle includes 1–2, in the appropriate direction,

$$\Phi^t = \sum_{\mathcal{C} \ni 1-2} \Phi^t(\mathcal{C}). \quad (180)$$

Thus, cycle currents constitute a refinement of the information contained in an observable current. We can now separate the contributions to a trajectory that come from cycles that contain 1–2, and those that come from simple cycles not containing 1–2. For example, considering the graph in Eq. 133), for

$$\omega^t = \{3 \xrightarrow{t_1} 2 \xrightarrow{t_2} 1 \xrightarrow{t_3} 4 \xrightarrow{t_4} 2 \xrightarrow{t_5} 5 \xrightarrow{t_6} 4 \xrightarrow{t_7} 3 \xrightarrow{t_8}\} \quad (181)$$

we obtain the two cycles $2 \rightarrow 1 \rightarrow 4 \rightarrow 2$ and $3 \rightarrow 5 \rightarrow 4 \rightarrow 3$ (notice that there also exists cycle $4 \rightarrow 2 \rightarrow 5 \rightarrow 4$, which is not independent of the other two—thus showing

that the representation in terms of cycle currents is not unique: it depends on *when* we start counting). We invert the first, obtaining the *hidden time-reversed trajectory*

$$\tilde{\omega}^t = \{3 \xrightarrow{t_8} 4 \xrightarrow{t_7} 5 \xrightarrow{t_6} 2 \xrightarrow{t_2} 1 \xrightarrow{t_3} 4 \xrightarrow{t_4} 2 \xrightarrow{t_5} 3 \xrightarrow{t_1}\} \quad (182)$$

and similarly we rearrange the waiting times. This procedure of hidden time-reversal of a trajectory has the following properties, analogous to that of full time-reversal:

- (i) It is involutive, hence in particular it is one-to-one;
- (ii) It inverts all of the cycle currents but those that pass through 1–2, hence it also preserves the observable current

$$\tilde{\Phi}^t := \Phi[\tilde{\omega}^t] = \Phi^t; \quad (183)$$

- (iii) It preserves all symmetric quantities, namely waiting times and activities.

Proposition 16 *The following FR for the cycle currents passing through edge 1–2 holds:*

$$\frac{P(\{\Phi^t(\mathcal{C})\}_{\mathcal{C} \ni 12})}{P(\{-\Phi^t(\mathcal{C})\}_{\mathcal{C} \ni 12})} \asymp \exp \sum_{\mathcal{C} \ni 12} \Phi^t(\mathcal{C}) \mathcal{A}(\mathcal{C}). \quad (184)$$

Proof Assuming that the trajectories are cyclic, $i_N = i_0$, with a few standard manipulations one arrives at

$$\frac{P(\omega^t)}{P(\tilde{\omega}^t)} = \exp \sum_{\mathcal{C} \ni 12} \Phi^t(\mathcal{C}) \mathcal{A}(\mathcal{C}). \quad (185)$$

We can then marginalize for the cycle currents. □

This comes as close as possible to obtaining a marginal FR that is localized on edge 1–2. Notice though that in general

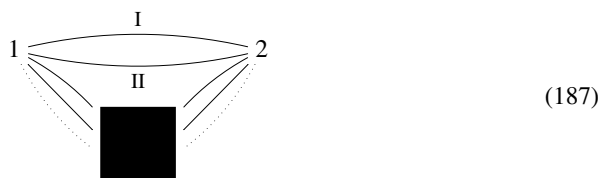
$$P(\tilde{\omega}^t) \neq \tilde{P}(\omega^t). \quad (186)$$

Understanding whether there exists relationship between the hidden TR trajectory and the hidden TR dynamics is a major open question.

4.10 The Case of Multiple Edges

For sake of notational ease we assumed so far that at most one transition per pair of sites was possible, i.e. that the network does not have multiple edges. In nonequilibrium thermodynamics it is necessary to consider the case where several different transitions are discernible, because they are due to the interaction with different reservoirs. In this section we show that the theory holds with few obvious modifications.

Let us consider the following network, which includes two transition mechanisms for 1–2, labeled I and II:



The tilted operator corresponding to the measurement of the current along edge $1 \xleftarrow{I} 2$ is given by

$$M^I(q) := \begin{pmatrix} -w_1 & e^{-q} w_{12}^I + w_{12}^{\Pi} & \dots \\ e^q w_{21}^I + w_{21}^{\Pi} & -w_2 & \\ \vdots & & \ddots \end{pmatrix}. \quad (188)$$

Apart from this little difference, a direct inspection of the proofs of Propositions 1 and 2 shows that all proceeds like above, with the only modification that the effective affinity now reads

$$Q = \log \frac{w_{12}^I p_2^{\text{st}}}{w_{21}^I p_1^{\text{st}}} \quad (189)$$

where \vec{p}^{st} is the steady-state in the network where edge $1 \xleftarrow{I} 2$ is removed. This argument can be easily scaled up to the case of several observational currents that will be the object of study of the rest of this paper. Since the procedure is fairly straightforward, we will not produce it explicitly.

5 Episode 3: Several Edge Currents

In this episode we take into consideration several currents, each supported on one edge, proving the analog of Propositions 1 and 14 and discussing some of their consequences. The most important new concept is that there is a *hierarchy* of marginal theories and of powerful inter-hierarchical fluctuation relations.

5.1 Main Result

We now consider the statistics of $|\mu|$ currents, each supported on a different edge $i_\mu j_\mu$. For the moment we assume that the removal of all such edges does not disconnect the network. We call j_μ the tail (or source) of edge $i_\mu j_\mu$ and i_μ its tip (or target). The set of currents is marginal in the sense explained in Sect. 3.4: it does not include a full set of chords of the network. Let $M(\{q_\mu\})$ be the tilted generator with entries

$$M(\{q_\mu\})_{i,j} = \begin{cases} -w_i, & \text{if } i = j \\ w_{i_\mu j_\mu} e^{-q_\mu}, & \text{if } \exists \mu \text{ s.t. } i = i_\mu, j = j_\mu \\ w_{j_\mu i_\mu} e^{q_\mu}, & \text{if } \exists \mu \text{ s.t. } i = j_\mu, j = i_\mu \\ w_{ij}, & \text{otherwise} \end{cases}. \quad (190)$$

From now on we omit to explicitly state “ $\exists \mu$ ”.

Let us consider the generator W_{hid} of the Markovian dynamics defined by setting all rates $w_{i_\mu j_\mu} = w_{j_\mu i_\mu} = 0$, *ossia* by removing the set of edges corresponding to the observable transitions. It has entries

$$[W_{\text{hid}}]_{i,j} = \begin{cases} -w_i + \sum_{k,\mu} w_{ki} (\delta_{k,j_\mu} \delta_{i,i_\mu} + \delta_{i,j_\mu} \delta_{k,i_\mu}), & \text{if } i = j, \\ 0, & \text{if } i_\mu j_\mu = ij, ji, \dots \\ w_{ij}, & \text{elsewhere} \end{cases}. \quad (191)$$

Such dynamics admits a unique normalized steady state \vec{p}^{st} . We define the effective affinities

$$Q_\mu := \log \frac{w_{i_\mu j_\mu} p_{j_\mu}^{\text{st}}}{w_{j_\mu i_\mu} p_{i_\mu}^{\text{st}}}. \quad (192)$$

Finally, let $P_{\text{st}} := \text{diag}\{p_i^{\text{st}}\}_i$.

Proposition 17 *The operator*

$$\tilde{W} := P_{\text{st}} M(\{Q_\mu\})^T P_{\text{st}}^{-1} \quad (193)$$

is a MJPG.

Proof We have

$$\tilde{W}(\{Q_\mu\})_{j,i} = \begin{cases} -w_i, & \text{if } i = j \\ w_{i_\mu j_\mu} \frac{p_{j_\mu}^{\text{st}}}{p_{i_\mu}^{\text{st}}} e^{-Q_\mu}, & \text{if } i = i_\mu, j = j_\mu \\ w_{j_\mu i_\mu} \frac{p_{i_\mu}^{\text{st}}}{p_{j_\mu}^{\text{st}}} e^{+Q_\mu}, & \text{if } i = j_\mu, j = i_\mu \\ w_{ij} \frac{p_j^{\text{st}}}{p_i^{\text{st}}}, & \text{elsewhere} \end{cases} \quad (194)$$

$$= \begin{cases} -w_i, & \text{if } i = j \\ w_{j_\mu i_\mu}, & \text{if } i = i_\mu, j = j_\mu \\ w_{i_\mu j_\mu}, & \text{if } i = j_\mu, j = i_\mu \\ w_{ij} \frac{p_j^{\text{st}}}{p_i^{\text{st}}}, & \text{elsewhere} \end{cases} \quad (195)$$

Off-diagonal entries are positive, diagonal entries are negative, therefore all we need to prove is that columns add up to zero. Summing over j , for each i , we obtain

$$\begin{aligned} \sum_j \tilde{W}(\{Q_\mu\})_{j,i} &= -w_i + \sum_{j \neq i} [W_{\text{hid}}]_{ij} \frac{p_j}{p_i} \\ &\quad + \sum_j w_{ji} \sum_\mu (\delta_{i,i_\mu} \delta_{j,j_\mu} + \delta_{i,j_\mu} \delta_{j,i_\mu}) \end{aligned} \quad (196)$$

where we recognized in the last term the off-diagonal entries of the generator W_{hid} , and in the second term in this expression precisely the off-diagonal correction that is needed to have

$$\sum_j \tilde{W}(\{Q_\mu\})_{j,i} = \sum_j [W_{\text{hid}}]_{ij} \frac{p_j}{p_i} = 0 \quad (197)$$

where we used the fact that \vec{p} is the steady state of W_{hid} . \square

Notice that, according to Eq. (195), forward and hidden TR generators can be written respectively as in Eqs. (116), where W_{hid} is the generator obtained by removing all edges $\{i_\mu j_\mu\}$ and W_{mar} is the complementary generator of the dynamics on those edges only.

Let us mention that by the construction of the so-called “Doob transform,” that we will briefly discuss in the conclusions, Sect. 8, there is no impediment in producing a proof of Proposition 17 for any values $q_\mu = q_\mu^*$ in the locus of zeroes of the SCGF, as mentioned in the Parade.

Similarly to the analysis in Sect. 4.4, effective affinities can be given a graphical interpretation. We will only give a simple example. Consider the complete graph on four nodes

$$\begin{array}{c} 1 \text{ --- } 2 \\ | \quad \diagup \quad \diagdown \\ | \quad \diagdown \quad \diagup \\ | \quad \diagup \quad \diagdown \\ 3 \text{ --- } 4 \end{array} \quad (198)$$

$$Q_{12} = \log \frac{\begin{array}{c} \overleftrightarrow{\downarrow} \overleftrightarrow{\downarrow} \overleftrightarrow{\downarrow} \overleftrightarrow{\downarrow} \\ \overleftrightarrow{\uparrow} \overleftrightarrow{\uparrow} \overleftrightarrow{\uparrow} \overleftrightarrow{\uparrow} \end{array}}{\begin{array}{c} \overleftrightarrow{\downarrow} \overleftrightarrow{\downarrow} \overleftrightarrow{\downarrow} \overleftrightarrow{\downarrow} \\ \overleftrightarrow{\uparrow} \overleftrightarrow{\uparrow} \overleftrightarrow{\uparrow} \overleftrightarrow{\uparrow} \end{array}} =: \log \frac{\psi_{12}}{\psi_{21}} \quad (199a)$$

The right-hand side defines the hidden fluxes. We recognize in these expressions those Hill cycle fluxes that are *not* in common between the two currents, which explicitly read (setting $\pi(\mathcal{G}) \stackrel{!}{=} 1$)

$$\begin{aligned} \phi_{12} = \psi_{\setminus 12} - \psi_{\setminus 21} &+ \begin{array}{|c|} \hline \overrightarrow{\hspace{1cm}} \\ \hline \overleftarrow{\hspace{1cm}} \\ \hline \end{array} + \begin{array}{|c|} \hline \overrightarrow{\hspace{1cm}} \\ \hline \overrightarrow{\hspace{1cm}} \\ \hline \end{array} + \begin{array}{|c|} \hline \overrightarrow{\hspace{1cm}} \\ \hline \overrightarrow{\hspace{1cm}} \\ \hline \end{array} + \begin{array}{|c|} \hline \overrightarrow{\hspace{1cm}} \\ \hline \overrightarrow{\hspace{1cm}} \\ \hline \end{array} \\ &- \begin{array}{|c|} \hline \overrightarrow{\hspace{1cm}} \\ \hline \overleftarrow{\hspace{1cm}} \\ \hline \end{array} - \begin{array}{|c|} \hline \overrightarrow{\hspace{1cm}} \\ \hline \overrightarrow{\hspace{1cm}} \\ \hline \end{array} - \begin{array}{|c|} \hline \overrightarrow{\hspace{1cm}} \\ \hline \overrightarrow{\hspace{1cm}} \\ \hline \end{array} - \begin{array}{|c|} \hline \overrightarrow{\hspace{1cm}} \\ \hline \overrightarrow{\hspace{1cm}} \\ \hline \end{array} \end{aligned} \quad (200a)$$

$$\begin{aligned} \phi_{34} = \psi_{\setminus 34} - \psi_{\setminus 43} &+ \begin{array}{|c|} \hline \overrightarrow{} \\ \hline \overleftarrow{} \\ \hline \end{array} + \begin{array}{|c|} \hline \overleftrightarrow{} \\ \hline \overleftrightarrow{} \\ \hline \end{array} + \begin{array}{|c|} \hline \overrightarrow{} \\ \hline \overrightarrow{} \\ \hline \end{array} + \begin{array}{|c|} \hline \overleftarrow{} \\ \hline \overleftarrow{} \\ \hline \end{array} \\ &- \begin{array}{|c|} \hline \overleftarrow{} \\ \hline \overleftarrow{} \\ \hline \end{array} - \begin{array}{|c|} \hline \overleftrightarrow{} \\ \hline \overleftrightarrow{} \\ \hline \end{array} - \begin{array}{|c|} \hline \overrightarrow{} \\ \hline \overrightarrow{} \\ \hline \end{array} - \begin{array}{|c|} \hline \overleftarrow{} \\ \hline \overleftarrow{} \\ \hline \end{array}. \end{aligned} \quad (200b)$$

Notice that from these expressions it is not so obvious (to the best of our understanding) that the mean marginal EPR $\phi_{12}Q_{12} + \phi_{34}Q_{34}$, should be non-negative, as we could prove directly in Sect. 4.5 for the single-edge case. This fact will rather follow from the IFR.

We derive the FR at all times using the formalism of the tilted generator and of the generating function, rather than the direct derivation in terms of path p.d.f.'s given in the previous Episode.

Proposition 18 *The marginal FR*

$$\frac{P(\{\Phi_\mu\})}{\tilde{P}(-\{\Phi_\mu\})} = \exp \sum Q_\mu \Phi_\mu \quad (201)$$

holds at all times, provided the initial configuration is sampled with probability $P(i^0 \equiv i) = p_i^{\text{st}}$.

Proof Rather than considering the probability itself, we consider the moment generating function of the currents $\zeta^t(\{q_\mu\})$ at time t , and let $\tilde{\zeta}^t(\{q_\mu\}) = (\zeta_i^t(\{q_\mu\}))_i$ be the moment generating function conditioned to being at site i at time t , defined as

$$\zeta_i^t(\{q_\mu\}) := \int \prod d\Phi_\mu P(\Phi_\mu^t \equiv \Phi_\mu, i^t \equiv i) \exp\left(-\sum \Phi_\mu q_\mu\right), \quad (202)$$

such that the moment generating function is recovered from the conditional one by $\zeta^t(\{q_\mu\}) = \vec{1} \cdot \tilde{\zeta}^t(\{q_\mu\})$. We can also introduce a conditional moment generating function for the hidden TR process $\tilde{\zeta}$. These two conditional functions evolve respectively by

$$\frac{d}{dt} \tilde{\zeta}^t(\{q_\mu\}) = M(\{q_\mu\}) \tilde{\zeta}^t(\{q_\mu\}), \quad (203a)$$

$$\frac{d}{dt} \tilde{\zeta}^t(\{q_\mu\}) = \tilde{M}(\{q_\mu\}) \tilde{\zeta}^t(\{q_\mu\}). \quad (203b)$$

Let $\zeta_i^0(\{q_\mu\}) = \tilde{\zeta}_i^0(\{q_\mu\}) = p_i^0$ be the probability of finding the system in site i at time $t = 0$. Defining $U^t(\{q_\mu\}) := \exp t M(\{q_\mu\})$ and $\tilde{U}^t(\{q_\mu\}) := \exp t \tilde{M}(\{q_\mu\})$ we have

$$\tilde{\zeta}^t(\{q_\mu\}) = U^t(\{q_\mu\}) \vec{p}^0. \quad (204)$$

Similarly, for the hidden TR dynamics, evaluating at $q_\mu \rightarrow Q_\mu - q_\mu$ we obtain

$$\tilde{\zeta}^t(Q_\mu - q_\mu) = \tilde{U}^t(Q_\mu - q_\mu) \vec{p}^0 \quad (205)$$

$$= P_{\text{st}} \left[\exp t M(\{q_\mu\})^T \right] P_{\text{st}}^{-1} \vec{p}^0. \quad (206)$$

We can now find the moment generating functions as

$$\zeta^t(\{q_\mu\}) = \vec{1} \cdot U^t(\{q_\mu\}) \vec{p}^0 \quad (207a)$$

$$\tilde{\zeta}^t(Q_\mu - q_\mu) = \vec{p}^0 \cdot P_{\text{st}}^{-1} U^t(\{q_\mu\}) P_{\text{st}} \vec{1} \quad (207b)$$

from which it follows that the choice of $\vec{p}^0 = \vec{p}^{\text{st}}$ makes the latter two expressions identical at all times, thus yielding an all-time fluctuation symmetry. Eq. (201) follows by properties of the Laplace transform. \square

As a straightforward corollary, we have the following asymptotic FRs.

Proposition 19 *The following marginal symmetry of the SCGF and IFR hold*

$$\lambda(\{Q_\mu - q_\mu\}) = \tilde{\lambda}(\{q_\mu\}), \quad (208)$$

$$\lambda(\{Q_\mu\}) = 0. \quad (209)$$

Proof The first follows from

$$\lambda(\{q_\mu\}) = \lim_{t \rightarrow \infty} \frac{1}{t} \zeta^t(\{q_\mu\}). \quad (210)$$

The second follows by evaluating at $q_\mu = Q_\mu$. \square

Just like the single-edge-current FR actually holds at the level of trajectories as in Eq. (173), we can also generalize the above proposition to a “complete” set of currents as follows.

Proposition 20 *The SCGF $\lambda(\{q_\mu\}_\mu, \{q_\alpha\}_{\alpha \geq |\mu|})$ of a “complete” set of currents satisfies the FR*

$$\lambda(\{q_\mu\}_\mu, \{q_\alpha\}_{\alpha \neq \mu}) = \tilde{\lambda}(\{Q_\mu - q_\mu\}_\mu, \{-q_\alpha\}_{\alpha \neq \mu}). \quad (211)$$

Proof Let us construct the tilted generator of the forward and backward dynamics in terms of the Hadamard product (see p.45) as (we drop the explicit range of the indices)

$$\begin{aligned}\tilde{M}(\{q_\mu\}, \{q_\alpha\}) &= P_{\text{st}} W_{\text{hid}}^T P_{\text{st}}^{-1} \circ T(\{q_\alpha\}) + W_{\text{mar}} \circ T(\{q_\mu\}) \\ &= P_{\text{st}} [W_{\text{hid}} \circ T(\{-q_\alpha\})]^T P_{\text{st}}^{-1} + W_{\text{mar}} \circ T(\{q_\mu\})\end{aligned}\quad (212)$$

where we used the fact that $T(\{q_\alpha\})^T = T(\{-q_\alpha\})$. We have

$$\begin{aligned}M(\{Q_\mu - q_\mu\}, \{-q_\alpha\}) &= \\ &= W_{\text{hid}} \circ T(\{q_\alpha\}) + W_{\text{mar}} \circ T(\{q_\mu\}) \\ &= W_{\text{hid}} \circ T(\{-q_\alpha\}) + W_{\text{mar}} \circ T(\{Q_\mu\}) \circ T(\{-q_\mu\}) \\ &= W_{\text{hid}} \circ T(\{-q_\alpha\}) + P_{\text{st}} W_{\text{mar}}^T P_{\text{st}}^{-1} \circ T(\{-q_\mu\}) \\ &= W_{\text{hid}} \circ T(\{-q_\alpha\}) + P_{\text{st}} [W_{\text{mar}} \circ T(\{q_\mu\})]^T P_{\text{st}}^{-1}\end{aligned}\quad (213)$$

where in the first passage we used the entry-wise associative property of the Hadamard product and in the second the main proposition Proposition 17. We then obtain

$$M(\{Q_\mu - q_\mu\}, \{-q_\alpha\}) = P_{\text{st}} \tilde{M}(\{q_\mu\}, \{q_\alpha\})^T P_{\text{st}}^{-1}\quad (214)$$

and the conclusion follows from matrix similarity. \square

5.4 The Hierarchy of Marginal Theories

Let us now suppose we follow a systematic procedure by which we first consider the marginal theory along edge $i_1 j_1$, then we expand our knowledge to $i_2 j_2$, and so on up to $i_{|\alpha|} j_{|\alpha|}$ when we cover a “complete” set of chord currents. Let $Q_\mu^{1, \dots, |\mu|}$ be the effective affinity along edge $i_\mu j_\mu$ obtained from the $|\mu|$ -th marginal theory in the hierarchy, where the superscript $1, \dots, |\mu|$ denotes the set of marginal currents. Each such theory gives an estimate of the entropy production rate

$$\sigma_{1, \dots, |\mu|} := \sum_{\mu=1}^{|\mu|} Q_\mu^{1, \dots, |\mu|} \phi_\mu.\quad (215)$$

The following proposition is a generalization of the result presented in Ref. [58].

Proposition 21 *The following hierarchy of inequalities is satisfied:*

$$0 \leq \sigma_1 \leq \sigma : 1, 2 \leq \dots \leq \sigma_{1, 2, \dots, |\alpha|}.\quad (216)$$

Proof Notice that the Proposition 20 can be written in terms of the currents’ p.d.f.’s as

$$\frac{P(\{\Phi_\alpha\}_{\alpha=1}^{|\alpha|})}{\tilde{P}(\{\Phi_\alpha\}_{\alpha=1}^{|\alpha|})} = \exp \left(\sum_{\alpha=1}^{|\alpha|} \mathcal{A}_\alpha \Phi_\alpha - \sum_{\mu=1}^{|\mu|} Q_\mu^{1, \dots, |\mu|} \Phi_\mu \right).\quad (217)$$

Let $\tilde{P}^{1, \dots, |\mu|}(\{\Phi_\mu\}_{\mu=1}^{|\mu'|})$ denote the probability of the first $|\mu'|$ currents according to the hidden TR dynamics of the $|\mu|$ -th theory, for $|\mu| > |\mu'|$ (notice that this is a new object: so far we only considered the probability of the first $|\mu|$ currents in the $|\mu|$ -th theory, and did

not specified the level of the hierarchy). Since the latter equality holds for all $|\mu|$, by dividing two such relations obtained for $|\mu|$ and $|\mu'|$, we obtain

$$\frac{\tilde{P}^{|\mu|}(\{\Phi_\mu\}_{\mu=1}^{|\mu|})}{\tilde{P}^{|\mu'|}(\{\Phi_\mu\}_{\mu=1}^{|\mu|})} = \exp\left(\sum_{\mu=1}^{|\mu|} Q_\mu^{1,\dots,|\mu|} \Phi_\mu - \sum_{\mu=1}^{|\mu'|} Q_\mu^{1,\dots,|\mu'|} \Phi_\mu\right) \quad (218)$$

We can then follow the usual route, obtaining the corresponding IFRs and using the Jensen inequality. \square

5.5 Stalling and Response at Stalling

The following result characterizes the stalling steady state.

Proposition 22 *The effective affinities vanish if and only if the mean marginal currents vanish.*

Proof This is evident given the explicit expression for the current $\phi_\mu = w_{i_\mu j_\mu} p_{j_\mu} - w_{j_\mu i_\mu} p_{i_\mu}$, and the fact that \vec{p}^{st} is the unique stalling state. \square

Notice that, with the hierarchy of theories in mind, the set of rates for which current ϕ_1 vanishes includes the set of rates for which currents ϕ_1, ϕ_2 vanish, and so on, until one reaches equilibrium. In this respect, equilibrium is just the stalling state of a marginal theory that happens to be complete.

Let us now parametrize the rates $w_{ij} \rightarrow w_{ij}(\mathbf{x})$ by a set of parameters $\mathbf{x} = (\{x_\mu\}_\mu, \mathbf{z})$ such that the first $|\mu|$ are *marginally thermodynamic* in the sense that

$$\frac{w_{i_\mu j_\mu}(\mathbf{x})}{w_{j_\mu i_\mu}(\mathbf{x})} = \exp x_\mu \quad (219)$$

and that no other rate depends on x_μ . The effective affinity now takes the form

$$Q_\mu(x) = x_\mu - x_\mu^{\text{st}}(\mathbf{z}) \quad (220)$$

where

$$x_\mu^{\text{st}}(\mathbf{z}) := \log \frac{p_{j_\mu}^{\text{st}}(\mathbf{z})}{p_{i_\mu}^{\text{st}}(\mathbf{z})}. \quad (221)$$

Notice that the stalling value of the μ -th parameter $x_\mu^{\text{st}}(\mathbf{z})$ might depend on the other parameters. The following result is crucial for an operational definition of the effective affinities.

Proposition 23 *At fixed dynamic parameters \mathbf{z} , the stalling steady state obtained by removing the observable edges is the same as the stalling steady state obtained by tuning the thermodynamic parameters to the stalling values $x_\mu \stackrel{!}{=} x_\mu^{\text{st}}(\mathbf{z})$.*

Proof Let us evaluate the forward generator on the stalling steady state:

$$\begin{aligned} W(\mathbf{x}) \vec{p}^{\text{st}}(\mathbf{x}) &= [W_{\text{hid}}(\mathbf{z}) + W_{\text{mar}}(\mathbf{x})] \vec{p}^{\text{st}} \\ &= W_{\text{mar}}(\mathbf{x}) \vec{p}^{\text{st}} \\ &= (0, \dots, \phi_\mu(\mathbf{x}), \dots, -\phi_\mu(\mathbf{x}), \dots, 0)^T \end{aligned} \quad (222)$$

where we used the definition of stalling $W_{\text{hid}} \vec{p}^{\text{st}} = 0$ and the fact that the hidden generator does not depend on the thermodynamic parameters. In the resulting vector we only highlighted

the entries corresponding to i_μ and j_μ . Then, in the light of Proposition 22 and of Eq. (220), this expression vanishes whenever the effective affinities vanish. \square

We now exploit the marginal fluctuation symmetry Eq. (208), that in its parametric form reads

$$\tilde{\lambda}(\{q_\mu\}; \mathbf{x}) = \lambda(x_\mu - x_\mu^{\text{st}}(z) - q_\mu; \mathbf{x}). \quad (223)$$

Taking the second mixed derivatives with respect to $q_\mu, q_\nu, x_\mu, x_\nu$ (omitting explicit dependencies) we obtain

$$\frac{\partial^2 \tilde{\lambda}}{\partial q_\mu \partial q_\nu} = \frac{\partial^2 \lambda}{\partial q_\mu \partial q_\nu}, \quad (224a)$$

$$\frac{\partial^2 \tilde{\lambda}}{\partial q_\mu \partial x_\nu} = -\frac{\partial^2 \lambda}{\partial q_\mu \partial q_\nu} - \frac{\partial^2 \lambda}{\partial q_\mu \partial x_\nu}, \quad (224b)$$

$$\frac{\partial^2 \tilde{\lambda}}{\partial x_\nu \partial x_\mu} = \frac{\partial^2 \lambda}{\partial x_\nu \partial x_\mu} + \frac{\partial^2 \lambda}{\partial q_\mu \partial q_\nu} + \frac{\partial^2 \lambda}{\partial x_\mu \partial q_\nu} + \frac{\partial^2 \lambda}{\partial q_\mu \partial x_\nu}. \quad (224c)$$

Evaluating at $\{q_\mu = 0\}$ and at $\{x_\mu = x_\mu^{\text{st}}\}$ we obtain the marginal FDRs

$$\tilde{\phi}_{\mu\nu} = \phi_{\mu\nu} = \phi_{\nu;\mu}^{\text{st}} + \phi_{\mu;\nu}^{\text{st}} \quad (225a)$$

$$= \tilde{\phi}_{\nu;\mu}^{\text{st}} + \tilde{\phi}_{\mu;\nu}^{\text{st}} \quad (225b)$$

$$= \phi_{\mu;\nu}^{\text{st}} + \tilde{\phi}_{\mu;\nu}^{\text{st}} \quad (225c)$$

$$= \phi_{\nu;\mu}^{\text{st}} + \tilde{\phi}_{\nu;\mu}^{\text{st}} \quad (225d)$$

and the marginal RRs

$$\phi_{\mu;\nu}^{\text{st}} = \tilde{\phi}_{\nu;\mu}^{\text{st}}. \quad (226a)$$

We can also consider the response to a perturbation of the κ -th kinetic parameter. By taking the first total derivative with respect to x_κ and evaluating at $\{q_\mu = 0\}$ we obtain the orthogonality relation

$$\sum_\mu \phi_\mu \frac{dx_\mu^{\text{st}}}{dx_\kappa} = 0. \quad (227)$$

The second mixed derivative with respect to q_μ, x_κ and to $x_\kappa, x_{\kappa'}$ yield

$$\phi_{\mu;\kappa} + \tilde{\phi}_{\mu;\kappa} = \sum_\nu \phi_{\mu\nu} \frac{dx_\nu^{\text{st}}}{dx_\kappa} \quad (228)$$

$$\sum_\mu \phi_{\mu;\kappa} \frac{dx_\mu^{\text{st}}}{dx_{\kappa'}} + \sum_\mu \phi_{\mu;\kappa'} \frac{dx_\mu^{\text{st}}}{dx_\kappa} = \sum_{\mu,\nu} \phi_{\mu\nu} \frac{dx_\nu^{\text{st}}}{dx_\kappa} \frac{dx_\mu^{\text{st}}}{dx_{\kappa'}}. \quad (229)$$

As regards higher-order response relations, it is clear that the treatment given in Sect. 3.5.5 can all be reproduced, with the only *caveat* that on the right-hand side of Eq. (104) there should appear the response coefficients of the hidden TR dynamics. We will not address this explicitly.

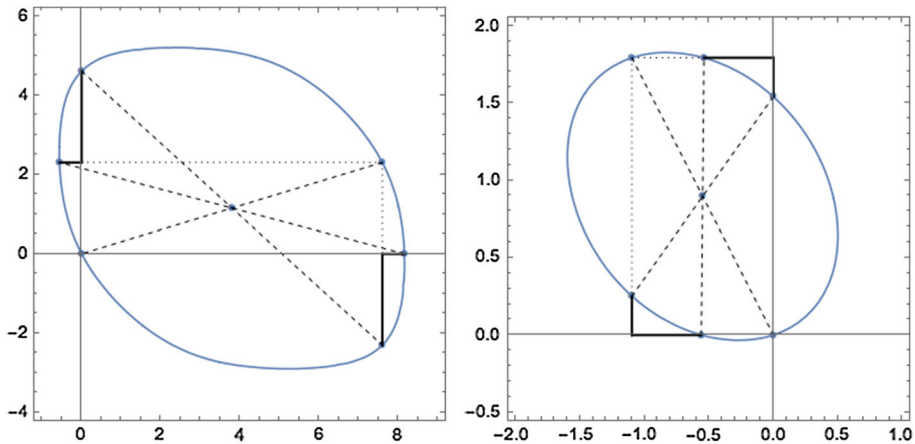


Fig. 2 Level curve (q_1^*, q_2^*) where $\lambda(q_1^*, q_2^*) = 0$, for two models. Bullets along the curves, starting from the origin and going counterclockwise, are for the plot on the left $(0, 0)$, $(\mathcal{A}_1, \mathcal{A}_2 - Q_2)$, $(Q_1, 0)$, $(\mathcal{A}_1, \mathcal{A}_2)$, $(0, Q_2)$, $(\mathcal{A}_1 - Q_1, \mathcal{A}_2)$, and for the plot on the right $(0, 0)$, $(0, Q_2)$, $(\mathcal{A}_1 - Q_1, \mathcal{A}_2)$, $(\mathcal{A}_1, \mathcal{A}_2)$, $(\mathcal{A}_1, \mathcal{A}_2 - Q_2)$, $(Q_1, 0)$. At the center of the figures is the symmetry point $(\mathcal{A}_1/2, \mathcal{A}_2/2)$. The model on the left has rates $w_{21} = w_{14} = w_{32} = w_{43} = w_{42} = 1$; $w_{12} = w_{41} = w_{23} = 10$; $w_{34} = w_{24} = 20$. The model on the right has rates $w_{21} = w_{41} = w_{23} = w_{43} = 1$; $w_{14} = w_{32} = w_{24} = 10$; $w_{34} = 20$; $w_{42} = 30$

5.6 Inequalities Between Effective Affinities

In this section we briefly sketch some ideas on how to derive inequalities between effective affinities at different orders in the hierarchy.

We first consider the case of a system with two “real” affinities $\mathcal{A}_1 (= Q_1^{1,2})$ and $\mathcal{A}_2 (= Q_2^{1,2})$ supported on the chords $i_1 j_1$ and $i_2 j_2$. Two marginal theories can be considered, one with effective affinity $Q_1 (= Q_1^1)$ defined along edge $i_1 j_1$ and one with $Q_2 (= Q_2^2)$ defined along edge $i_2 j_2$.

Proposition 24 *Either both “real” affinities are smaller in modulus than the effective affinities, $|\mathcal{A}_1| \leq |Q_1|$ and $|\mathcal{A}_2| \leq |Q_2|$, or they are both larger, $|\mathcal{A}_1| \geq |Q_1|$ or $|\mathcal{A}_2| \geq |Q_2|$.*

Proof We sketch an idea of the proof based on visual inspection of the two plots in Fig. 2. They represent level curves where the SCGF vanishes, $\lambda(q_1^*, q_2^*) = 0$, for two specific models. Six points are highlighted along the curve: the origin $(0, 0)$ the effective affinities $(Q_1, 0)$ and $(0, Q_2)$, and their dual under the mirror symmetry with respect to the point $(\mathcal{A}_1/2, \mathcal{A}_2/2)$, which need to belong to the curve as a consequence of the fluctuation symmetry. Any level curve of a convex function is a convex set. Consider the two highlighted segments. Their coordinates are either $Q_1 - \mathcal{A}_1 > 0$ and $Q_2 - \mathcal{A}_2 > 0$ for the left-hand curve, or $\mathcal{A}_2 - Q_2 > 0$, $Q_1 - \mathcal{A}_1 > 0$ for that on the right. Since for the right curve both $\mathcal{A}_1, Q_1 < 0$, we are consistent with the claim of the proposition. A similar analysis of other special cases will convince the reader it is not possible to draw six such special points along a convex curve that do not satisfy the claim. \square

Pushing this method beyond $|\alpha| = 2$ is far from trivial and it would make for an interesting mathematical inquiry. Let us briefly sketch some thoughts concerning the case $|\alpha| = 3$. In Fig. 3 we plot the polytopes obtained by triangulations of the points $(0, 0, 0)$ —the universal zero of the SCGF—, the point $(Q_1^{1,2,3}, Q_2^{1,2,3}, Q_3^{1,2,3}) =: A$ —the “real” affinities, the

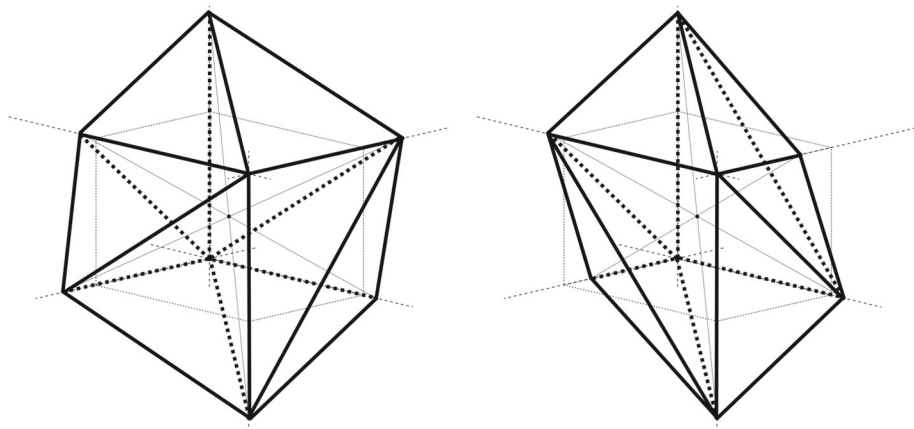


Fig. 3 In the space (q_1, q_2, q_3) , two polytopes having the effective affinities corresponding to level $|\alpha| = 1$ and $|\alpha| = 3$ marginal theories as their vertices. The image illustrates the conjectured fact that there always exist convex Delaunay triangulations of the vertices for any values of such effective affinities. However, the affinities corresponding to the $|\alpha| = 2$ will need to satisfy constraints

points $(Q_1^1, 0, 0)$, $(0, Q_2^2, 0)$, $(0, 0, Q_3^3)$ —the effective affinities of the single-edge marginal theories—and their mirror images through point $A/2$ (the small dot at the center of the polygon)—given by $(Q_1^3 - Q_1^1, Q_2^3, Q_3^3)$, $(Q_1^3, Q_2^3 - Q_2^1, Q_3^3)$, $(Q_1^{1,2,3}, Q_2^{1,2,3}, Q_3^{1,2,3} - Q_3^1)$. In the plots, all (effective and “real”) affinities are positive, without loss of generality. In the left-hand plot effective affinities are all larger than the “real” affinities, $(Q_1^1 > Q_1^{1,2,3}, Q_2^2 > Q_2^{1,2,3}, Q_3^3 > Q_3^{1,2,3})$, in the right-hand one one of the effective affinities is smaller, $(Q_1^1 < Q_1^{1,2,3}, Q_2^2 > Q_2^{1,2,3}, Q_3^3 > Q_3^{1,2,3})$. In both cases one can find a convex envelope that includes all such points. Now consider that must belong to such envelope also the six further points corresponding to the second level in the hierarchy $|\mu| = 2$, that is the three $(Q_1^{1,2}, Q_2^{1,2}, 0)$, $(0, Q_2^{2,3}, Q_3^{2,3})$ and $(Q_1^{1,3}, 0, Q_3^{1,3})$, and their FR-duals $(Q_1^{1,2,3} - Q_1^{1,2}, Q_2^{1,2,3} - Q_2^{1,2}, 0)$, $(0, Q_2^{1,2,3} - Q_2^{2,3}, Q_3^{1,2,3} - Q_3^{2,3})$ and $(Q_1^{1,2,3} - Q_1^{1,3}, 0, Q_3^{1,2,3} - Q_3^{1,3})$. It is clear that requiring that there exists a convex polytope that also includes such points as corners would pose strict inter-affinity relationships, whose determination however goes far astray from our scope.

5.7 Gauge Invariance

We have the following result.

Proposition 25 *Under the condition that the removal of the marginal edges does not disconnect the network, the effective affinities are invariant under a gauge transformation of the rates $w_{ij} \rightarrow w'_{ij} = w_{ij} e^{-a_j}$.*

Proof By the matrix-tree theorem, the effective affinities can be expressed in terms of rooted oriented spanning trees on the graph where the observable edges are removed. Notice that for every rooted oriented spanning tree \mathcal{T}_k with root at site k , the gauge transformation gives

$$\tau'(\mathcal{T}_k) = \prod_{ij \in \mathcal{T}_k} w'_{ij} = \prod_{i \neq k} e^{a_i} \prod_{ij \in \mathcal{T}_k} w_{ij} = e^{-a_k + a} \tau(\mathcal{T}_k) \quad (230)$$

where $a = \sum_i a_i$. This is due to the fact that in a rooted oriented spanning tree, each site but the root is the tail of exactly one directed edge. This factor drops from the expression of the steady probability because of normalization. Thus

$$w'_{ij} p_j^{\text{st}'} = w'_{ji} p_i^{\text{st}'} \quad (231)$$

and therefore the effective affinities are invariant. \square

Since currents are also gauge-invariant, it follows that the effective EPR is also invariant. Gauges affect the shape of the initial state that needs to be chosen for an all-time FR to hold.

A different, but related, acceptance of gauge invariance of the SCGF is that put forward by Wachtel et al. [110]. Let us review it here for it will have implications as regards the marginal theory, that we will analyze in Sect. 5.9.

Proposition 26 *The SCGF for all the currents $\lambda(\{q_{ij}\}_{ij \in \mathcal{E}})$ is invariant under the gauge transformation $q_{ij} \rightarrow q_{ij} + \sum c_{ij}^{\alpha*} Q_{\alpha*}$, where the vectors $\{c_{ij}^{\alpha*}\}_{ij \in \mathcal{E}}$ span the cocycle space of the graph.*

Proof Let us omit to specify $ij \in \mathcal{E}$. Using a trick employed in Ref. [64] to prove the FR, we consider the characteristic polynomial of tilted generator

$$\Delta(s; \{q_{ij}\}) := \det(W(\{q_{ij}\}; \mathbf{x}) - sI). \quad (232)$$

We employ the well-known expansion of the determinant of a matrix A [116]

$$\det A = \sum_{\pi} (-1)^{|\pi|} \prod_{i \in \mathcal{J}} A_{i, \pi(i)} \quad (233)$$

in terms of permutations $\pi : \mathcal{J} \rightarrow \mathcal{J}$ the graph's sites, where $|\pi|$ is the permutation's parity. Any permutation admits a cycle decomposition [116] in terms of singlets $i' \in \mathcal{J}^{\text{sing}}(\pi) \subseteq \mathcal{J}$ such that $\pi(i') = i'$, and a certain number $|\iota|$ of simple cycles, $\mathcal{C}_{\iota}(\pi)$, which do not cross each other and which cover all of the remaining sites in $\mathcal{J} \setminus \mathcal{J}^{\text{sing}}(\pi)$. One then obtains

$$\prod_{i \in \mathcal{J}} A_{i, \pi(i)} = \prod_{i' \in \mathcal{J}^{\text{sing}}(\pi)} A_{i' i'} \cdot \prod_{\iota} \prod_{ij} A_{ij}^{c_{\iota}^{ij}(\pi)}, \quad (234)$$

yielding

$$\begin{aligned} & \Delta\left(s; \{q_{ij} + \sum c_{ij}^{\alpha*} Q_{\alpha*}\}\right) \\ &= \sum_{\pi} (-1)^{|\pi|} \prod_{i' \in \mathcal{J}^{\text{sing}}(\pi)} (-s - w_{i'}) \cdot \prod_{\iota} \prod_{ij} (w_{ij} e^{-q_{ij}})^{c_{\iota}^{ij}(\pi)} \\ & \quad \times \exp - \sum_{\alpha*, \iota} Q_{\alpha*} \sum_{ij} c_{ij}^{\alpha*} c_{\iota}^{ij}(\pi). \end{aligned} \quad (235)$$

The latter term is unity because cycles are orthogonal to cocycles, and therefore the characteristic polynomial is gauge invariant:

$$\Delta\left(s; \{q_{ij} + \sum c_{ij}^{\alpha*} Q_{\alpha*}\}\right) = \Delta(s; \{q_{ij}\}) \quad (236)$$

In fact, it follows that all of the eigenvalues are also gauge invariant, including the SCGF. \square

The relationship between gauge invariance of cycle affinities and gauge invariance of the SCGF is as follows. A gauge transformation of the rates as described above transforms the log-ratio of the rates as $\log w_{ij}/w_{ji} \rightarrow \log w_{ij}/w_{ji} + a_i - a_j$ (a “gauge connection”). Any such transformation can be seen as a cocyclic transformation, as it can be shown that $a_i - a_j = \sum_{\alpha^*} c_{ij}^{\alpha^*} a_{\alpha^*}$ for some a_{α^*} . Indeed, using the FR for all the currents and imposing that it only depends on the cycle affinities, one obtains the latter invariance.

5.8 Marginally Thermodynamic Parametrizations

The local parametrization Eq. (219) automatically complies with the requirement B0 mentioned in the Parode. In this paragraph we analyze in more detail which parametrizations are *thermodynamic* in the sense that they satisfy the requirement A0 for the “complete” theory and, within such parametrization schemes, which are *marginally thermodynamic*. More precisely, in the first case our goal is to find the most general parametrization of the rates $\mathbf{x} \rightarrow w_{ij}(\mathbf{x})$ such that

$$\frac{\partial}{\partial x_\kappa} \mathcal{A}_\alpha(\mathbf{x}) = \delta_{\alpha,\kappa}. \quad (237)$$

This question will find a conclusive answer in Proposition 27, while Proposition 28 characterizes the marginal parametrizations that satisfy

$$\frac{\partial}{\partial x_{\mu'}} Q_\mu(\mathbf{x}) = \delta_{\mu,\mu'}. \quad (238)$$

Notice that while Eq. (237) involves all of the parameters, the second is much looser since it only involves the first $|\mu|$ parameters, and in general

$$\frac{\partial}{\partial x_\kappa} Q_\mu(\mathbf{x}) \neq 0, \quad \kappa > |\mu|. \quad (239)$$

Let us now pave the way to Propositions 27 and 28. It is an obvious fact that transition rates can be uniquely decomposed as

$$w_{ij}(\mathbf{x}) = v_{ij}(\mathbf{x}) e^{-a_{ij}(\mathbf{x})}, \quad (240)$$

with an antisymmetric term $a_{ij} = -a_{ji}$ and a symmetric positive term $v_{ij} = v_{ji} > 0$. Clearly, by definition “real” affinities are only sensible to transformations of the antisymmetric part. Furthermore, as explained in Sect. 5.7, they are invariant under so-called *gauge transformations* $a_{ij} \rightarrow a'_{ij} = a_{ij} + a_i - a_j$ [75,110], hence the general dependence that the symmetric part of the rates can take is

$$a_{ij}(\mathbf{x}) = a'_{ij}(\{x_\alpha\}) + a_j(\mathbf{x}) - a_i(\mathbf{x}). \quad (241)$$

in terms of some *gauge connection* a'_{ij} that only depends on the first $|\alpha|$ parameters, and of a *pure gauge* a_j on which, for the moment, we impose no restriction. Notice that the affinity can be written in terms of the gauge connection only:

$$\mathcal{A}_\alpha(\{x_{\alpha'}\}) = \sum_{ij \in \mathcal{C}_\alpha} a'_{ij}(\{x_{\alpha'}\}). \quad (242)$$

Since the effect of a variation of the thermodynamic parameters on the affinities is via the gauge connection, we now need to impose that variations of x_α only modify the α -th affinity according to Eq. (237). The basic intuition is that any perturbation of the gauge connection

along edges that are shared among different cycles will modify the affinities of both cycles. Let us then introduce the *peripheral* set as the set of edges that belong to cycle \mathcal{C}_α and to no other fundamental cycle $\mathcal{D}_\alpha = \mathcal{C}_\alpha \cap (\mathcal{E} \setminus \bigcup_{\alpha' \neq \alpha} \mathcal{C}_{\alpha'})$. In practice, these consist of the generating chord e_α and of all edges that are only separated from the chord by vertices of degree 2 (sometimes called bridges [20]). Then, we can at best distribute the dissipative parameter x_α among the edges of the corresponding peripheral set, in such a way that

$$a'_{ij}(\{x_\alpha\}) = \begin{cases} a''_{ij} + \lambda_\alpha^{ij} x_\alpha, & ij \in \mathcal{D}_\alpha \\ a''_{ij}, & ij \notin \bigcup_\alpha \mathcal{D}_\alpha \end{cases} \quad (243)$$

where a''_{ij} are independent of all parameters and $\lambda_\alpha^{ij} = -\lambda_\alpha^{ji}$ are real numbers such that, for all α ,

$$\sum_{ij \in \mathcal{D}_\alpha} \lambda_\alpha^{ij} = 1. \quad (244)$$

This provides the general structure of a thermodynamic parametrization. Notice that no constraint is imposed on the dependence of the symmetric part and of the pure gauge on the parameters.

Proposition 27 (Thermodynamic parametrization) *A parametrization of the transition rates is thermodynamic, in the sense that it satisfies Eq. (237), if and only if there exist (possibly non-unique) symmetric terms, a gauge connection and a pure gauge obeying Eq. (243) such that*

$$w_{ij}(\mathbf{x}) = \sqrt{v_{ij}(\mathbf{x})} \exp \frac{1}{2} \left[a_j(\mathbf{x}) - a_i(\mathbf{x}) + a'_{ij}(\{x_\alpha\}) \right]. \quad (245)$$

Proof Sufficiency is trivial: from Eq. (242) we have

$$\mathcal{A}_\alpha(\{x_{\alpha'}\}) = \sum_{ij \in \mathcal{D}_\alpha} \lambda_\alpha^{ij} x_\alpha + \sum_{ij \in \mathcal{C}_\alpha} a''_{ij} = x_\alpha + \sum_{ij \in \mathcal{C}_\alpha} a''_{ij}. \quad (246)$$

Necessity is inbuilt in the above construction: the symmetric/antisymmetric splitting is unique; perturbations of the antisymmetric part that do not affect the affinities are accounted for by the pure gauge; hence we can only focus on any explicit dependencies of the gauge connection on the parameters that will affect the affinities. In particular, any perturbation of the rates along an edge that is in common among several cycles will affect the affinities of both cycles, therefore violating thermodynamic parametrization. \square

Notice that, as anticipated in Sect. 2.1, the symmetric term and the pure gauge, to which we referred as the *internal landscape*, can undergo global arbitrary transformations, while once the pure gauge is fixed, the parametrization of the gauge connection is local and quite strict.

We now want to inquire how much stricter is a marginally thermodynamic parametrization obeying Eq. (238) with respect to a thermodynamic parametrization. We already know that effective affinities are gauge invariant. Yet they do depend on the kinetic parameters all over the network. Then, all we need is to make the dependency of the symmetric terms local like that of the gauge connection.

Proposition 28 (Marginal thermodynamic parametrization) *A parametrization of the transition rates is marginally thermodynamic, in the sense that it satisfies Eq. (239), if it is thermodynamic and, additionally, if symmetric terms only depend locally on the thermodynamic parameters, meaning that only those along $ij \in \mathcal{D}_\mu$ depend on x_μ .*

Proof All boils down to showing that the effective affinity can be written as

$$Q_\mu(\mathbf{x}) = \log \left(\frac{p_{i'_\mu}^{\text{st}}(\mathbf{x})}{p_{j'_\mu}^{\text{st}}(\mathbf{x})} \prod_{ij \in \mathcal{D}_\mu} \frac{w_{ij}}{w_{ji}} \right) \quad (247)$$

where i'_μ and j'_μ are respectively the source of the first edge in \mathcal{D}_μ and the target of the last edge in \mathcal{D}_μ . Notice that these depend in a contrived way on all other rates not belonging to \mathcal{D}_μ . Having singled out the rates in \mathcal{D}_μ , we notice that symmetric terms still cancel among themselves, hence we conclude. \square

This latter proposition is not “if and only if” only because of subtle special cases of no practical interest.

5.9 Disconnecting the Configuration Space

So far we have assumed that the removal of the edges that support observable currents does not disconnect the configuration space. Let us relax this assumption.

The removal of edges from the network may result into several disconnected subgraphs \mathcal{G}_κ . If that occurs, then each connected subgraph has its own unique normalized steady state $\vec{p}^{\text{st},\kappa}$ and any convex combination of them

$$\vec{p}^{\text{st}}(\boldsymbol{\pi}) = \sum_\kappa \pi_\kappa \vec{p}^{\text{st},\kappa}, \quad \sum_\kappa \pi_\kappa = 1 \quad (248)$$

is a steady state for the stalling dynamics. Here $\pi_\kappa > 0$ is the probability of finding the system in subgraph \mathcal{G}_κ at the moment of the preparation of the initial ensemble. As regards the case of several currents each supported on one edge analyzed in Sect. 5, the effective affinities introduced in Eq. (192) are still well-defined, and an inspection of the proof of Proposition 17 and of Proposition 19 reveals that all that matters to go through is that $\vec{p}^{\text{st}}(\boldsymbol{\pi})$ is *some* steady state. The only main difference is that in all expressions there will be an explicit dependence on the parameters $\boldsymbol{\pi}$, in particular via $P_{\text{st}}(\boldsymbol{\pi})$ and the effective affinities $Q_\mu(\boldsymbol{\pi})$. Consider the function $i \rightarrow \kappa(i)$ mapping a site i to the connected component it belongs to. We then have

$$Q_\mu(\boldsymbol{\pi}) = \log \frac{w_{i_\mu j_\mu} p_{j_\mu}^{\text{st},\kappa(j_\mu)}}{w_{j_\mu i_\mu} p_{i_\mu}^{\text{st},\kappa(i_\mu)}} + \log \frac{\pi_{\kappa(j_\mu)}}{\pi_{\kappa(i_\mu)}} =: Q'_\mu + \log \frac{\pi_{\kappa(j_\mu)}}{\pi_{\kappa(i_\mu)}} \quad (249)$$

where the right-hand side defines a preferred “ground” value of the effective affinity Q'_μ . The finite-time FR holds unmodified, provided one chooses the right effective affinity and the right initial ensemble (as we will analyze in a future publication, and has already been observed in the case of the full FR by Rao [73], in the preparation there is a difference between removing edges, and tuning parameters to values where the currents stall).

As regards the asymptotic FR, notice that in the derivation of the fluctuation symmetry the matrix $P_{\text{st}}(\boldsymbol{\pi})$ does not enter the game as it only contributes a similarity transformation. Hence the full dependence of the relation on the $\boldsymbol{\pi}$ is through the effective affinities:

$$\tilde{\lambda}(\{q_\mu\}) = \lambda(\{Q_\mu(\boldsymbol{\pi}) - q_\mu\}). \quad (250)$$

But since the left-hand side of this expression does not depend on $\boldsymbol{\pi}$ altogether, we obtain that the SCGF has a symmetry

$$\lambda(\{q_\mu\}) = \lambda(\{q_\mu + \log \pi_{\kappa(j_\mu)} / \pi_{\kappa(i_\mu)}\}). \quad (251)$$

This also implies that certain linear combinations of mean currents always vanish at the steady state. Consider for example the case where two subgraphs are separated by two observable edges:

$$\begin{array}{c} \text{I} \\ \curvearrowright \\ \mathcal{G}_1 \quad \mathcal{G}_2 \\ \curvearrowleft \\ \text{II} \end{array} . \quad (252)$$

The symmetry of the rate function reads

$$\lambda(\{q_I, q_{II}\}) = \lambda(\{q_I + \alpha, q_{II} + \alpha\}). \quad (253)$$

where $\alpha = \log \pi_2 / \pi_1$. Taking the derivative with respect to α and evaluating at $q_I = q_{II} = \alpha = 0$ we obtain $\phi_I + \phi_{II} = 0$.

Notice that Proposition 22 then needs to be reformulated, as it is not sufficient that the effective affinity vanishes for some values of the π to have a vanishing stalling current. This is a consequence of the marginal gauge invariance of the SCGF, that we now analyze.

The SCGF of all the marginal currents is obtained by setting $q_{ij} = 0$ for all $ij \in \mathcal{E} \setminus \mathcal{E}_{\text{mar}}$. With reference to Proposition 26, clearly any gauge transformation that involves cocycles that have edges in the hidden sector would break this condition, thus it is not allowed. But, since by definition the removal of a (simple) cocycle disconnects the graph in two subgraphs (we can always take the basis cocycles to be simple in the sense described in Sect. 3.4.3), then if the marginal configuration space contains a cocycle, then the hidden configuration space gets disconnected. In fact, in the above example the two edges form a simple cocycle. We thus see that marginal gauge invariance coincides with the freedom of choice of the stalling steady state in the definition of the effective affinities. While, according to the tenets of Schnakenberg's theory, when dealing with single-edge currents one can always take observable currents along a subset of chords, and thus arrive at uniquely defined and invariant affinities, in the case of currents defined over multiple edges it might often be the case that gauge invariance plays a role. This depends on the specific context.

6 Episode 4: Phenomenological Currents

Finally, we consider a set of marginal phenomenological currents. The theory exposed in the previous sections holds unchanged provided an additional strict condition, that we call *marginal consistency*, is verified. We characterize marginal consistency in physical terms in terms of hidden entropy production.

6.1 Marginal Phenomenological Currents

Phenomenological currents are linear combinations of edge currents

$$\Phi_{\mu}^t = \sum_{ij} \varphi_{\mu}^{ij} \Phi_{ij}^t, \quad (254)$$

where $\varphi_{\mu}^{ij} = -\varphi_{\mu}^{ji}$. Let $\mathcal{E}_{\mu} = \{ij \in \mathcal{E} \mid \varphi_{\mu}^{ij} \neq 0\}$ be the set of edges that support the μ -th marginal current, and $\mathcal{E}_{\text{mar}} := \bigcup_{\mu} \mathcal{E}_{\mu}$ be the *marginal edge set* of all edges that support a marginal phenomenological current. Furthermore let \mathcal{J}_{mar} be set of sites of the graph that are

boundaries of some edge in \mathcal{E}_{mar} . For sake of notational simplicity, we assume that $\mathcal{E} \setminus \mathcal{E}_{\text{mar}}$ is connected; the general case can be built after the considerations in Sect. 5.9.

We assume that the phenomenological currents defined above are independent, in the sense that there is no vector $\ell = (\ell_\mu)_\mu$ such that

$$\sum_{\mu} \ell_{\mu} \Phi_{\mu}^t \asymp 0 \quad (255)$$

in the long time limit (at steady state). This amounts to ask that the matrix $(\sum_{ij} \varphi_{\mu}^{ij} c_{ij}^{\alpha})_{\alpha, \mu}$ has full rank. The theory of FRs for a complete set of phenomenological currents in the presence of conservation laws has been analyzed in Ref. [71] at the steady state, and in Ref. [73] at finite times. The generalization of our marginal theory to systems with conserved quantities, either within or across the marginal/hidden sectors of the configuration space, is an interesting direction for future research.

6.2 Marginal Consistency

Let $\Lambda(\{q_{\mu}\})$ be the SCGF of the marginal phenomenological currents with respect to the forward dynamics. It is the dominant eigenvalue of the tilted operator $M(\{q_{\mu}\})$ obtained by replacing the off-diagonal entries w_{ij} by $w_{ij} \exp -\sum \varphi_{\mu}^{ij} q_{\mu}$ and by keeping the diagonal entries the same. Similarly we can introduce the SCGF $\tilde{\Lambda}(\{q_{\mu}\})$ of the currents with respect to the hidden TR dynamics.

Also, let $\lambda(\{q_{ij}\}_{ij \in \mathcal{E}_{\text{mar}}})$, $\tilde{\lambda}(\{q_{ij}\}_{ij \in \mathcal{E}_{\text{mar}}})$ be the SCGFs of all the edge currents in the marginal sector of the configuration space. We hereby exploit the contraction principle in the theory of large deviations, which states that we can obtain the SCGF of a coarser observable by replacing $q_{ij} \rightarrow \sum \varphi_{\mu}^{ij} q_{\mu}$,

$$\Lambda(\{q_{\mu}\}) = \lambda \left(\left\{ \sum \varphi_{\mu}^{ij} q_{\mu} \right\}_{ij \in \mathcal{E}_{\text{mar}}} \right), \quad (256a)$$

$$\tilde{\Lambda}(\{q_{\mu}\}) = \tilde{\lambda} \left(\left\{ \sum \varphi_{\mu}^{ij} q_{\mu} \right\}_{ij \in \mathcal{E}_{\text{mar}}} \right). \quad (256b)$$

By the theory exposed in Sect. 5, we know that the marginal fluctuation symmetry holds

$$\lambda(\{q_{ij}\}_{ij \in \mathcal{E}_{\text{mar}}}) = \tilde{\lambda}(\{Q_{ij} - q_{ij}\}_{ij \in \mathcal{E}_{\text{mar}}}), \quad (257)$$

where the effective affinities are given by the usual expression $Q_{ij} = \log w_{ij} p_j^{\text{st}} / w_{ji} p_i^{\text{st}}$, the stalling steady state being defined as the unique steady state in the network where all edges in \mathcal{E}_{mar} are removed.

We now inquire under which conditions this relation extends to the SCGFs of the marginal phenomenological currents.

Proposition 29 *A sufficient condition for the marginal phenomenological fluctuation symmetry to hold is the condition of marginal consistency, i.e. that there exist phenomenological effective affinities Q_{μ} such that*

$$Q_{ij} = \sum \varphi_{\mu}^{ij} Q_{\mu}, \quad \forall ij \in \mathcal{E}_{\text{mar}}. \quad (258)$$

Proof Straightforward given the definitions and the contraction principle:

$$\begin{aligned}\Lambda(\{q_\mu\}) &= \lambda \left(\left\{ \sum \varphi_\mu^{ij} q_\mu \right\}_{ij \in \mathcal{E}_{\text{mar}}} \right) = \tilde{\lambda} \left(\left\{ \sum \varphi_\mu^{ij} (Q_\mu - q_\mu) \right\}_{ij \in \mathcal{E}_{\text{mar}}} \right) \\ &= \tilde{\Lambda}(\{Q_\mu - q_\mu\}).\end{aligned}\quad (259)$$

□

The condition is not strictly necessary, for the reason noted in Sect. 5.9 that the effective affinities might not be unique when the removal of the observable edges disconnects the network, which is usually the case for phenomenological currents.

Marginal consistency poses a strong constraint on the shape of the stalling steady state. In fact, for all sites $i \in \mathcal{I}_{\text{mar}}$ in the marginal edge set, the stalling steady state must satisfy

$$\frac{p_j^{\text{st}}}{p_i^{\text{st}}} = \frac{w_{ji}}{w_{ij}} \exp \sum \varphi_\mu^{ij} Q_\mu. \quad (260)$$

This balance will not be granted for any thermodynamically consistent model, as we will show by a simple example in Sect. 6.5. At the highest level of the hierarchy of marginal theories, where the effective affinities are the “real” affinities of the complete theory, the above condition boils down to prescribing certain values for the rate ratio w_{ij}/w_{ji} , up to a pure gauge, see Eq. (66). This is due to the peculiar property of equilibrium steady states of writing just in terms of w_{ij}/w_{ji} , see Eq. (62). Instead, whilst gauge invariant, the effective affinities do not depend directly on the rate ratio. As a consequence, marginal consistency requires a fine tuning of the rates all over the network. This entails an unprecedented relationship between symmetric and antisymmetric contributions to the rates, i.e. an interplay between the physics that shapes the internal landscape of the system, and the thermodynamic parameters.

Proposition 30 *If the condition of marginal consistency holds, then*

$$\tilde{W} := P_{\text{st}} M(\{Q_\mu\})^T P_{\text{st}}^{-1} \quad (261)$$

is a MJPG.

Proof We have, using Eq. (260)

$$\tilde{w}_{ij} = \frac{p_i^{\text{st}}}{p_j^{\text{st}}} w_{ji} \exp \sum \varphi_\mu^{ij} Q_\mu = w_{ij}, \quad \forall ij \in \mathcal{E}_{\text{mar}} \quad (262a)$$

$$\tilde{w}_{ij} = \frac{p_i^{\text{st}}}{p_j^{\text{st}}} w_{ji}, \quad \forall ij \notin \mathcal{E}_{\text{mar}}. \quad (262b)$$

The proof proceeds as usual by summing over the columns of \tilde{W} and using the fact that \vec{p}^{st} is the stalling state. □

From this result follows that, if the condition of marginal consistency holds, basically everything that has been discussed in Sect. 5 holds without modifications.

6.3 Internal Stalling

The above condition of marginal consistency so far remains formal. Notice that a state of *internal stalling*, where all the microscopic currents supporting the marginal phenomenological currents vanish, is also a state of *phenomenological stalling*:

$$\phi_{ij} = 0, \forall ij \in \mathcal{E}_\mu \Rightarrow \phi_\mu = 0. \quad (263)$$

The inverse implication is not generally true.

Proposition 31 *If marginal consistency holds, then at a phenomenological stalling steady state, internal currents stall.*

Proof By the fluctuation symmetry, if $Q_\mu = 0$ then $\phi_\mu = 0$. From Eq. (258) follows that if $Q_\mu = 0$ then $Q_{ij} = 0$, and therefore $\phi_{ij} = 0$. \square

While the reverse implication does not hold, for the simple reason that stalling states are a subclass of all possible marginally consistent states, the following statement gives a physical characterization of marginal consistency.

Proposition 32 *If at a phenomenological stalling steady state the marginal stochastic entropy production vanishes, then the effective affinities satisfy marginal consistency.*

Proof We are requiring that

$$\sum_{ij \in \mathcal{E}_{\text{mar}}} Q_{ij} \Phi_{ij}^t = 0 \quad (264)$$

vanishes for all values of $\Phi_{ij}^t \equiv t\phi_{ij}$ satisfying the steady state condition $\partial\phi = 0$ (or, equivalently, up to $O(1)$ transient contributions). Let us impose phenomenological stalling. We have

$$\sum_{ij \in \mathcal{E}_\mu} \varphi_\mu^{ij} \phi_{ij} = 0. \quad (265)$$

Therefore ϕ is in the kernel of matrix $\varphi = (\varphi_\mu^{ij})_{ij \in \mathcal{E}_{\mu,\mu}}$. Therefore it can be an arbitrary vector in the orthogonal complement of the image of φ . Since Eq. (264) must vanish for any such vector, then Q_{ij} must live in the image of φ , that is, satisfy Eq. (258). \square

6.4 Response

Let us endow ourselves with a marginally thermodynamic parametrization x_{ij} of the edges $ij \in \mathcal{E}_{\text{mar}}$. We can consider the phenomenological currents' covariance at phenomenological stalling

$$\phi_{\mu\mu'}^{\text{phen. st.}} = \sum_{ij, ij'} \varphi_\mu^{ij} \varphi_{\mu'}^{ij'} \phi_{ij, ij'}^{\text{phen. st.}}. \quad (266)$$

Notice that we would not be able to proceed further if phenomenological stalling did not imply internal stalling, because we do not generally have a FDR for nonvanishing currents. Under the assumption of marginal consistency we obtain

$$\begin{aligned} \phi_{\mu\mu'}^{\text{st}} &= \sum_{ij, ij'} \varphi_\mu^{ij} \varphi_{\mu'}^{ij'} \left(\frac{\partial}{\partial x_{ij}} \phi_{ij'}^{\text{st}} + \frac{\partial}{\partial x_{ij'}} \phi_{ij}^{\text{st}} \right) \\ &= \nabla_\mu \phi_{\mu'}^{\text{st}} + \nabla_{\mu'} \phi_\mu^{\text{st}} \end{aligned} \quad (267)$$

where

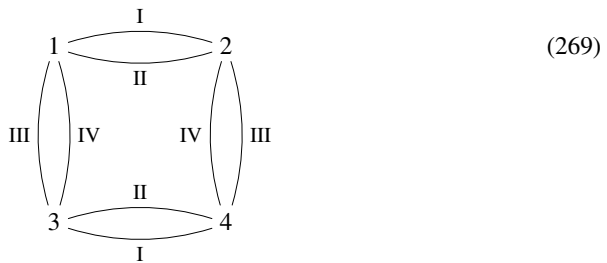
$$\nabla_\mu := \sum_{ij} \varphi_\mu^{ij} \frac{\partial}{\partial x_{ij}} \quad (268)$$

is the directional (Lie) derivative along vector $\{\varphi_\mu^{ij}\}_{ij}$. This means that there is a preferred set of directions where the variation of the parameters should be performed, hence that we need to sub-parametrize rates in terms of parameters x_μ in a way that $\nabla_\mu = \partial/\partial x_\mu$ becomes a proper derivative.

6.5 Complete Versus Marginal Consistency

In this section we show by a simple example that marginal consistency is stricter “complete” consistency. While this example is not representative of a physical apparatus, the gap between marginal and complete thermodynamic consistency appears to us to be the most interesting open question left aside in this paper.

We take into consideration a simple model whose configuration space is depicted by the following graph with rates labeled by a reservoir index taking values in $\rho = \text{I, II, III, IV}$:



where we assume that the rates satisfy

$$w_{12}^{\text{I}}/w_{21}^{\text{I}} = w_{34}^{\text{I}}/w_{43}^{\text{I}} \quad (270a)$$

$$w_{12}^{\text{II}}/w_{21}^{\text{II}} = w_{34}^{\text{II}}/w_{43}^{\text{II}} \quad (270b)$$

$$w_{13}^{\text{III}}/w_{31}^{\text{III}} = w_{24}^{\text{III}}/w_{42}^{\text{III}} \quad (270c)$$

$$w_{13}^{\text{IV}}/w_{31}^{\text{IV}} = w_{24}^{\text{IV}}/w_{42}^{\text{IV}} \quad (270d)$$

The physical rationale is that rates of type ρ are due to the interaction with a bath at inverse temperature β^ρ and satisfy the condition of local detailed balance [100,117]

$$\frac{w_{ij}}{w_{ji}} = \exp(-\beta^\rho(\epsilon_i - \epsilon_j)). \quad (271)$$

In the above model it is implied that the energy gaps $\epsilon_1 - \epsilon_2 = \epsilon_3 - \epsilon_4$ are the same. The physical currents are given by (dropping the dependency on t)

$$\Phi^\rho = \Phi_{12}^\rho + \Phi_{34}^\rho, \quad \rho = \text{I, II} \quad (272a)$$

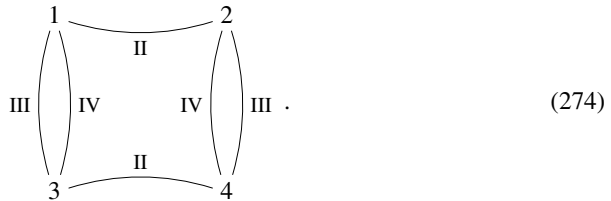
$$\Phi^\rho = \Phi_{13}^\rho + \Phi_{24}^\rho, \quad \rho = \text{III, IV}. \quad (272b)$$

The above parametrization Eq. (270) grants that the entropy production along a trajectory is a functional of the phenomenological currents only. The situation is different for the effective

description of an observer who, for example, only measures Φ^I . The effective affinities along the two edges supporting this current are given by

$$Q_{12}^I = \log \frac{w_{12}^I p_2^{\text{st}}}{w_{21}^I p_1^{\text{st}}}, \quad Q_{34}^I = \log \frac{w_{34}^I p_4^{\text{st}}}{w_{43}^I p_3^{\text{st}}} \quad (273)$$

where \vec{p}^{st} is the stalling steady state on the network



Marginal consistency requires that the two effective affinities coincide, $Q_{12}^I = Q_{34}^I$, which is only the case if

$$\frac{p_1^{\text{st}}}{p_2^{\text{st}}} = \frac{p_3^{\text{st}}}{p_4^{\text{st}}}. \quad (275)$$

This condition is *not* implied by local detailed balance and it requires a fine tuning of rates all over the network. This has broad implications on the overall flows measured in the system. In the above example, notice that in view of Eq. (270b), marginal thermodynamic consistency implies that the two steady forces along edges II should also coincide:

$$\log \frac{w_{21}^{\text{II}} p_1^{\text{st}}}{w_{12}^{\text{II}} p_2^{\text{st}}} = \log \frac{w_{43}^{\text{II}} p_3^{\text{st}}}{w_{34}^{\text{II}} p_4^{\text{st}}}. \quad (276)$$

As a consequence, at the stalling steady state both transitions of type II are in same direction. This is obviously impossible, therefore the steady currents $\phi_{12}^{\text{II}}, \phi_{34}^{\text{II}}$ must vanish as well at stalling. The only current at a phenomenological stalling steady state is along cycles formed by edges of type III and IV.

7 Stasimon: A Negative Result

Marginal currents play a role in the formulation of so-called *uncertainty relations* between a current's variance and the full EPR, recently formulated [82] and proven under quite general conditions [83,84]. Remarkably, the results hold for marginal currents of any kind, though it has been argued that the bound is only strict when the current is the entropy production itself [85]. Therefore it is interesting to inquire whether stricter bounds in terms of marginal measures of the entropy production, rather than the full entropy production, might hold [85].

The uncertainty relation states that for any current

$$\frac{\langle \phi^2 \rangle}{\langle \phi \rangle^2} \geq \frac{2}{\sum \mathcal{A}_\alpha \langle \phi_\alpha \rangle}. \quad (277)$$

The bound is significative when the current is the entropy production, $\phi = \sum \mathcal{A}_\alpha \phi_\alpha$. Otherwise the bound performs poorly, and one might want to lower the measure of EPR so to

make the right-hand side of this inequality as high as possible. Since $\sum \mathcal{F}_\alpha \langle \phi_\alpha \rangle \geq \mathcal{Q} \langle \phi \rangle$, one tempting hypothesis is that the effective EPR might also satisfy the bound:

$$\frac{\langle \phi^2 \rangle}{\langle \phi \rangle} \geq \frac{2}{\mathcal{Q}}. \quad (278)$$

The uncertainty relation derives from the following quadratic bound on the rate function [83]

$$\lambda(\{q_\alpha\}) \geq \sum q_\alpha \langle \phi_\alpha \rangle \left(\frac{\sum q_{\alpha'} \langle \phi_{\alpha'} \rangle}{\sum \mathcal{A}_{\alpha''} \langle \phi_{\alpha''} \rangle} - 1 \right). \quad (279)$$

Choosing, for definitiveness, the first current as our observable $\phi = \phi_1$, and setting $q_\alpha = \delta_{\alpha,1} q$, one obtains

$$\lambda(q) \geq q \langle \phi \rangle \left(\frac{q \langle \phi \rangle}{\sigma} - 1 \right). \quad (280)$$

The bound is found by taking second derivatives and evaluating at $q = 0$. In our marginal theory, the quadratic function $q \langle \phi \rangle (q/\mathcal{Q} - 1)$ that would yield the relation Eq.(278) in a similar way does not bound $\lambda(q)$, rather it approximates it as a quadratic taking the same vanishing values at $q = 0, \mathcal{Q}$. In fact, not only the bound on the SCGF does not hold, but by randomly inspecting the behavior of several example systems we were able to find several cases where there is a (mild) violation of the uncertainty relation. One such case is given by the generator

$$W = \begin{pmatrix} -15 & 2 & 6 & 1 \\ 3 & -12 & 10 & 0 \\ 7 & 10 & -21 & 5 \\ 5 & 0 & 5 & -6 \end{pmatrix}. \quad (281)$$

Thus we rule out a reasonable hypothesis.

8 Exode: Conclusions

Generic considerations on future perspectives were presented in Sect. 2.3. In this section we provide a more prosaic list of partial results and of technical issues related to the derivations of the results that are either unresolved and/or that might be of some interest from a mathematical point of view.

8.1 Relationship to Doob's Transform

The tilted operator $M(\{q_\mu\})$ is generally not similar to a MJPG. This has implications, for example, for the efficient computation of the SCGF, which cannot be performed by a direct application of the Gillespie algorithm but requires more advanced techniques [118,119]. Nevertheless, for all values of the tilting fields $\{q_\mu\}$ one can build a MJPG $W(\{q_\mu\})$ out of the tilted operator $M(\{q_\mu\})$, by performing the so-called Doob transform [120]

$$W(\{q_\mu\}) = R(\{q_\mu\}) M(\{q_\mu\})^T R^{-1}(\{q_\mu\}) - \lambda(\{q_\mu\}) I, \quad (282)$$

where $R(\{q_\mu\})$ is the diagonal matrix whose entries are the components of the right Perron–Frobenius eigenvector of $M(\{q_\mu\})$. In a way, the typical currents according to this new

dynamics reproduce the statistics of rare currents of the original dynamics, somehow realizing Onsager's regression hypothesis. Furthermore, it can be shown that the diagonal matrix $R(\{q_\mu\})$ [resp. $L(\{q_\mu\})$] whose entries are the components of the right [resp. left] eigenvector of $M(\{q_\mu\})$ can be interpreted as the probability of a fluctuation conditioned on being at some configuration at the final [resp. initial] time [81].

Our theory implies that the hidden TR generator \tilde{W} is the Doob transform evaluated at $\lambda(\{Q_\mu\}) = 0$, where the Doob's transform and the forward generator are related by a similarity transformation. Furthermore, this implies that

$$P_{\text{st}} = R(\{Q_\mu\}), \quad (283)$$

which yields a different interpretation of the right eigenvector for that particular value of the tilting parameters.

A complementary perspective on tilting techniques regards Markov processes that are *conditioned* upon observing a certain event, as thoroughly discussed in Ref. [121]. While such conditional processes are not Markovian on their own, there exists the possibility of constructing Markovian generators that best reproduce their behavior. Then the question is open what particular rare events are typical of hidden time-reversal generators, and what kind of conditioning on Markov processes they represent.

Knowledge of the effective affinity might be implemented in algorithms for the reconstruction of the SCGF based on cloning and/or on iterative procedures, e.g. to pinpoint certain specific values of the SCGF.

8.2 Direct Derivations of the FDR and Activity

Let us consider the derivation of nonequilibrium FDRs carried over e.g. in Ref. [94] and compare it to our own approach at stalling steady states. We focus on the single-edge scenario detailed in Sect. 4. The trajectory p.d.f. Eq. (81) can be written in terms of two terms, one that is time-symmetric and one that is time-antisymmetric:

$$P[\omega^t](x) = \exp \frac{1}{2} (\Theta[\omega^t](x) + \Sigma[\omega^t](x)) P_{i_0}^0. \quad (284)$$

where Θ is a suitably defined time-symmetric term that contains both the Poisson-type waiting-time distribution, and the activities (symmetrized fluxes). We made explicit the dependency on parameter x . We have

$$\frac{\partial}{\partial x} \Sigma[\omega^t](x) = \Phi_{12}[\omega^t]. \quad (285)$$

Let's look at the response of the average steady-state current:

$$\frac{\partial}{\partial x} \phi_{12}(x) = \lim_{t \rightarrow \infty} \frac{1}{t} \int \mathcal{D}\omega^t \Phi_{12}[\omega^t] \frac{\partial}{\partial x} P[\omega^t](x) \quad (286)$$

$$= \frac{1}{2} \langle \phi_{12}^2 \rangle + \lim_{t \rightarrow \infty} \frac{1}{2t} \left\langle \Phi_{12} \frac{\partial}{\partial x} \Theta(x) \right\rangle. \quad (287)$$

This shows that in general to the response of a current contributes the self-correlation and the correlation of the current with a time-symmetric observable, that contributes in important ways out of equilibrium.

Let us see how from this we can derive the FDR at equilibrium, that is, assuming there exists a value $x = x^{\text{eq}}$ such that $P[\omega^t](x^{\text{eq}}) = P[\bar{\omega}^t](x^{\text{eq}})$ (up to boundary terms). Key to the result is that the time symmetry is not affected by the derivative, therefore

$$\frac{\partial \Theta}{\partial x}[\omega^t](x^{\text{eq}}) = \frac{\partial \Theta}{\partial x}[\bar{\omega}^t](x^{\text{eq}}). \quad (288)$$

Then:

$$\begin{aligned} \left\langle \Phi_{12} \frac{\partial \Theta}{\partial x} \right\rangle^{\text{eq}} &= \int \mathcal{D}\omega^t P[\omega^t](x^{\text{eq}}) \Phi_{12}[\omega^t] \frac{\partial \Theta}{\partial x}[\omega^t](x^{\text{eq}}) \\ &= \int \mathcal{D}\omega^t P[\bar{\omega}^t](x^{\text{eq}}) \Phi_{12}[\bar{\omega}^t] \frac{\partial \Theta}{\partial x}[\bar{\omega}^t](x^{\text{eq}}) \\ &= - \left\langle \Phi_{12} \frac{\partial \Theta}{\partial x} \right\rangle^{\text{eq}} \end{aligned} \quad (289)$$

and therefore it must vanish.

It would be desirable to have a similar direct proof that the active response vanishes at stalling steady-states. We could use, instead of the time-reversed trajectory, the hidden time-reverse trajectory introduced in Sect. 4.9. However, if we go through the same passages as above the proof halts as we do not have a trajectory such that $P[\tilde{\omega}^t] = \tilde{P}[\omega^t]$ (not even at stalling). We believe that finding a proper way to deal with such problem would disclose a whole new set of techniques that could be used to analyze marginal systems.

Furthermore, it would be interesting to consider the large deviation functions of the joint activities and currents, which in the case of a “complete” theory is known analytically [122].

8.3 Gauge Invariance

In Quantum Field Theory, Wilson loops of gauge connections are the fundamental gauge-invariant quantities. Importantly, they satisfy the reconstruction property [123,124]: given the Wilson loops, one can reconstruct the gauge connection up to gauge transformations. The analogue of Wilson loops in nonequilibrium thermodynamics are the “real” affinities, as argued in Refs. [75,76]. When considering a system exchanging energy with several of heat reservoirs at different temperatures (possibly a continuum of them), “real” affinities take the well-known form $\oint \delta Q/k_B T$.

The marginal theory though deals with new objects, the effective affinities, that are not defined along a single loop, are not defined only in terms of the gauge connection alone, and somehow have a “renormalized” character, they are “dressed”. Yet they are gauge invariant and they constitute important observables. So, it would be interesting to ponder how this construction might go back to gauge theories, in particular as comes to non-Abelian gauge theories, to the Mandelstam identities, and to the reconstruction property.

8.4 The Deletion-Contraction Paradigm

Some passages in our theory offer a connection to the paradigm of deletion-contraction in algebraic graph theory [125], that has applications to as remote areas as knot polynomials [126] and Feynman diagrams [127,128]. Deletion-contraction formulas apply in particular to spanning-tree polynomials [129]. We notice in passing that most results in this area regard symmetric Laplacians, while MJPGs are not necessarily symmetric. Then, the deletion-contraction formula we employed in Eq. (127) rooted oriented spanning tree only holds if the root is one of the two contracted vertices. It is very easy to generate counterexamples for other roots. This opens up the question whether it is possible to prove more general deletion-contraction formulas for weighted oriented graphs.

Acknowledgements We are thankful to Bernhard Altaner for introducing the first version of the problem that inspired this work. We enjoyed discussions with Artur Wachtel and all other group members. MP also thanks the bartenders and attendees of Tramways for moral support and company, and Alexandra Elbakyan for furnishing a fundamental tool for scientific documentation. The research was supported by the European Research Council, project NanoThermo (ERC-2015-CoG Agreement No. 681456).

References

- Collins, H.: *Gravity's Ghost and Big Dog: Scientific Discovery and Social Analysis in the Twenty-First Century*. University of Chicago Press, Chicago (2014)
- Einstein, A.: *Autobiographical Notes* (1946), quoted in G. Holton and Y. Elkana, *Albert Einstein: Historical and Cultural Perspectives* (Princeton University Press, 1982)
- Seifert, U.: Stochastic thermodynamics, fluctuation theorems and molecular machines Rep. Progr. Phys. **75**, 126001 (2012)
- Collin, D., Ritort, F., Jarzynski, C., Smith, S.B., Tinoco, I., Bustamante, C.: Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. *Nature* **437**, 231 (2005)
- Nyquist, H.: Thermal agitation of electric charge in conductors. *Phys. Rev.* **32**, 110 (1928)
- Green, M.S.: Markoff random processes and the statistical mechanics of time-dependent phenomena. *J. Chem. Phys.* **20**, 1281 (1952)
- Onsager, L.: Reciprocal relations in irreversible processes. I. *Phys. Rev.* **37**, 405 (1931)
- Onsager, L.: Reciprocal relations in irreversible processes. II. *Phys. Rev.* **38**, 2265 (1931)
- Kubo, R.: The fluctuation-dissipation theorem. *Rep. Progr. Phys.* **29**, 255 (1966)
- Miller, D.G.: Thermodynamics of irreversible processes. The experimental verification of the Onsager reciprocal relations. *Chem. Rev.* **60**, 15 (1960)
- Bochkov, G.N., Kuzovlev, Y.E.: Nonlinear fluctuation-dissipation relations and stochastic models in nonequilibrium thermodynamics: I. Generalized fluctuation-dissipation theorem. *Physica A* **106**, 443 (1981)
- Bochkov, G.N., Kuzovlev, Y.E.: Nonlinear fluctuation-dissipation relations and stochastic models in nonequilibrium thermodynamics: II. Kinetic potential and variational principles for nonlinear irreversible processes. *Physica A* **106**, 480 (1981)
- Kurchan, J.: Fluctuation theorem for stochastic dynamics. *J. Phys. A* **31**, 3719 (1998)
- Maes, C.: The fluctuation theorem as a Gibbs property. *J. Stat. Phys.* **95**, 367 (1999)
- Lebowitz, J.L., Spohn, H.: A Gallavotti–Cohen-type symmetry in the large deviation functional for stochastic dynamics. *J. Stat. Phys.* **95**, 333 (1999)
- Allen, W.: *Whatever works* (2009)
- Gallavotti, G.: A local fluctuation theorem. *Physica A* **263**, 39–50 (1999)
- García-García, R., Sourabh, L., Lacoste, D.: Thermodynamic inference based on coarsened-grained data or noisy measurements. *Phys. Rev. E* **93**, 032103 (2016)
- Wächter, C., Strassberg, P., Brandes, T.: Stochastic thermodynamics based on incomplete information: generalized Jarzynski equality with measurement errors with or without feedback. *New J. Phys.* **18**, 113042 (2016)
- Altaner, B., Vollmer, J.: Fluctuation-preserving coarse graining for biochemical systems. *Phys. Rev. Lett.* **108**, 228101 (2012)
- Gupta, D., Sabhapandit, S.: Fluctuation theorem for entropy production of a partial system in the weak-coupling limit. *Eur. Phys. Lett.* **115**, 60003 (2016)
- Uhl, M., Pietzonka, P., Seifert, U.: Fluctuations of apparent entropy production in networks with hidden slow degrees of freedom. *J. Stat. Mech.* **2018**(2), 023203 (2018)
- Bravi, B., Sollich, P.: Statistical physics approaches to subnetwork dynamics in biochemical systems. *Phys. Biol.* **14**(4), 045010 (2017)
- Strasberg, P., Schaller, G., Brandes, T., Esposito, M.: Thermodynamics of a physical model implementing a Maxwell demon. *Phys. Rev. Lett.* **110**, 040601 (2013)
- Mandal, D., Jarzynski, C.: Work and information processing in a solvable model of Maxwell's demon. *Proc. Nat. Acad. Sci.* **109**, 11641 (2012)
- Parrondo, J.M.R., Horowitz, J.M., Sagawa, Takahiro: Thermodynamics of information. *Nat. Phys.* **11**, 131 (2015)
- Frenzel, M.F., Sagawa, T.: Coarse-grained hidden entropy production in partially inaccessible quantum jump trajectories. [arXiv:1609.08628](https://arxiv.org/abs/1609.08628)
- Crooks, G.E., Still, S.: Marginal and Conditional Second Laws of Thermodynamics. [arXiv:1611.04628](https://arxiv.org/abs/1611.04628)

29. Bechhoefer, J.: Hidden Markov models for stochastic thermodynamics. *New J. Phys.* **17**, 075003 (2015)
30. Partovi, M.: Entanglement versus Stosszahlansatz: disappearance of the thermodynamic arrow in a high-correlation environment. *Phys. Rev. E* **77**, 021110 (2008)
31. Bera, M.N., Riera, A., Lewenstein, M., Winter, A.: Generalized laws of thermodynamics in the presence of correlations. *Nat. Commun.* **8**, 2180 (2017)
32. Jarzynski, C.: Stochastic and macroscopic thermodynamics of strongly coupled systems. *Phys. Rev. X* **7**, 011008 (2017)
33. Cerino, L., Puglisi, A.: Entropy production for velocity-dependent macroscopic forces: the problem of dissipation without fluctuations. *Eur. Phys. Lett* **111**, 40012 (2015)
34. Mehl, J., Lander, B., Bechinger, C., Blickle, V., Seifert, U.: Role of hidden slow degrees of freedom in the fluctuation theorem. *Phys. Rev. Lett.* **108**, 220601 (2012)
35. Esposito, M.: Stochastic thermodynamics under coarse graining. *Phys. Rev. E* **85**, 041125 (2012)
36. Bo, S., Celani, A.: Multiple-scale stochastic processes: decimation, averaging and beyond. *Phys. Rep.* **670**, 1 (2017)
37. Wang, S.-W., Kawaguchi, K., Sasa, S.-i., Tang, L.-H.: Entropy production of nanosystems with time scale separation. *Phys. Rev. Lett.* **117**, 070601 (2016)
38. Puglisi, A., Pigolotti, S., Rondoni, L., Vulpiani, A.: Entropy production and coarse graining in Markov processes. *J. Stat. Mech.* E11001
39. Qian, H.: A simple theory of motor protein kinetics and energetics. *Biophys. Chem.* **263**, 67 (1997)
40. Kolomeisky, A.B., Fisher, M.E.: Simple mechanochemistry describes the dynamics of kinesin molecules. *Proc. Nat. Acad. Sci.* **98**, 7748 (2001)
41. Dubi, Y., Di Ventra, M.: Colloquium: heat flow and thermoelectricity in atomic and molecular junctions. *Rev. Mod. Phys.* **83**, 131 (2011)
42. Brandner, K., Seifert, U.: Multi-terminal thermoelectric transport in a magnetic field: bounds on Onsager coefficients and efficiency. *New J. Phys.* **15**, 105003 (2013)
43. Sánchez, R., Sothmann, B., Jordan, A.N., Büttiker, M.: Powerful energy harvester based on resonant-tunneling quantum wells. *New J. Phys.* **15**, 125001 (2013)
44. Yamamoto, K., Entin-Wohlman, O., Aharony, A., Hatano, N.: Efficiency bounds on thermoelectric transport in magnetic fields: the role of inelastic processes. *Phys. Rev. B* **94**, 121402 (2016)
45. Leitmann, S., Franosch, T.: Nonlinear response in the driven lattice lorentz gas. *Phys. Rev. Lett.* **111**, 190603 (2013)
46. Falasco, G., Cossetto, T., Penocchio, E., Esposito, M.: Negative differential response in chemical reactions. [arXiv:1812.11245](https://arxiv.org/abs/1812.11245)
47. Gomez-Solano, J.R., Petrosyan, A., Ciliberto, S., Chetrite, R., Gawędzki, K.: Experimental verification of a modified fluctuation-dissipation relation for a micron-sized particle in a nonequilibrium steady state. *Phys. Rev. Lett.* **103**, 040601 (2009)
48. Cugliandolo, L.F., Kurchan, J., Peliti, L.: Energy flow, partial equilibration, and effective temperatures in systems with slow dynamics. *Phys. Rev. E* **55**, 3898 (1997)
49. Crisanti, A., Picco, M., Ritort, F.: Fluctuation relation for weakly ergodic aging systems. *Phys. Rev. Lett.* **110**, 080601 (2013)
50. Poletтини, M., Esposito, M.: Effective thermodynamics for a marginal observer. *Phys. Rev. Lett.* **119**, 240601 (2017)
51. Altaner, B., Poletтини, M., Esposito, M.: Fluctuation-dissipation relations far from equilibrium. *Phys. Rev. Lett.* **117**, 180601 (2016)
52. Lau, A.W.C., Lacoste, D., Mallick, K.: Nonequilibrium fluctuations and mechanochemical couplings of a molecular motor. *Phys. Rev. Lett.* **99**, 158102 (2007)
53. Lacoste, D., Lau, A.W.C., Mallick, K.: Fluctuation theorem and large deviation function for a solvable model of a molecular motor. *Phys. Rev. E* **78**, 011915 (2008)
54. Bulnes Cuetara, G., Esposito, M., Schaller, G., Gaspard, P.: Effective fluctuation theorems for electron transport in a double quantum dot coupled to a quantum point contact. *Phys. Rev. B* **88**, 115134 (2013)
55. Shiraishi, N., Sagawa, T.: Fluctuation theorem for partially masked nonequilibrium dynamics. *Phys. Rev. E* **91**, 012130 (2015)
56. Rosinberg, M.L., Horowitz, J.M.: Continuous information flow fluctuations. *Eur. Phys. Lett.* **116**, 10007 (2016)
57. Hartich, D., Barato, A.C., Seifert, U.: Stochastic thermodynamics of bipartite systems: transfer entropy inequalities and a Maxwell's demon interpretation. *J. Stat. Mech.* **2**, P02016 (2014)
58. Bisker, G., Poletтини, M., Gingrich, T.R., Horowitz, J.M.: Hierarchical bounds on entropy production inferred from partial information. *J. Stat. Mech.* **2017**(9), 093210 (2017)
59. Biggs, N.: Algebraic graph theory. Cambridge University Press, Cambridge (1993)
60. Touchette, H.: The large deviation approach to statistical mechanics. *Phys. Rep.* **478**, 1 (2009)

61. Baez, A.V.: *The New College Physics: A Spiral Approach*. W. H. Freeman & Co., San Francisco (1967)
62. Poletini, M., Esposito, M.: Transient fluctuation theorem for the currents and initial equilibrium ensembles. *J. Stat. Mech.* **2014**(10), P10033 (2014)
63. Faggionato, A., Di Pietro, D.: Gallavotti–Cohen-type symmetry related to cycle decompositions for Markov chains and biochemical applications. *J. Stat. Phys.* **143**, 11 (2011)
64. Andrieux, D., Gaspard, P.: Fluctuation theorem for currents and Schnakenberg network theory. *J. Stat. Phys.* **127**, 107 (2007)
65. Maes, C.: On the origin and the use of fluctuation relations for the entropy. *Sémin. Poincaré* **2**, 29–62 (2003)
66. Andrieux, D., Gaspard, P.: Fluctuation theorem and Onsager reciprocity relations. *J. Chem. Phys.* **121**, 6167 (2004)
67. Andrieux, D., Gaspard, P.: A fluctuation theorem for currents and non-linear response coefficients. *J. Stat. Mech.* **2007**(02), P02006 (2007)
68. Gao, C.Ya., Limmer, D.T.: Nonlinear transport coefficients from large deviation functions. [arXiv:1812.01470](https://arxiv.org/abs/1812.01470)
69. Schnakenberg, J.: Network theory of microscopic and macroscopic behavior of master equation systems. *Rev. Mod. Phys.* **48**, 571 (1976)
70. Poletini, M.: Cycle/cocycle oblique projections on oriented graphs. *Lett. Math. Phys.* **105**, 89 (2015)
71. Poletini, M., Bulnes Cuetara, G., Esposito, M.: Conservation laws and symmetries in stochastic thermodynamics. *Phys. Rev. E* **94**, 052117 (2016)
72. Hänggi, P., Talkner, P., Borkovec, M.: Reaction-rate theory: fifty years after Kramers. *Rev. Mod. Phys.* **62**, 251 (1990)
73. Rao, R., Esposito, M.: Conservation laws shape dissipation. *New J. Phys.* **20**, 023007 (2018)
74. Poletini, M., Esposito, M.: Irreversible thermodynamics of open chemical networks I: emergent cycles and broken conservation laws. *J. Chem. Phys.* **141**, 024117 (2014)
75. Poletini, M.: Nonequilibrium thermodynamics as a gauge theory. *Eur. Phys. Lett.* **97**, 30003 (2012)
76. Poletini, M.: Of dice and men. Subjective priors, gauge invariance, and nonequilibrium thermodynamics. In: *Proceedings of the 12th Joint European Thermodynamics Conference* (2013)
77. Ben-Naim, A.: *A Farewell to Entropy: Statistical Thermodynamics Based on Information*. S. World Scientific, Singapore (2008)
78. Rao, R., Esposito, M.: Detailed fluctuation theorems: a unifying perspective. *Entropy* **20**, 635 (2018)
79. Hill, T.L.: *Free Energy Transduction and Biochemical Cycle Kinetics*. Dover, New York (2005)
80. Hill, T.L.: *Studies in irreversible thermodynamics*. IV. Diagrammatic representation of steady state fluxes for unimolecular systems. *J. Theor. Biol.* **10**, 442 (1966)
81. Lazarescu, A.: The physicist's companion to current fluctuations: one-dimensional bulk-driven lattice gases. *J. Phys. A* **48**, 503001 (2015)
82. Barato, A.C., Seifert, U.: Thermodynamic uncertainty relation for biomolecular processes. *Phys. Rev. Lett.* **114**, 158101 (2015)
83. Pietzonka, P., Barato, A.C., Seifert, U.: Universal bounds on current fluctuations. *Phys. Rev. E* **93**, 052145 (2016)
84. Gingrich, T.R., Horowitz, J.M., Perunov, N., England, J.: Dissipation bounds all steady-state current fluctuations. *Phys. Rev. Lett.* **116**, 120601 (2016)
85. Poletini, M., Lazarescu, A., Esposito, M.: Tightening the uncertainty principle for stochastic currents. *Phys. Rev. E* **94**, 052104 (2016)
86. Pal, A., Rahav, S.: Integral fluctuation theorems for stochastic resetting systems. *Phys. Rev. E* **96**, 062135 (2017)
87. Harada, T., Sasa, S.: Equality connecting energy dissipation with a violation of the fluctuation-response relation. *Phys. Rev. Lett.* **95**, 130602 (2005)
88. Harada, T., Sasa, S.: Energy dissipation and violation of the fluctuation-response relation in nonequilibrium Langevin systems. *Phys. Rev. E* **73**, 026131 (2006)
89. Poletini, M.: Macroscopic constraints for the minimum entropy production principle. *Phys. Rev. E* **84**, 051117 (2011)
90. Poletini, M.: Fact-checking Ziegler's maximum entropy production principle beyond the linear regime and towards steady states. *Entropy* **15**, 2570 (2013)
91. Murashita, Y., Funo, K., Ueda, M.: Nonequilibrium equalities in absolutely irreversible processes. *Phys. Rev. E* **90**, 042110 (2014)
92. Garrahan, J.P.: Simple bounds on fluctuations and uncertainty relations for first-passage times of counting observables. *Phys. Rev. E* **95**, 032134 (2017)
93. Baiesi, M., Falasco, G.: Inflow rate, a time-symmetric observable obeying fluctuation relations. *Phys. Rev. E* **92**, 042162 (2015)

94. Baiesi, M., Maes, C., Wynants, B.: Fluctuations and response of nonequilibrium states. *Phys. Rev. Lett.* **103**, 010602 (2009)
95. Baiesi, M., Maes, C., Wynants, B.: The modified Sutherland-Einstein relation for diffusive non-equilibria. *Proc. R. Soc. A* **467**, 2792 (2011)
96. Prost, J., Joanny, J.-F., Parrondo, J.M.R.: Generalized fluctuation-dissipation theorem for steady-state systems. *Phys. Rev. Lett.* **103**, 090601 (2009)
97. Spinney, R.E., Ford, I.J.: Nonequilibrium thermodynamics of stochastic systems with odd and even variables. *Phys. Rev. Lett.* **108**, 170603 (2012)
98. Speck, T., Seifert, U.: Restoring a fluctuation-dissipation theorem in a nonequilibrium steady state. *Europhys. Lett.* **74**, 391 (2006)
99. Chetrite, R., Gawędzki, K.: Eulerian and Lagrangian pictures of non-equilibrium diffusions. *J. Stat. Phys.* **137**, 890 (2009)
100. Esposito, M., van den Broeck, C.: The three faces of the second law: I. Master equation formulation. *Phys. Rev. E* **82**, 011143 (2010)
101. Gaveau, B., Schulman, L.S.: Master equation based formulation of nonequilibrium statistical mechanics. *J. Math. Phys.* **37**, 3897 (1996)
102. Barbier, M., Gaspard, P.: Microreversibility, nonequilibrium current fluctuations, and response theory. *J. Phys. A* **51**(35), 355001 (2018)
103. Chaiken, S.: A combinatorial proof of the all minors matrix tree theorem. *SIAM J. Algebr. Discret.* **3**, 319 (1982)
104. Poletini, M., Esposito, M.: Nonconvexity of the relative entropy for Markov dynamics: a Fisher information approach. *Phys. Rev. E* **88**, 012112 (2013)
105. Zia, R.K.P., Schmittmann, B.: Probability currents as principal characteristics in the statistical mechanics of nonequilibrium steady states. *J. Stat. Mech.* **2007**(07), P07012 (2007)
106. Anderson, D.F., Kurtz, T.G.: *Stochastic Analysis of Biochemical Systems*. Springer, Berlin (2015)
107. Ethier, S.N., Kurtz, T.G.: *Markov Processes: Characterization and Convergence*. Wiley, New York (2009)
108. Weber, M.F., Frey, E.: Master equations and the theory of stochastic path integrals. *Rep. Progr. Phys.* **80**, 046601 (2017)
109. Poletini, M.: BEST statistics of Markovian fluxes: a tale of Eulerian tours and Fermionic ghosts. *J. Phys. A* **48**, 365005 (2015)
110. Wachtel, A., Vollmer, J., Altaner, B.: Fluctuating currents in stochastic thermodynamics I. Gauge invariance of asymptotic statistics. *Phys. Rev. E* **92**, 042132 (2015)
111. Koza, Z.: General technique of calculating drift velocity and diffusion coefficient in arbitrary periodic systems. *J. Phys. A* **32**, 7637 (1999)
112. Altaner, B., Wachtel, A., Vollmer, J.: Fluctuating currents in stochastic thermodynamics II. Energy conversion and nonequilibrium response in Kinesin models. *Phys. Rev. E* **92**, 042133 (2015)
113. Jia, C., Jiang, D., Qian, M.: Cycle symmetries and circulation fluctuations for discrete-time and continuous-time Markov chains. *Ann. Appl. Probab.* **26**, 2454 (2016)
114. Bulnes-Cuetara, G., Esposito, M., Imparato, A.: Exact fluctuation theorem without ensemble quantities. *Phys. Rev. E* **89**, 052119 (2014)
115. Cover, T.M., Thomas, J.A.: *Elements of Information Theory*. Wiley, New York (2006)
116. Skiena, S.: The cycle structure of permutations, §1.2.4. In: *Implementing Discrete Mathematics: Combinatorics and Graph Theory with Mathematica*, pp. 20–24. Addison-Wesley, Reading, MA (1990)
117. Maes, C., Netočný, K., Wynants, B.: On and beyond entropy production: the case of Markov jump processes. *Markov Proc. Rel. Fields.* **14**, 445 (2008)
118. Nemoto, T., Bouchet, F., Jack, R.L., Lecomte, V.: Population-dynamics method with a multiccanonical feedback control. *Phys. Rev. E* **93**, 062123 (2016)
119. Giardinà, C., Kurchan, J., Peliti, L.: Direct evaluation of large-deviation functions. *Phys. Rev. Lett.* **96**, 120603 (2006)
120. Verley, G.: Nonequilibrium thermodynamic potentials for continuous-time Markov chains. *Phys. Rev. E* **93**, 012111 (2016)
121. Chetrite, R., Touchette, H.: Nonequilibrium Markov processes conditioned on large deviations. *Ann. Henri Poincaré* **16**, 2005 (2015)
122. Bertini, L., Faggionato, A., Gabrielli, D.: Flows, currents, and cycles for Markov Chains: large deviation asymptotics. *Stoch. Proc. Appl.* **125**, 2786 (2015)
123. Gambini, R., Pullin, J.: *Loops, Knots, Gauge Theories and Quantum Gravity*. Cambridge University Press, Cambridge (2000)
124. Giles, R.: Reconstruction of gauge potentials from Wilson loops. *Phys. Rev. D* **24**, 2160 (1981)
125. Sokal, A.D.: The multivariate Tutte polynomial (alias Potts model) for graphs and matroids. *Surv. Comb.* **327**, 173 (2005)

126. Kauffman, L.H.: A Tutte polynomial for signed graphs. *Discret. Appl. Math.* **25**, 105 (1989)
127. Nakanishi, N.: *Graph Theory and Feynman Integrals*. Gordon and Breach, New York (1971)
128. Aluffi, P., Marcolli, M.: Feynman motives and deletion-contraction relations. In: *Topology of Algebraic Varieties and Singularities: Conference in Honor of Anatoly Libgober's 60th Birthday*, vol. 538. American Mathematical Society, Providence (2011)
129. Bollobás, B., Pebody, L., Riordan, O.: Contraction-deletion invariants for graphs. *J. Combin. Theory Ser. B* **80**, 320 (2000)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.