Journal of Statistical Physics (2023) 190:119
https://doi.org/10.1007/s10955-023-03136-0

Trajectorial Dissipation of Φ-Entropies for Interacting Particle Systems

Benedikt Jahnel[1,2] · Jonas Köppl[2]

Received: 2 February 2023 / Accepted: 29 June 2023 / Published online: 11 July 2023
© The Author(s) 2023

Abstract
A classical approach for the analysis of the long-time behaviour of Markov processes is to consider suitable Lyapunov functionals like the variance or more generally Φ-entropies. Via purely analytic arguments it can be shown that these functionals are indeed non-increasing in time under quite general assumptions on the process. We refine these classical results via a more probabilistic approach and show that dissipation is already present on the level of individual trajectories for spatially extended systems of infinitely many interacting particles with arbitrary underlying geometry and compact local spin spaces. This extends previous results from the setting of finite-state Markov chains or diffusions in $\mathbb{R}^n$ to an infinite-dimensional setting with weak assumptions on the dynamics.

Keywords Interacting particle systems · Phi-entropy · Time-reversal · Martingale representation

Mathematics Subject Classification Primary 82C20 · Secondary 60K35

1 Introduction

There are many different techniques to study the long-time behaviour of Markov processes that excel in different situations. One very common and powerful technique is the use of Lyapunov functionals, i.e., functionals that are monotone in time. An example of such a functional is the variance

$$\mathbf{Var}_\mu(f) := \mathbb{E}_\mu[f^2] - \mathbb{E}_\mu[f]^2, \quad f \in L^2(\mu),$$

Communicated by Christian Maes.

✉ Jonas Köppl
jonas.koeppl@wias-berlin.de

Benedikt Jahnel
benedikt.jahnel@tu-braunschweig.de

1 Institut für Mathematische Stochastik, Technische Universität Braunschweig, Universitätsplatz 2, 38106 Brunswick, Germany

2 Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstraße 39, 10117 Berlin, Germany

where μ is an invariant measure for some Markov process $(X_t)_{t\geq 0}$ with semigroup $(P_t)_{t\geq 0}$. If we now fix an observable f and consider the function

$$[0,\infty) \ni t \mapsto \mathbf{Var}_\mu(P_t f) \in [0,\infty),$$

then it is easy to see that this is non-increasing and under some further assumptions one can even show that it is strictly decreasing for all non-constant observables f. This whole viewpoint is purely based on functional analytic arguments and one does not even need to speak about the underlying Markov process itself to carry out the corresponding calculations. Looking at this result from a different perspective, we observe that *on average* the process $(P_t f(X_t)^2)_{t\geq 0}$ is non-increasing. However, the purely analytic approach does not provide us with any insight on the behaviour of this process on the level of single trajectories. Of course, in general we cannot expect that every trajectory of this process is non-increasing, but we can hope that the process exhibits some stochastic form of monotonicity, such as being non-increasing in *conditional mean*, i.e., a supermartingale. From a probabilistic standpoint, these limitations of the purely analytic tools are somewhat dissatisfying. Consequently, we seek to enhance this coarse approach by applying a more detailed, probabilistic technique that allows us to extend these results to a trajectorial level, by which we mean results on the behaviour of *single realisations* of a stochastic process. By doing so, we uncover more of the underlying probabilistic mechanisms governing the decay of variance, or more generally, the decay of Φ-entropies. For this, we will first briefly recall the notion of Φ-entropies and then explain our main results and ideas with the help of the simple example of a continuous-time Markov chain on a finite state space. The rest of the article is then devoted to extending these ideas to the setting of spatially extended systems of infinitely many interacting particles as e.g. considered in [13].

1.1 Φ-Entropies and Their Decay Under Markovian Dynamics

Let $\Phi : I \to \mathbb{R}$ be a smooth and *convex* function defined on a not necessarily bounded interval $I \subset \mathbb{R}$. Let $(E, \mathcal{B}(E))$ be a Polish space equipped with its Borel σ-algebra and assume that μ is a probability measure on $(E, \mathcal{B}(E))$. The Φ-entropy functional is then defined on the set of μ-integrable functions $f : E \to I$ by

$$\mathbf{Ent}_\mu^\Phi(f) := \int_E \Phi(f) d\mu - \Phi\left(\int_E f d\mu\right) = \mathbb{E}_\mu\left[\Phi(f)\right] - \Phi\left(\mathbb{E}_\mu\left[f\right]\right).$$

By Jensen's inequality one can immediately deduce that the Φ-entropy functional takes its values in $\mathbb{R}_+ \cup \{+\infty\}$. Moreover, $\mathbf{Ent}_\mu^\Phi(f)$ vanishes if its argument is constant and if Φ is *strictly* convex, then the converse is also true. For special choices of Φ one can recover the classical variance and relative entropy functionals since we have

$$\mathbf{Ent}_\mu^{u\mapsto u^2} = \mathbf{Var}_\mu, \quad \mathbf{Ent}_\mu^{u\mapsto u\log u} = h(\cdot|\mu).$$

Now let $(X(t))_{t\geq 0}$ be a Markov process on our Polish space E with associated semigroup $(P_t)_{t\geq 0}$ acting on $C_b(E;\mathbb{R})$, the space of continuous and bounded real-valued functions on E. Let us assume that there exists an invariant probability measure μ and denote by $\mathscr{L}$ the generator of the semigroup $(P_t)_{t\geq 0}$ with domain $\mathrm{dom}(\mathscr{L}) \subset C_b(E;\mathbb{R})$.

By invariance of μ and Jensen's inequality one can now deduce that for all $f \in C_b(E;\mathbb{R})$

$$\mathbf{Ent}_\mu^\Phi(P_t f) = \mathbb{E}_\mu\left[\Phi(P_t f)\right] - \Phi\left(\mathbb{E}_\mu\left[P_t f\right]\right) \leq \mathbb{E}_\mu\left[P_t \Phi(f)\right] - \Phi\left(\mathbb{E}_\mu\left[f\right]\right) = \mathbf{Ent}_\mu^\Phi(f).$$

This tells us that the Φ-entropy is non-increasing as a function of t and can be used as a Lyapunov function. More precisely, with purely analytic arguments, one can even deduce the following general result about the decay of Φ-entropies.

Proposition 1.1 (DeBruijn like property for Markov semigroups, [2]) *Let $(X(t))_{t\geq 0}$ be a Markov process on a Polish space E equipped with its Borel σ-algebra $\mathcal{B}(E)$ and let $(P_t)_{t\geq 0}$ be the associated Markov semigroup with generator $\mathscr{L}$. Assume that μ is an invariant probability measure. Then, for any continuous and bounded function $f : E \to I$ and any $t > 0$, it holds that*

$$\partial_t \, \mathbf{Ent}_\mu^\Phi(P_t f) = \mathbb{E}_\mu \left[\Phi'(P_t f)\mathscr{L}(P_t f)\right] \leq 0.$$

This result is classical, but we nevertheless recall its short analytic proof.

Proof The chain rule and the definition of the generator $\mathscr{L}$ directly imply that

$$\partial_t \, \mathbf{Ent}_\mu^\Phi(P_t f) = \mathbb{E}_\mu \left[\Phi'(P_t f)\frac{d}{dt}(P_t f)\right] = \mathbb{E}_\mu \left[\Phi'(P_t f)\mathscr{L}(P_t f)\right].$$

To see that the left-hand side is actually non-positive, it suffices to observe that the convexity of Φ implies via Jensen's inequality for conditional expectations

$$\Phi(P_{t+s} g) \leq P_t(\Phi(P_s g))$$

for any $s, t \geq 0$ and hence for all f we have

$$\mathbf{Ent}_\mu^\Phi(P_{t+s} f) \leq \mathbf{Ent}_\mu^\Phi(P_s f),$$

by invariance of μ. □

By integrating with respect to the time variable one obtains the following classical corollary, which links exponential decay of Φ-entropies and functional inequalities involving Φ-entropies.

Corollary 1.2 *In the setting of Proposition 1.1, the following two statements are equivalent.*

i *There exists a constant $c > 0$ such that for all $f \in \mathrm{dom}(\mathscr{L})$*

$$\mathbf{Ent}_\mu^\Phi(f) \leq -c\mathbb{E}_\mu \left[\Phi'(f)\mathscr{L} f\right].$$

ii. *There exists a constant $c > 0$ such that for all continuous and bounded $f : E \to I$*

$$\mathbf{Ent}_\mu^\Phi(P_t f) \leq e^{-\frac{t}{c}} \mathbf{Ent}_\mu^\Phi(f).$$

Note that, in the special case $\Phi : u \mapsto u^2$, one recovers the Poincaré inequality

$$\mathbf{Var}_\mu(f) \leq -\frac{c}{2} \langle f, \mathscr{L} f\rangle_{L^2(\mu)},$$

which is well-known to be equivalent to exponential L^2 ergodicity, see e.g. [7]. For a more detailed review of Φ-entropies and further results we refer the interested reader to [2].

1.2 A Finite State-Space Example for the Trajectorial Approach

As one can see, the results above can be obtained without even mentioning the underlying stochastic process and just dealing with the semigroup and its generator. While the simplicity of this method is certainly attractive, it solely provides information about averaged quantities, i.e., the *mean* behaviour of the process, but it does not offer any insights into the behaviour of individual realisations. We therefore want to complement this perspective with a more refined and probabilistic approach that enables us to derive results at a trajectorial level.

For simplicity, we will first discuss the main ideas for the example of a continuous-time Markov chain on a finite state space. More precisely, let $(X_t)_{t\geq 0}$ be a Markov chain on a finite set E with irreducible generator $\mathscr{L}$ and strictly positive invariant measure μ. Hence, the corresponding Markov semigroup is given by the matrix exponential $(e^{t\mathscr{L}})_{t\geq 0}$. Denote the underlying probability space by $(\Omega, \mathcal{A}, \mathbb{P})$ and assume that $X_0 \sim \mu$ under $\mathbb{P}$.

It is easy to check that for all bounded $f : [0, \infty) \times E \to \mathbb{R}$ such that for all $x \in E$ the partial derivatives $\partial_t f(\cdot, x)$ are continuous and bounded, the process defined by

$$f(t, X_t) - \int_0^t (\partial_s + \mathscr{L}) f(s, X_s) ds, \quad t \geq 0, \tag{1.1}$$

is a martingale with respect to the canonical filtration generated by $(X_t)_{t\geq 0}$, see e.g. [14, Lemma IV.4.20].

If we now fix a finite time horizon $T > 0$ and consider the time-reversal $(\hat{X}_t)_{0\leq t\leq T}$ of $(X_t)_{t\geq 0}$, where $\hat{X}_t = X_{T-t}$, then under $\mathbb{P}$ the time-reversed process is again a time-homogeneous Markov process with generator $\hat{\mathscr{L}}$, where

$$\hat{\mathscr{L}}(x, y) = \frac{\mu(y)}{\mu(x)} \mathscr{L}(y, x).$$

A short calculation now shows that, for each bounded $g : E \to \mathbb{R}$ and $T > 0$, the process $(P_{T-s} g(\hat{X}_s))_{0\leq s\leq T}$ is a $((\hat{\mathcal{F}}_t)_{0\leq t\leq T}, \mathbb{P})$-martingale, where $\hat{\mathcal{F}}_t = \sigma(X_{T-s} : 0 \leq s \leq t)$. Indeed, we can use the chain rule to calculate

$$\partial_t P_{T-t} g(x) = -\hat{\mathscr{L}} P_{T-t} g(x),$$

so the correction term in (1.1) vanishes. Note that it is crucial to use the time-reversed process here, since the correction term does not cancel out if one uses the forward process.

By convexity this directly implies that the time-reversed trajectorial Φ-entropy, i.e., the process defined by

$$\Phi(P_{T-s} g(X(T-s))), \quad 0 \leq t \leq T,$$

is a submartingale. The submartingale property of this process should be thought of as a stochastic monotonicity because it tells us that almost surely we have

$$\mathbb{E}\left[\Phi(P_{T-s} g(X(T-s)))\middle|\hat{\mathcal{F}}_t\right] \geq \Phi(P_{T-t} g(X(T-t))), \quad 0 \leq t \leq s \leq T. \tag{1.2}$$

This provides us with valuable insights into the behavior of Φ-entropy functionals along individual trajectories and can be viewed as a *trajectorial refinement* of DeBruijn's Theorem, which we can recover from (1.2) by taking expectations with respect to $\mathbb{P}$. The submartingale property of the backward dynamics can be interpreted as follows: Regardless of our knowledge about the system's future trajectory, we expect the Φ-entropy to have decreased in the past. Hence, in a sense, the knowledge of a system's trajectory does not influence the DeBruijn-like decay of Φ-entropies when looking backwards in time.

Moreover, we can now apply the standard machinery of martingale inequalities to get concentration bounds on the fluctuations around the mean. For example, Doob's classical submartingale inequality, see [14, Theorem II.52.1], implies that for all $C > 0$ we have

$$\mathbb{P}\left[\sup_{0\le t\le T} \Phi(P_t f(\eta_t)) \ge C\right] \le \frac{\int \Phi(f)d\mu}{C}.$$

Since the right-hand side does not depend on $T > 0$, we even get a tail bound for unbounded time intervals

$$\mathbb{P}\left[\sup_{0\le t<\infty} \Phi(P_t f(\eta_t)) \ge C\right] \le \frac{\int \Phi(f)d\mu}{C}.$$

An analogous time-reversed martingale structure has been identified in the field of stochastic thermodynamics, specifically in the investigation of the decay of non-equilibrium free energies under Markovian dynamics, see [15, Sect. 9.1.4]. Exploring the underlying reasons for why the submartingale property appears exclusively in reversed time is a subject of ongoing research. For readers interested in delving deeper into this topic from a physics perspective, we recommend the comprehensive review article [15].

The main work is now to establish that a similar argument as in the case of a finite state space can also be made rigorous to treat infinite-dimensional systems like the interacting particle systems we consider. To our best knowledge, the first results of this kind, in the context of diffusions in $\mathbb{R}^n$, have been achieved in [4]. More recently, starting with [9], these results have been extended to more and more classes of Markov processes, including continuous time Markov chains on countable state spaces, see [10]. The works [11, 16] are also in a similar spirit.

The setting will be made precise in Sect. 2, but roughly speaking, we consider continuous-time Markov jump processes on general configuration spaces $\Omega = \Omega_0^S$, where S is an arbitrary countable set and Ω_0 is a compact Polish space. We will refer to the elements of S as *sites* and call Ω_0 the *local state-space*. In most examples considered in the literature, S is the vertex set of some graph like the d-dimensional hypercubic lattice $\mathbb{Z}^d$, a tree or the Cayley graph of a group. This underlying spatial geometry dictates which particles can interact with each other and we are therefore not in the setting of mean-field systems but in an infinite-dimensional setting. This of course brings with it its own set of technical difficulties which need to be dealt with for making the time-reversal arguments work.

The main technical difficulties come from making sure that the time-reversal is again a well-defined interacting particle system and from obtaining a description of its generator. This is made possible by assuming some local regularity of the local conditional distributions of the time-stationary measure μ. Namely, by the assumption that μ is actually a Gibbs measure with respect to a quasilocal specification that additionaly satisfies a certain smoothness condition. This condition is e.g. satisfied if the specification is given in terms of a potential $\Phi = (\Phi_B)_{B\Subset S}$ such that

$$\sup_{x\in S} \sum_{B\Subset S:\ B\ni x} |B|\, \|\Phi_B\|_\infty < \infty,$$

where the notation $B \Subset S$ means that B is a *finite* subset of S. Note that this condition is for example satisfied for any translation-invariant finite-range potential, so our theory applies to a fairly large class of models.

1.3 Organisation of the Manuscript

The rest of this article is organised as follows. We will first collect the necessary notation and formulate our main results in Sect. 2. Then, as a first step, we investigate the time-reversal of interacting particle systems in equilibrium in Sect. 3 with the main goal of obtaining an explicit representation of the (formal) generator of the time-reversed dynamics. In Sect. 4, we will then apply these results to establish pathwise properties of general Φ-entropy functionals.

2 Setting and Main Results

Let $(\Omega_0, \mathcal{B}_0)$ be a compact Polish space equipped with its Borel σ-algebra and λ_0 a probability measure on $(\Omega_0, \mathcal{B}_0)$, which will serve as our reference measure. We will consider Markovian dynamics on the configuration space $\Omega = \Omega_0^S$, where S is some countable set whose elements we will refer to as *sites*. In most applications this will be the set of vertices of some graph, e.g. $\mathbb{Z}^d$ or a tree. We equip Ω with the product topology and corresponding Borel σ-algebra $\mathcal{F}$. Note that $\mathcal{F}$ coincides with the product σ-algebra $\otimes_{x\in S}\mathcal{B}_0$. For $\Delta \subset S$ we will also write $\Omega_\Delta := \Omega_0^\Delta$ for the set of partial configurations. We will also equip Ω_Δ with the product σ-algebra and the probability measure $\lambda_\Delta = \otimes_{x\in\Delta}\lambda_0$. For $\Lambda \subset S$, let $\mathcal{F}_\Lambda$ be the sub-σ-algebra of $\mathcal{F}$ that is generated by the projections $\omega \mapsto \omega_\Delta \in \Omega_\Delta$ for $\Delta \Subset \Lambda$, where we write $\Subset$ to signify that a set is a *finite* subset of another set. For $\Delta \subset S$ and (partial) configurations $\eta_{\Delta^c} \in \Omega_{\Delta^c}$ and $\xi_\Delta \in \Omega_\Delta$, we will write $\xi_\Delta\eta_{\Delta^c}$ for the configuration that is defined on all of S and agrees with η_{Δ^c} on Δ^c and with ξ_Δ on Δ. For a topological space E, we will denote its Borel σ-algebra by $\mathcal{B}(E)$ and the space of continuous real-valued functions on E by $C(E)$. The space of non-negative measures on E, or more precisely on $\mathcal{B}(E)$, will be denoted by $\mathcal{M}(E)$ and is equipped with the topology of weak-convergence. The subset of probability measures, i.e., non-negative measures with total mass equal to one, will be denoted by $\mathcal{M}_1(E)$. The total variation distance on $\mathcal{M}(E)$ will be denoted by $\|\cdot\|_{\mathrm{TV}}$.

2.1 Interacting Particle Systems and Gibbs Measures

2.1.1 Interacting Particle Systems

We will consider time-continuous Markovian dynamics on Ω, namely interacting particle systems characterised by time-homogeneous generators $\mathscr{L}$ with domain $\mathrm{dom}(\mathscr{L}) \subset C(\Omega)$ and the associated Markovian semigroup $(P_t)_{t\geq 0}$ on $C(\Omega)$. For interacting particle systems we adopt the notation and exposition of the standard reference [13, Chapter I].

In our setting, the generator $\mathscr{L}$ is given by a collection of transition measures $(c_\Delta(\cdot, d\xi))_{\Delta\Subset S}$ in finite volumes $\Delta \Subset S$, i.e., mappings

$$\Omega \ni \eta \mapsto c_\Delta(\eta, d\xi_\Delta) \in \mathcal{M}(\Omega_\Delta).$$

These transition measures can be interpreted as the infinitesimal rates at which the particles inside Δ switch from the configuration η_Δ to ξ_Δ, given that the rest of the system is currently in state η_{Δ^c}. The full dynamics of the interacting particle system is then given as the superposition of these local dynamics,

$$\mathscr{L} f(\eta) = \sum_{\Delta\Subset S} \int_{\Omega_\Delta} [f(\xi_\Delta\eta_{\Delta^c}) - f(\eta)]\, c_\Delta(\eta, d\xi_\Delta). \tag{2.1}$$

In [13, Chapter I] it is shown that the following conditions are sufficient to guarantee well-definedness.

(L1) For each $\Delta \Subset \Omega$ the mapping

$$\Omega \ni \eta \mapsto c_\Delta(\eta, d\xi_\Delta) \in \mathcal{M}(\Omega_\Delta)$$

is continuous.

(L2) The total rate at which a single particle switches its state is uniformly bounded, i.e.,

$$\sup_{x\in S} \sum_{\Delta \ni x} \sup_{\eta\in\Omega} c_\Delta(\eta, \Omega_\Delta) < \infty.$$

(L3) The total influence of all other particles on the transition rates of a single particle is uniformly bounded, i.e.,

$$M := \sup_{x\in S} \sum_{\Delta\ni x} \sum_{y\neq x} \delta_y c_\Delta < \infty,$$

where

$$\delta_y c_\Delta := \sup\left\{\|c_\Delta(\eta,\cdot) - c_\Delta(\xi,\cdot)\|_{\mathrm{TV}} : \ \eta_{y^c} = \xi_{y^c}\right\}.$$

Under these conditions, the core of the operator $\mathscr{L}$ is given by

$$D(\Omega) := \left\{ f \in C(\Omega) : \ |||f||| := \sum_{x\in S} \delta_x(f) < \infty \right\},$$

where for $x \in S$

$$\delta_x(f) := \sup_{\eta,\xi:\ \eta_{x^c}=\xi_{x^c}} |f(\eta) - f(\xi)|$$

is the oscillation of a function $f : \Omega \to \mathbb{R}$ at the site x. In addition, one can show the following estimates for $\mathscr{L}$ and the action of the semigroup $(P_t)_{t\geq 0}$ generated by $\mathscr{L}$. We will need these later on.

Lemma 2.1 *Assume that the generator $\mathscr{L}$ satisfies* **(L1)** $-$ **(L3)** *and denote by $(P_t)_{t\geq 0}$ the semigroup generated by $\mathscr{L}$.*

i. *For $f \in D(\Omega)$ we have $P_t f \in D(\Omega)$ for all $t \geq 0$ and*

$$|||P_t f||| \leq \exp\left((M-\varepsilon)t\right) |||f|||.$$

ii. *For all $f \in D(\Omega)$ it holds that*

$$\|\mathscr{L} f\|_\infty \leq \left(\sup_{x\in S} \sum_{\Delta\ni x} \sup_{\eta\in\Omega} c_\Delta(\eta, \Omega_\Delta)\right) |||f|||.$$

The constants are explicitly given by

$$M = \sup_{x\in S} \sum_{\Delta\ni x} \sum_{y\neq x} \delta_y c_\Delta < \infty,$$

$$\varepsilon = \inf_{x\in S} \inf_{\eta,\zeta:\ \eta_{x^c}=\zeta_{x^c}, \eta_x\neq\zeta_x} \sum_{\Delta\ni x} \left(\sum_{\xi_\Delta:\xi_x=\zeta_x} c_\Delta(\eta,\xi_\Delta) + \sum_{\xi_\Delta:\xi_x=\eta_x} c_\Delta(\zeta,\xi_\Delta) \right).$$

Proof Combine the results from Proposition 3.2(a) and Theorem 3.9.(d) in [13, Chapter I]. □

For our purposes, the mere well-definedness of an interacting particle system is not sufficient and we need to assume some more regularity. All the additional assumptions we put will be used to make the generator of the time-reversal well-defined.

(R1) For each Δ and $\eta \in \Omega$ the measure $c_\Delta(\eta, d\xi_\Delta)$ is absolutely continuous with respect to the reference measure $\lambda_\Delta(d\xi_\Delta)$ with density $c_\Delta(\eta, \cdot)$.

(R3) For each $\Delta \in \Omega$ the map

$$\Omega \times \Omega_\Delta \ni (\eta, \xi_\Delta) \mapsto c_\Delta(\eta, \xi_\Delta) \in \mathbb{R}$$

is continuous with respect to the product topology.

(R3) The total rate of transition for a single site is uniformly bounded from above

$$\sup_{x \in S} \sum_{\Delta \ni x} \sup_{\eta \in \Omega} \|c_\Delta(\eta, \cdot)\|_\infty < \infty.$$

(R4) The condition (**L3**) is satisfied, i.e.,

$$\sup_{x \in S} \sum_{\Delta \ni x} \sum_{y \neq x} \delta_y c_\Delta < \infty.$$

(R5) There exists an $R > 0$ such that for all $\Delta \Subset \mathbb{Z}^d$ with $|\Delta| > R$ we have

$$\sup_{\eta \in \Omega, \xi_\Delta \in \Omega_\Delta} c_\Delta(\eta, \xi_\Delta) = 0.$$

We will comment on where and why we need these assumptions and their connection to the classical conditions (**L1**)–(**L3**) at the end of Sect. 2.1.2, after we have stated our assumptions on the local conditional distribution of the time-stationary measure μ.

2.1.2 Gibbs Measures and the DLR Formalism

We will mainly be interested in the situation where the process generated by $\mathscr{L}$ admits a time-stationary measure μ with a well-behaved local representation, namely that μ is a Gibbs measure with respect to to a sufficiently nice specification γ. Let us therefore first recall the general definition of a specification.

Definition 2.2 A specification $\gamma = (\gamma_\Lambda)_{\Lambda \Subset S}$ is a family of probability kernels γ_Λ from Ω_{Λ^c} to $\mathcal{M}_1(\Omega)$ that additionally satisfies the following properties.

i. Each γ_Λ is *proper*, i.e., for all $B \in \mathcal{F}_{\Lambda^c}$ it holds that

$$\gamma_\Lambda(B|\cdot) = \mathbf{1}_B(\cdot).$$

ii. The probability kernels are *consistent* in the sense that if $\Delta \subset \Lambda \Subset S$, then for all $A \in \mathcal{F}$

$$\gamma_\Lambda \gamma_\Delta(A|\cdot) = \gamma_\Lambda(A|\cdot),$$

where the concatenation of two probability kernels is defined as usual via

$$\gamma_\Lambda \gamma_\Delta(A|\eta) = \int_\Omega \gamma_\Delta(A|\omega) \gamma_\Lambda(d\omega|\eta).$$

For the existence and further properties of Gibbs measures with specification γ one needs to impose some conditions on the specification γ. One sufficient condition for the existence of a Gibbs measure for a specification γ is *quasilocality*, see e.g. [5] or [6]. For the following sections we will need to assume some more regularity for the specification γ. In particular, these assumptions will guarantee that γ is quasilocal.

(S1) For each $\Delta \Subset S$ and $\eta \in \Omega$, the probability measure $\gamma_\Delta(d\xi_\Delta|\eta)$ is absolutely continuous with respect to the reference measure $\lambda_\Delta(d\xi_\Delta)$ with density $\gamma_\Delta(\cdot|\eta)$.

(S2) For all $\Delta \Subset S$, the map

$$\Omega \times \Omega_\Delta \ni (\eta, \xi_\Delta) \mapsto \gamma_\Delta(\xi_\Delta|\eta_{\Delta^c}) \in [0, \infty)$$

is continuous (with respect to the product topology).

(S3) The conditional densities on the single spin spaces are uniformly bounded away from zero and infinity, i.e.,

$$0 < \delta \leq \inf_{x\in S} \inf_{\eta\in\Omega} \gamma_x(\eta_x|\eta_{x^c}) \leq \sup_{x\in S} \sup_{\eta\in\Omega} \gamma_x(\eta_x|\eta_{x^c}) \leq \delta^{-1} < \infty.$$

(S4) We have

$$\sup_{x\in S} \sum_{\Delta\ni x:\ c_\Delta>0} \sum_{y\neq x} \delta_y \gamma_\Delta < \infty,$$

where

$$\delta_y \gamma_\Delta = \sup\left\{ \|\gamma_\Delta(d\xi_\Delta|\eta) - \gamma_\Delta(d\xi_\Delta|\zeta)\|_{\mathrm{TV}} :\ \eta_{y^c} = \zeta_{y^c} \right\}.$$

Remark 2.3 Now that we have stated all of the conditions that we need, let us briefly comment on why and where we need them.

i. Assumption (**R3**) clearly implies (**L2**) and together with (**R4**) ensures that the interacting particle system is well-defined.
ii. Assumption (**R1**) and (**S1**) allow us to write down the local transition density of the time-reversal and (**S3**) makes sure that we are not performing a division by zero.
iii. The further regularity assumptions (**R3**), (**R5**) (**S3**), (**S4**) and the continuity assumptions (**R2**) and (**S2**) make sure that the local transition densities of the time-reversal also give rise to a well-defined interacting particle system.
iv. The quantity in (**S4**) is similar to the classical Dobrushin uniqueness condition, see [6]. However, we only need it to be finite and not strictly smaller than one.

Example 2.4 One particular class of models to which our theory can be applied to are spin systems for which the specification γ is defined via a potential $\Phi = (\Phi_B)_{B\Subset S}$ that satisfies

$$\sup_{x\in S} \sum_{B\ni x} |B|\, \|\Phi_B\|_\infty < \infty,$$

and where the rates are of the form

$$c_\Delta(\eta, \xi_\Delta) = \begin{cases} \exp\left(-\beta \sum_{B:B\cap\Delta\neq\emptyset} \Phi_B(\xi_\Delta \eta_{\Delta^c})\right), & \text{if } |\Delta| = 1, \\ 0, & \text{otherwise.} \end{cases}$$

Instead of these single-site updates one could also consider updates in larger regions with a bounded diameter. Then the rates satisfy (**R1**) − (**R5**) and the specification satisfies (**S1**) − (**S4**), as one can see by using similar arguments as in the proof of [5, Lemma 6.28]. Note that

this class of examples includes models with long-range pairwise interactions $(\Phi_{\{x,y\}})_{x,y\in\mathbb{Z}^d}$ that satisfy

$$\left\|\Phi_{\{x,y\}}\right\|_\infty \sim \|x-y\|_1^{-\alpha}$$

for some $\alpha > d$.

2.2 The Time-Reversal of an Interacting Particle System

In the notation of above, assume that $\mu \in \mathscr{G}(\gamma)$ is Gibbs measure for a quasilocal specification γ, i.e., assume that μ satisfies the DLR equations

$$\mu(f) = \mu(\gamma_\Lambda(f|\cdot))$$

for all $\Lambda \Subset S$ and bounded measurable functions f. Further assume that μ is time-stationary with respect to the Markovian dynamics with generator $\mathscr{L}$. Denote the semigroup generated by $\mathscr{L}$ by $(P_t)_{t\geq 0}$ and the corresponding process on Ω by $(\eta(t))_{t\geq 0}$. As discussed in Sect. 1.2, for each fixed $T > 0$ the process $(\eta(T-t))_{0\leq t\leq T}$ is again a time-homogeneous Markov process and under some suitable assumptions its associated semigroup has a generator $\hat{\mathscr{L}}$. But what does this generator look like? For general Markov processes it is not possible to give a closed form expression, but in our case we can use the special structure of $\mathscr{L}$ as the superposition of local dynamics in finite volumes. In each of these finite volumes, it is clear how the time-reversal with respect to μ should look and we can hope that we can again write $\hat{\mathscr{L}}$ as the superposition of finite volume processes. With this Ansatz, the probabilistic intuition dictates the educated guess

$$\hat{c}_\Delta(\eta,\xi_\Delta) = c_\Delta(\xi_\Delta\eta_{\Delta^c},\eta_\Delta)\frac{\gamma_\Delta(\xi_\Delta|\eta_{\Delta^c})}{\gamma_\Delta(\eta_\Delta|\eta_{\Delta^c})} \tag{2.2}$$

for the transition densities appearing in the generator of the time-reversed interacting particle system. However, at this stage, it is not obvious that the generator of the time-reversed system is again of the form (2.1) and has precisely these rates. For Markov processes on finite state spaces this is an easy calculation but we have to put in some more work, which will be carried out in Sect. 3. We obtain the following result that extends results from [12] to a much more general setting.

Proposition 2.5 (Time-reversal generator) *Let the rates of an interacting particle system with generator $\mathscr{L}$ satisfy* (**R1**) − (**R5**) *and assume that μ is a time-stationary measure for the corresponding Markov semigroup $(P(t))_{t\geq 0}$ on $C(\Omega)$ that is generated by $\mathscr{L}$ such that μ is a Gibbs measure with respect to a specification γ that satisfies* (**S1**) − (**S4**). *Then, the time-reversed process has a generator $\hat{\mathscr{L}}$ whose transition densities (with respect to the reference measure λ_0) are given by*

$$\hat{c}_\Delta(\eta,\xi_\Delta) = c_\Delta(\xi_\Delta\eta_{\Delta^c},\eta_\Delta)\frac{\gamma_\Delta(\xi_\Delta|\eta_{\Delta^c})}{\gamma_\Delta(\eta_\Delta|\eta_{\Delta^c})}.$$

The proof of this can be found at the end of Sect. 3.

2.3 Trajectorial Decay of Φ-Entropies

With this auxiliary result at hand, we can then obtain the following result, which describes the dissipation of general Φ-entropies on a trajectorial level. Before we state the theorem,

let us introduce some further notation to express the main equation in a cleaner way. The Bregman Φ-divergence associated with $\Phi : I \to \mathbb{R}$ is defined as

$$\mathrm{div}^{\Phi}(p|q) := \Phi(p) - \Phi(q) - (p-q)\Phi'(q), \quad p, q \in I.$$

This is precisely the difference between the value of Φ at the point p and the value of the first-order Taylor expansion of Φ around q, evaluated at p and is non-negative, since we assumed that Φ is convex. Bregman divergences are sometimes also referred to as Bregman distances, despite not being a metric since they are in general not symmetric and do not satisfy the triangle inequality. They are however still useful for applying techniques from optimisation theory in more general contexts, e.g. in statistical learning theory [1]. As an example, consider the classical entropy function $\Phi : u \mapsto u \log u$. Then the associated Bregman divergence is given by

$$\mathrm{div}^{\Phi}(p|q) = p \log p - p \log q + p - q = p \log\left(\frac{p}{q}\right) + p - q.$$

Note that we now have to be careful with the probability space and filtration we are working with, since we are talking about results on a trajectorial level.

Theorem 2.6 (Trajectorial decay of Φ-entropies) *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space on which the interacting particle system $(\eta(s))_{s\geq 0}$ is defined. Denote the generator of the interacting particle system by $\mathscr{L}$ and assume that its rates satisfy* (**R1**) − (**R5**) *and assume that under $\mathbb{P}$ we have $\eta_0 \sim \mu$, where μ is a time-stationary measure for the corresponding Markov semigroup $(P_t)_{t\geq 0}$ on $C(\Omega)$ that is generated by $\mathscr{L}$ such that μ is a Gibbs measure with respect to a specification γ that satisfies* (**S1**) − (**S4**). *Let $I \subset \mathbb{R}$ be an interval and $\Phi : I \to \mathbb{R}$ be a continuously differentiable convex function. Then, for any $f \in D(\Omega)$ such that $f(\Omega) \subset I$, and $T > 0$, the process defined by*

$$L^{\Phi,f}(s) := \Phi(P_{T-s} f(\eta_{T-s})), \quad 0 \leq s \leq T, \tag{2.3}$$

is a $((\hat{\mathcal{G}}_t)_{0\leq t\leq T}, \mathbb{P})$-submartingale, where $\hat{\mathcal{G}}_t = \sigma(\eta(T-s) : 0 \leq s \leq t)$. Its Doob–Meyer decomposition is given by

$$\begin{aligned} L^{\Phi,f}(t) &= M^{\Phi,f}(t) \\ &\quad + \int_0^t \sum_{\Delta \Subset S} \int_{\Omega_\Delta} \hat{c}_\Delta(\eta(T-s), \xi_\Delta) \mathrm{div}^{\Phi} \\ &\quad \times (P_{T-s}(f(\xi_\Delta \eta_{\Delta^c}(T-s)) | P_{T-s} f(\eta(T-s)))\lambda_\Delta(d\xi_\Delta) ds. \end{aligned} \tag{2.4}$$

The proof of this theorem can be found at the end of Sect. 4. While one can recover the classical DeBruijn-like decay of Φ-entropies as stated in Proposition 1.1 by taking expectations, let us stress that Theorem 2.6 gives us a much more precise description of the process. Indeed, the submartingale property provides us with information on the behaviour of the process in *conditional* mean, and not just on the level of ensemble averages. So we can actually say that, irrespective of the current state and future trajectory of our system, we expect that the Φ-entropy process has decreased in the past. Furthermore, the submartingale property allows us to make use of the strong machinery of (sub)martingale inequalities and thereby derive new results on the pathwise deviations from the ensemble average. Indeed, from Doob's submartingale inequality, see [14, Theorem II.52.1], we immediately get the

following explicit concentration bound for the trajectories:

$$\forall C > 0: \quad \mathbb{P}\left[\sup_{0\le t<\infty} L^{\Phi,f}(t) \ge C\right] \le \frac{\int_\Omega \Phi(f)d\mu}{C}.$$

This allows us to bound the probability of ever seeing large pathwise deviations from the ensemble average, as described by DeBruijn's Theorem.

For the sake of concreteness, let us write out the result from Theorem 2.6 explicitly for one of the simplest cases, namely the trajectorial decay of variance, corresponding to $\Phi : u \mapsto u^2$.

Corollary 2.7 (Trajectorial decay of variance) *In the setting of Theorem 2.6, we have that, for any $f \in D(\Omega)$ and $T > 0$, the process defined by $(P_{T-s} f(\eta_{T-s}))^2$, $0 \le s \le T$, is a $((\hat{\mathcal{G}}_t)_{0\le t\le T}, \mathbb{P})$-submartingale, where $\hat{\mathcal{G}}_t = \sigma(\eta(T-s) : 0 \le s \le t)$. Its Doob–Meyer decomposition is given by*

$$(P_{T-s} f(\eta_{T-s}))^2 = M^f(t) + \int_0^t \sum_{\Delta \Subset S} \int_{\Omega_\Delta} \hat{c}_\Delta(\eta(T-s), \xi_\Delta) \times [f(\xi_\Delta \eta_{\Delta^c}(T-s)) - f(\eta(T-s))]^2 \lambda_\Delta(d\xi_\Delta) ds.$$

2.4 Outlook

Even though we were able to show the trajectorial decay for the relative entropy under quite general assumptions on the dynamics, these results are not fully satisfactory in the context of statistical mechanics. The usually more interesting Lyapunov functional in this setting is the so-called *relative entropy density*, as e.g. considered in [5], which is not only defined for measures ν that are absolutely continuous with respect to μ. Therefore, it would be much more natural to work with this functional $h(\cdot|\mu) : \mathcal{M}_1^{inv}(\Omega) \to \mathbb{R}$ instead and one can show that is also a Lyapunov function for interacting particle systems under quite general assumptions, see [8], but it is somewhat unclear how to even formulate conjectures about the trajectorial properties of this functional, since one cannot naively evaluate it pointwise.

As we already saw in the case of a continuous time Markov chain on a finite state space, the main ingredient for this type of result is to obtain an explicit description of the generator of the time-reversed process. Another class of processes that could be of interest and is not covered by our results are systems which evolve continuously on their single spin spaces, as opposed to our pure-jump processes. The first example that comes to mind are of course systems of (infinitely-many) interacting diffusions, e.g. indexed by $\mathbb{Z}^d$. We expect that, if a given system of interacting diffusions admits a Gibbs measure as an invariant probability measure, then a combination of the techniques in [9] and this article should yield analogous results—of course under some suitable regularity conditions on the coefficients.

3 The Time-Reversed Interacting Particle System and Its Generator

The main goal of this section is to prove Proposition 2.5, thereby establishing that the generator of the time-reversal is indeed given by $\hat{\mathscr{L}}$. For this we will need to establish some regularity properties for the transition densities as defined in (2.2).

3.1 Upper and Lower Bounds for the Conditional Densities

Since we will need to deal with quotients involving the conditional densities γ_Δ on arbitrary finite subsets $\Delta \Subset S$, we will need to lift the upper and lower bounds from (**S3**) to this more general case. This is essentially the content of the following lemma.

Lemma 3.1 *Let γ be a specification that satisfies* (**S1**) *and* (**S3**). *Then, there exists a constant $C \in (0, \infty)$ such that for all $\Delta \Subset S$ we have the estimate*

$$e^{-C|\Delta|} \leq \inf_{\eta\in\Omega} \gamma_\Delta(\eta_\Delta|\eta_{\Delta^c}) \leq \sup_{\eta\in\Omega} \gamma_\Delta(\eta_\Delta|\eta_{\Delta^c}) \leq e^{C|\Delta|}.$$

This constant is precisely given by $C = |\log \delta|$.

Proof For this, fix an enumeration $i_1, \ldots, i_k$ of the elements of Δ and introduce the notation

$$[i_j, i_k] := \left\{i_j, i_{j+1}, \ldots, i_k\right\}, \quad 1 \leq j \leq k.$$

With this notation at hand, we can use the chain rule for conditional probability densities to write

$$\gamma_\Delta(\eta_\Delta|\eta_{\Delta^c}) = \prod_{j=1}^{k} \gamma_{[i_1,i_j]}(\eta_{i_j}|\eta_{[i_{j+1},i_k]}\eta_{\Delta^c}), \tag{3.1}$$

where $\gamma_{[i_1,i_j]}(\eta_{i_j}|\eta_{[i_{j+1},i_k]}\eta_{\Delta^c})$ is the marginal density of the measure $\gamma_{[i_1,i_j]}(d\eta_{[i_1,i_j]}|\eta_{[i_{j+1},i_k]}\eta_{\Delta^c})$ with respect to the site i_j. But, using consistency of the specification γ, we have

$$\gamma_{[i_1,i_j]}(\eta_{i_j}|\eta_{[i_{j+1},i_k]}\eta_{\Delta^c}) = \int \gamma_{[i_1,i_j]}(d\xi_{[i_1,i_j]}|\eta_{[i_{j+1},i_k]}\eta_{\Delta^c})\gamma_{i_j}(\eta_{i_j}|\xi_{[i_1,i_{j-1}]}\eta_{[i_{j+1},i_k]}\eta_{\Delta^c}),$$

which is, by assumption, upper bounded by δ^{-1} and lower bounded by δ. In conjunction with the representation (3.1) this implies the desired upper and lower bound where the constant C is explicitly given by $C = |\log(\delta)|$. □

As a corollary we now get the following estimate for the quotients that appear in the definition of transition density of the time-reversal (2.2).

Lemma 3.2 *Let $\Delta \Subset S$ and γ be a specification that satisfies* (**S1**) *and* (**S3**). *Then, for all $\Delta \Subset S$, $\eta \in \Omega$ and $\xi_\Delta \in \Omega_\Delta$, we have*

$$0 < e^{-2C|\Delta|} \leq \frac{\gamma_\Delta(\xi_\Delta|\eta_{\Delta^c})}{\gamma_\Delta(\eta_\Delta|\eta_{\Delta^c})} \leq e^{2C|\Delta|}.$$

3.2 The Switching Lemma

Now that we can be sure that the densities as in (2.2) are actually well-defined and we are not performing a division by zero, we can start showing that $\hat{\mathscr{L}}$ is indeed the generator of the time-reversed process. The main technical tool will be the following lemma.

Lemma 3.3 *Let the rates of a well-defined interacting particle system with generator $\mathscr{L}$ satisfy* (**R1**) *and assume that μ is a time-stationary measure for $\mathscr{L}$ and μ is a Gibbs measure*

with respect to a specification γ that satisfies (**S1**) *and* (**S3**). *Then, we have for all bounded and measurable $f, g : \Omega \to \mathbb{R}$ and $\Delta \Subset S$ that*

$$\begin{aligned}&\int_{\Omega_\Delta}\int_{\Omega} c_\Delta(\omega,\xi_\Delta) f(\omega) g(\xi_\Delta\omega_{\Delta^c})\mu(d\omega)\lambda_\Delta(d\xi_\Delta)\\ &\quad=\int_{\Omega_\Delta}\int_{\Omega} \hat{c}_\Delta(\omega,\xi_\Delta) f(\xi_\Delta\omega_{\Delta^c}) g(\omega)\mu(d\omega)\lambda_\Delta(d\xi_\Delta),\end{aligned} \tag{3.2}$$

where

$$\hat{c}_\Delta(\eta,\xi_\Delta) = c_\Delta(\xi_\Delta\eta_{\Delta^c},\eta_\Delta)\frac{\gamma_\Delta(\xi_\Delta|\eta_{\Delta^c})}{\gamma_\Delta(\eta_\Delta|\eta_{\Delta^c})}.$$

To keep the notation for conditional expectations in the upcoming proof simple, we will denote integration with respect to μ by $\mathbb{E}[\cdot]$.

Proof As a first step, note that for fixed $\Delta \Subset S$ and $\xi_\Delta \in \Omega_\Delta$ the maps

$$\Omega \ni \omega \mapsto g(\xi_\Delta\omega_{\Delta^c}) \in \mathbb{R}, \quad \Omega \ni \omega \mapsto f(\xi_\Delta\omega_{\Delta^c}) \in \mathbb{R},$$

are $\mathcal{F}_{\Delta^c}$-measurable. Therefore, we can use that γ is the local conditional distribution of μ and the definition of the rates $\hat{c}$ to obtain the μ-almost-sure identity

$$\begin{aligned}&\mathbb{E}\left[c_\Delta(\cdot,\xi_\Delta) f(\cdot) g(\xi_\Delta\cdot_{\Delta^c})|\mathcal{F}_{\Delta^c}\right](\omega)\\ &\quad= g(\xi_\Delta\omega_{\Delta^c})\mathbb{E}\left[c_\Delta(\cdot,\xi_\Delta) f(\cdot)|\mathcal{F}_{\Delta^c}\right](\omega)\\ &\quad= g(\xi_\Delta\omega_{\Delta^c})\int_{\Omega_\Delta}\gamma_\Delta(\zeta_\Delta|\omega_{\Delta^c}) c_\Delta(\zeta_\Delta\omega_{\Delta^c},\xi_\Delta) f(\zeta_\Delta\omega_{\Delta^c})\lambda_\Delta(d\zeta_\Delta)\\ &\quad= g(\xi_\Delta\omega_{\Delta^c})\int_{\Omega_\Delta}\gamma_\Delta(\xi_\Delta|\omega_{\Delta^c})\hat{c}_\Delta(\xi_\Delta\omega_{\Delta^c},\zeta_\Delta) f(\zeta_\Delta\omega_{\Delta^c})\lambda_\Delta(d\zeta_\Delta).\end{aligned}$$

If we now integrate over ξ_Δ, exchange the order of integration (via Fubini) and apply the same arguments in reverse— with f taking the role of g and vice versa—we get

$$\begin{aligned}&\int_{\Omega_\Delta}\mathbb{E}\left[c_\Delta(\cdot,\xi_\Delta) f(\cdot) g(\xi_\Delta\cdot_{\Delta^c})|\mathcal{F}_{\Delta^c}\right](\eta)\lambda_\Delta(d\xi_\Delta)\\ &\quad=\int_{\Omega_\Delta}\mathbb{E}\left[\hat{c}_\Delta(\cdot,\zeta_\Delta) f(\zeta_\Delta\cdot_{\Delta^c}) g(\cdot)|\mathcal{F}_{\Delta^c}\right](\eta)\lambda_\Delta(d\zeta_\Delta).\end{aligned}$$

By integrating both sides with respect to μ, exchanging the order of integration, and applying the law of total expectation we obtain

$$\begin{aligned}&\int_{\Omega_\Delta}\int_{\Omega_\Delta}(\omega,\xi_\Delta) f(\omega) g(\xi_\Delta\omega_{\Delta^c})\mu(d\omega)\lambda_\Delta(d\xi_\Delta)\\ &\quad=\int_{\Omega_\Delta}\int_{\Omega}\hat{c}_\Delta(\omega,\zeta_\Delta) f(\zeta_\Delta\omega_{\Delta^c}) g(\omega)\mu(d\omega)\lambda_\Delta(d\zeta_\Delta),\end{aligned}$$

as desired. □

3.3 Regularity of the Time-Reversal Rates

To make sure that $\hat{\mathscr{L}}$ is actually the generator of a well-defined interacting particle system we now show that the collection of transition measures $(\hat{c}_\Delta(\cdot,\cdot))_{\Delta\Subset S}$ satisfies the three conditions (**L1**)–(**L3**).

Proposition 3.4 *Let the rates of an interacting particle system with generator $\mathscr{L}$ satisfy* (**R1**) − (**R5**) *and assume that μ is a time-stationary measure for $\mathscr{L}$ and such that μ is a Gibbs measure with respect to a specification γ that satisfies (**S1**)–(**S4**). Then, the transition measures $(\hat{c}_\Delta(\cdot, d\xi_\Delta))_{\Delta \Subset \mathbb{Z}^d}$ with λ_Δ-densities given by*

$$\hat{c}_\Delta(\eta, \xi_\Delta) = c_\Delta(\xi_\Delta \eta_{\Delta^c}, \eta_\Delta) \frac{\gamma_\Delta(\xi_\Delta|\eta_{\Delta^c})}{\gamma_\Delta(\eta_\Delta|\eta_{\Delta^c})}$$

*satisfy the conditions (**L1**)–(**L3**).*

Proof *Ad* (**L1**): This follows from the continuity assumptions (**R2**) and (**S2**), together with assumption (**S3**) and Lemma 3.2.
Ad (**L2**): Note that for fixed $\Delta \Subset S$, $\xi_\Delta \in \Omega_\Delta$ and $\eta \in \Omega$ we have by Lemma 3.2 and assumption (**R5**)

$$|c_\Delta(\eta, \xi_\Delta)| = \left| c_\Delta(\xi_\Delta \eta_{\Delta^c}, \eta_\Delta) \frac{\gamma_\Delta(\xi_\Delta|\eta_{\Delta^c}}{\gamma_\Delta(\eta_\Delta|\eta_{\Delta^c})} \right| \leq \frac{1}{\delta} e^R c_\Delta(\xi_\Delta \eta_{\Delta^c}, \eta_\Delta).$$

So we get

$$\begin{aligned}\sup_{\eta \in \Omega} \hat{c}_\Delta(\eta, \Omega_\Delta) = \sup_{\eta \in \Omega} \int_{\Omega_\Delta} \hat{c}_\Delta(\eta, \xi_\Delta) \lambda_\Delta(d\xi_\Delta) &\leq \frac{1}{\delta} e^R \sup_{\eta \in \Omega} \int_{\Omega_\Delta} c_\Delta(\xi_\Delta \eta_{\Delta^c}, \eta_\Delta) \lambda_\Delta(d\xi_\Delta) \\ &\leq \frac{1}{\delta} e^R \sup_{\eta \in \Omega} \|c_\Delta(\eta, \cdot)\|_\infty .\end{aligned}$$

Therefore, assumption (**R3**) implies that (**L2**) is also satisfied.
Ad (**L3**): Fix $\Delta \Subset S$, $y \in S$ and two configurations η^1, η^2 that only disagree at y. Then, for any $\xi_\Delta \in \Omega_\Delta$ we have

$$\begin{aligned}&\left|\hat{c}_\Delta\left(\eta^1, \xi_\Delta\right) - \hat{c}_\Delta\left(\eta^2, \xi_\Delta\right)\right| \\ &= \left| c_\Delta\left(\xi_\Delta \eta^1_{\Delta^c}, \eta^1_\Delta\right) \frac{\gamma_\Delta\left(\xi_\Delta|\eta^1_{\Delta^c}\right)}{\gamma_\Delta\left(\eta^1_\Delta|\eta^1_{\Delta^c}\right)} - c_\Delta\left(\xi_\Delta \eta^2_{\Delta^c}, \eta^2_\Delta\right) \frac{\gamma_\Delta\left(\xi_\Delta|\eta^2_{\Delta^c}\right)}{\gamma_\Delta\left(\eta^2_\Delta|\eta^2_{\Delta^c}\right)} \right| \\ &\leq \left|c_\Delta\left(\xi_\Delta \eta^1_{\Delta^c}, \eta^1_\Delta\right)\right| \left| \frac{\gamma_\Delta\left(\xi_\Delta|\eta^1_{\Delta^c}\right)}{\gamma_\Delta\left(\eta^1_\Delta|\eta^1_{\Delta^c}\right)} - \frac{\gamma_\Delta\left(\xi_\Delta|\eta^2_{\Delta^c}\right)}{\gamma_\Delta\left(\eta^2_\Delta|\eta^2_{\Delta^c}\right)} \right| \\ &\quad + \left| \frac{\gamma_\Delta\left(\xi_\Delta|\eta^2_{\Delta^c}\right)}{\gamma_\Delta\left(\eta^2_\Delta|\eta^2_{\Delta^c}\right)} \right| \left|c_\Delta\left(\xi_\Delta \eta^1_{\Delta^c}, \eta^1_\Delta\right) - c_\Delta\left(\xi_\Delta \eta^2_{\Delta^c}, \eta^2_\Delta\right)\right| .\end{aligned}$$

To estimate this further, we will have to make a case distinction over whether the site y is contained in Δ or not. If y is contained in Δ, then we can naively use Lemma 3.2 and assumption (**R5**) to obtain the rough estimate

$$\left|\hat{c}_\Delta\left(\eta^1, \xi_\Delta\right) - \hat{c}_\Delta\left(\eta^2, \xi_\Delta\right)\right| \leq 4 \frac{1}{\delta} e^R \sup_{\eta \in \Omega, \xi_\Delta \in \Omega_\Delta} |c_\Delta\left(\eta, \xi_\Delta\right)| \leq \frac{4 e^R K(c)}{\delta}.$$

In the case where y is not contained in Δ, we can (and have to) be a bit more precise. Via the elementary algebraic rule

$$ac - bd = \frac{1}{2}\left[(a-b)(c+d) + (a+b)(c-d)\right],$$

and Lemma 3.2 plus assumption (**R5**) one obtains

$$\begin{aligned}
&\left|c_\Delta\left(\xi_\Delta\eta^1_{\Delta^c},\eta^1_\Delta\right)\right|\left|\frac{\gamma_\Delta\left(\xi_\Delta|\eta^1_{\Delta^c}\right)}{\gamma_\Delta\left(\eta^1_\Delta|\eta^1_{\Delta^c}\right)}-\frac{\gamma_\Delta\left(\xi_\Delta|\eta^2_{\Delta^c}\right)}{\gamma_\Delta\left(\eta^2_\Delta|\eta^2_{\Delta^c}\right)}\right|+\left|\frac{\gamma_\Delta\left(\xi_\Delta|\eta^2_{\Delta^c}\right)}{\gamma_\Delta\left(\eta^2_\Delta|\eta^2_{\Delta^c}\right)}\right|\left|c_\Delta\left(\xi_\Delta\eta^1_{\Delta^c},\eta^1_\Delta\right)-c_\Delta\left(\xi_\Delta\eta^2_{\Delta^c},\eta^2_\Delta\right)\right|\\
&=\frac{1}{2}\left|c_\Delta\left(\xi_\Delta\eta^1_{\Delta^c},\eta^1_\Delta\right)\right|\left|\frac{1}{\gamma_\Delta\left(\eta^1_\Delta|\eta^1_{\Delta^c}\right)\gamma_\Delta\left(\eta^2_\Delta|\eta^2_{\Delta^c}\right)}\right|\\
&\quad\times\left|\gamma_\Delta\left(\xi_\Delta|\eta^1_{\Delta^c}\right)-\gamma_\Delta\left(\xi_\Delta|\eta^2_{\Delta^c}\right)\right|\left|\gamma_\Delta\left(\eta^1_\Delta|\eta^1_{\Delta^c}\right)+\gamma_\Delta\left(\eta^2_\Delta|\eta^2_{\Delta^c}\right)\right|\\
&\quad+\left|\frac{\gamma_\Delta\left(\xi_\Delta|\eta^2_{\Delta^c}\right)}{\gamma_\Delta\left(\eta^2_\Delta|\eta^2_{\Delta^c}\right)}\right|\left|c_\Delta\left(\xi_\Delta\eta^1_{\Delta^c},\eta^1_\Delta\right)-c_\Delta\left(\xi_\Delta\eta^2_{\Delta^c},\eta^2_\Delta\right)\right|\\
&\le\frac{1}{2\delta^2}e^{2R}K(c)K(\gamma)\left|\gamma_\Delta\left(\xi_\Delta|\eta^1_{\Delta^c}\right)-\gamma_\Delta\left(\xi_\Delta|\eta^2_{\Delta^c}\right)\right|+\frac{1}{\delta}e^R\left|c_\Delta\left(\xi_\Delta\eta^1_{\Delta^c},\eta^1_\Delta\right)-c_\Delta\left(\xi_\Delta\eta^2_{\Delta^c},\eta^2_\Delta\right)\right|.
\end{aligned}$$

Now, by integrating this pointwise difference of the densities over ξ_Δ, we obtain via all of the other assumptions that

$$\sup_{x\in S}\sum_{\Delta\ni x}\sum_{y\neq x}\delta_y\hat{c}_\Delta<\infty.$$

But this is precisely (**L3**) and the proof is finished. □

With these two intermediate results at hand, we can now show that $\hat{\mathscr{L}}$ is indeed the generator of the time-reversal of $(\eta_t)_{t\ge0}$ (with respect to the time-stationary measure μ).

Proof of Proposition 2.5 It only remains to show that for all $f,g\in D(\Omega)$ we have

$$\int_\Omega f(\omega)\mathscr{L}g(\omega)\mu(d\omega)=\int_\Omega\left(\hat{\mathscr{L}}f\right)(\omega)g(\omega)\mu(d\omega),$$

since then the claimed time-reversal duality follows from Lemma A.4.

For this, we first note that it suffices to show that the duality relation for the generators holds for all local functions $f,g:\Omega\to\mathbb{R}$. Indeed, if it holds for all pairs of local functions, we can then extend it to functions with bounded total oscillation by using the estimates from Lemma 2.1 and dominated convergence. Therefore, let f,g be two local functions and let $\Lambda\Subset S$ be sufficiently large such that both f and g only depend on coordinates in Λ. By first applying Lemma 3.3 and then using that μ is time-stationary with respect to the Markovian dynamics generated by $\mathscr{L}$, we see that

$$\begin{aligned}
&\int_\Omega f(\omega)\mathscr{L}g(\omega)\mu(d\omega)-\int_\Omega\left(\hat{\mathscr{L}}f(\omega)\right)g(\omega)\mu(d\omega)\\
&=\sum_{\Delta\cap\Lambda\neq\emptyset}\int_{\Omega_\Delta}\int_\Omega c_\Delta(\omega,\xi_\Delta)[f\cdot g(\xi_\Delta\omega_{\Delta^c})-f\cdot g(\omega)]\mu(d\omega)\lambda_\Delta(d\xi_\Delta)\\
&=\int_\Omega\mathscr{L}(f\cdot g)(\omega)\mu(d\omega)=0,
\end{aligned}$$

which finishes the proof. □

4 Trajectorial Decay of Φ-Entropies

In this section we use the time-reversed process and a martingale argument to prove Theorem 2.6.

4.1 The Time-Dependent Martingale Property

The main technical tool will be the following lemma which can be seen as an extension of [14, Lemma IV.20.12] to our setting.

Lemma 4.1 *Let $\mathscr{L}$ be the generator of an interacting particle system $(\eta(s))_{s\geq 0}$ such that its transition rates satisfy* **(L1)** – **(L3)** *and let μ be a time-stationary measure with respect to $\mathscr{L}$. Then, for all $f : [0,\infty) \times \Omega \to \mathbb{R}$ such that*

i. *$f(\cdot,\eta) \in C^1([0,\infty))$ for all $\eta \in \Omega$ and*
ii. *for all $T > 0$ it holds that*

$$\sup_{0\leq t\leq T} |||f(t,\cdot)||| < \infty,$$

the process defined by

$$f(t,\eta(t)) - \int_0^t (\partial_s + \mathscr{L}) f(s,\eta(s))ds$$

is a martingale with respect to the filtration $\mathcal{G}_t := \sigma(\eta(u) : 0 \leq u \leq t)$.

The proof of this lemma is not difficult but hard to find in the existing literature, we therefore give it in some detail.

Proof For functions f as above, we define

$$M(s) := f(s,\eta(s)) - \int_0^s (\partial_u + \mathscr{L}) f(u,\eta(u))du, \quad s \geq 0$$

and denote by $(P_t)_{t\geq 0}$ the Markov semigroup generated by $(\partial_s + \mathscr{L})$. Then, for $s < t$, the Markov property and the elementary identity

$$\frac{d}{dt} P_t = P_t(\partial_t + \mathscr{L}) = (\partial_t + \mathscr{L})P_t,$$

give us

$$\begin{aligned}
&\mathbb{E}\left[f(t,\eta(t)) - \int_0^t (\partial_u + \mathscr{L}) f(u,\eta(u))du \middle| \mathcal{G}_s \right] \\
&= P_{t-s} f(s,\eta(s)) - \int_0^s (\partial_u + \mathscr{L}) f(u,\eta(u))du - \int_s^t P_{u-s}(\partial_u + \mathscr{L}) f(s,\eta(s))du \\
&= P_{t-s} f(s,\eta(s)) - \int_0^s (\partial_u + \mathscr{L}) f(u,\eta(u))du - \int_s^t \frac{d}{du} P_{u-s} f(s,\eta(s))du \\
&= f(s,\eta(s)) - \int_0^s (\partial_u + \mathscr{L}) f(u,\eta(u))du.
\end{aligned}$$

This shows that the process $(M(s))_{s\geq 0}$ is indeed a martingale. □

This abstract tool now lets us establish the analogue of the first step in the case of a finite state space considered in Sect. 1.2.

Proposition 4.2 *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space on which the interacting particle system $(\eta(s))_{s\geq 0}$ is defined. Denote the generator of $(\eta(s))_{s\geq 0}$ by $\mathscr{L}$ and assume that the rates satisfy* ***(R1)–(R5)*** *and assume that under $\mathbb{P}$ we have $\eta(0) \sim \mu$, where μ is a time-stationary measure for the corresponding Markov semigroup $(P_t)_{t\geq 0}$ on $C(\Omega)$ that is generated by $\mathscr{L}$ and that*

*μ is a Gibbs measure with respect to a specification γ that satisfies (**S1**)–(**S4**). Then, for all $f \in D(\Omega)$ and $T > 0$, the process defined by*

$$P_{T-s} f(\eta(T-s)), \quad 0 \le s \le T,$$

is a $((\hat{\mathcal{G}}_t)_{0 \le t \le T}, \mathbb{P})$-martingale, where $\hat{\mathcal{G}}_t = \sigma(\eta(T-s) : 0 \le s \le t)$.

Proof Note that by Lemma 2.1 we can apply Lemma 4.1 to the function

$$[0, T] \times \Omega \ni (s, \eta) \mapsto P_{T-s} f(\eta).$$

But since we have by the chain rule

$$\partial_s P_{T-s} f = -\hat{\mathscr{L}} P_{T-s} f,$$

the correction term cancels out and we obtain the claimed martingale property. □

4.2 Trajectorial Decay of Φ-Entropies

With this preliminary result in place, we can now come to the proof of our main result.

Proof of Theorem 2.6 *Submartingale property:* By Jensen's inequality and Proposition 4.2 we immediately see that the process $(L^{\Phi,f}(t))_{t \ge 0}$, as defined in (2.3), is a submartingale. *Doob–Meyer decomposition:* Here we want to apply Lemma 4.1 to the function

$$g : [0, T] \times \Omega \ni (s, \eta) \mapsto \Phi(P_{T-s} f) \in \mathbb{R}.$$

Via the chain rule we see that

$$\partial_s g(s, \cdot) = \partial_s \Phi(P_{T-s} f(\cdot)) = -\Phi'(P_{T-s} f(\cdot)) \hat{\mathscr{L}} P_{T-s} f(\cdot).$$

Applying the generator $\hat{\mathscr{L}}$ for fixed $s \in [0, T]$ yields

$$\hat{\mathscr{L}} g(s, \eta) = \sum_{\Delta \Subset S} \int_{\Omega_\Delta} \hat{c}_\Delta(\eta, \xi_\Delta) \left[\Phi(P_{T-s} f(\xi_\Delta \eta_{\Delta^c})) - \Phi(P_{T-s} f(\eta)) \right] \lambda_\Delta(d\xi_\Delta).$$

By putting these two ingredients together and using the previously introduced notation for the Bregman Φ-divergence we obtain

$$(\partial_s + \hat{\mathscr{L}}) g(s, \eta) = \sum_{\Delta \Subset S} \int_{\Omega_\Delta} \hat{c}_\Delta(\eta, \xi_\Delta) \mathrm{div}^\Phi(P_{T-s} f(\xi_\Delta \eta_{\Delta^c}) | P_{T-s} f(\eta)) \lambda_\Delta(d\xi_\Delta).$$

The claimed Doob–Meyer decomposition (2.4) of the submartingale $L^{\Phi,f}$ now follows from Lemma 4.1. □

Acknowledgements The authors acknowledge the financial support of the Leibniz Association within the Leibniz Junior Research Group on *Probabilistic Methods for Dynamic Communication Networks* as part of the Leibniz Competition. Furthermore, the authors would like to thank the two anonymous referees for their constructive feedback that helped to substantially improve the manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data Availability Data sharing not applicable to this article as no data sets were generated or analysed during the current study.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Appendix A: The Time-Reversal of Markov Processes in Equilibrium

In this section, we briefly summarise some properties of the time-reversal of a Markov process with respect to a time-stationary measure. These results are classical but not particularly easy to find in the literature, at least in this formulation.

We start by making precise what we mean by time-reversal of a stochastic process. Recall that any time-stationary stochastic process $(X_t))_{t\geq 0}$ can be extended (in law) to a process $(X_t)_{-\infty<t<\infty}$ via Kolmogorov's extension theorem.

Definition A.1 Let $(X_t)_{t\geq 0}$ be a time-stationary stochastic process. We call a process $(Y_t)_{t\geq 0}$ the *time-reversal* of X if

$$\mathrm{Law}((X_{-t})_{t\geq 0}) = \mathrm{Law}((Y_t)_{t\geq 0}).$$

The intuition behind this definition is that forward in time the process Y looks like the process X run backwards. For Markov processes this notion can be characterised in terms of the semigroups and generators as follows.

Proposition A.2 *Let $X = (X_t)_{t\geq 0}$ and $Y = (Y_t)_{t\geq 0}$ be Markov processes on a compact topological space E with associated semigroups $(T_t)_{t\geq 0}$ and $(S_t)_{t\geq 0}$. Assume that X has a time-stationary measure ν and for all $f, g \in C(E)$ we have*

$$\int_E (T_t f) g d\nu = \int_E f(S_t g) d\nu. \tag{A.1}$$

Then ν is also time-stationary for Y and Y is the time-reversal of X (with respect to ν).

Proof By the duality relation (A.1), ν is also time-stationary for Y and we can extend it to a process $(Y_t)_{-\infty<t<\infty}$ with $Y_0 \sim \nu$. To show that Y is the time-reversal of X (with respect to ν) it suffices to show that for arbitrary $n \in \mathbb{N}$ and times $t_0 < t_2 < \cdots < t_n < t_{n+1}$ and functions $f_1, \ldots, f_n \in C(E)$ it holds that

$$\mathbb{E}_{Y_0\sim\nu}\left[f_1(Y_{t_1})\cdots f_n(Y_{t_n})\right] = \mathbb{E}_{X_0\sim\nu}\left[f_1(X_{-t_1})\cdots f_n(X_{-t_n})\right].$$

We will do this as follows. First, we introduce the notation $f_0 = f_{n+1} \equiv 1$ and define functions $g_0, \ldots, g_{n+1}$ and $h_0, \ldots, h_{n+1}$ by $g_0 = h_{n+1} \equiv 1$,

$$g_l(x) := \mathbb{E}_x\left[\prod_{k=0}^{l-1} f_k(X_{t_l - t_k})\right], \quad 1 \leq l \leq n+1$$

and

$$h_l(y) := \mathbb{E}_y\left[\prod_{k=l+1}^{n+1} f_k(Y_{t_k - t_l})\right], \quad 0 \leq l \leq n.$$

With this notation it suffices to show that the quantity

$$\alpha_l := \int_E g_l(x) f_l(x) h_l(x) \nu(dx), \quad 0 \le l \le n+1,$$

does not depend on l. Indeed, by stationarity of X and Y, this will then imply that

$$\begin{aligned}\mathbb{E}_{Y_0 \sim \nu}\left[f_1(Y_{t_1}) \cdots f_n(Y_{t_n})\right] &= \int_E h_0 d\nu = \alpha_0 \\ &= \alpha_{n+1} = \int_E g_{n+1} d\nu = \mathbb{E}_{X_0 \sim \nu}\left[f_1(X_{-t_1}) \cdots f_n(X_{-t_n})\right],\end{aligned}$$

exactly as we wanted. In order to show that α_l does not depend on l, we use the duality relation as follows. For $0 \le l \le n$ it holds that

$$\begin{aligned}\alpha_l &= \int_E g_l f_l h_l d\nu \\ &= \int_E g_l(x) f_l(x) T_{t_{l+1}-t_l}[f_{l+1} h_{l+1}](x) \nu(dx) \\ &= \int_E S_{t_{l+1}-t_l}[g_l f_l](x) f_{l+1}(x) h_{l+1}(x) \nu(dx) \\ &= \int_E g_{l+1} f_{l+1} h_{l+1} d\nu = \alpha_{l+1}.\end{aligned}$$

Therefore, α_l does not depend on l and the claim follows. □

The duality relation (A.1) for the semigroups can also be verified by checking a similiar property on the level of generators. The main technical tool for establishing this relation between the generator of a Markov process and its semigroup will unsurprisingly be the celebrated Hille–Yosida theorem which we recall here.

Theorem A.3 (Hille–Yosida) *There is a one-to-one correspondence between Markov generators on $C(E)$ and Markov semigroups on $C(E)$. This correspondence is explicitly given by*

i.

$$\begin{aligned}\mathrm{dom}(\mathscr{L}) &= \left\{ f \in C(E) : \lim_{t \downarrow 0} \frac{S_t f - f}{t} \text{ exists} \right\}, \quad \text{and} \\ \mathscr{L} f &:= \lim_{t \downarrow 0} \frac{S_t f - f}{t}, \quad f \in \mathrm{dom}(\mathscr{L}).\end{aligned}$$

ii.

$$S_t f = \lim_{n \to \infty} \left(Id - \frac{t}{n} \mathscr{L} \right)^{-n} f, \quad f \in C(E), t \ge 0.$$

Moreover,

iii. *if $f \in \mathrm{dom}(\mathscr{L})$, then $S_t f \in \mathrm{dom}(\mathscr{L})$ for all $t \ge 0$ and $\frac{d}{dt} S_t f = \mathscr{L} S_t f = S_t \mathscr{L} f$ and*

iv. *for $g \in C(E)$ and $\lambda \ge 0$, the solution to the resolvent equation $f - \lambda \mathscr{L} f = g$ is given by*

$$f = \int_0^\infty e^{-t} S_{\lambda t} g dt.$$

Proof See Theorem 1.2.6 and Theorem 4.2.2 in [3]. □

With this at hand, we can now formulate the time-reversal duality for generators.

Lemma A.4 *Let $(T_t)_{t\geq 0}$ and $(S_t)_{t\geq 0}$ be two Markov semigroups on $C(E)$ where E is a compact topological space with time-stationary measure ν. Let $(\mathscr{A}, \mathrm{dom}(\mathscr{A}))$ and $(\mathscr{B}, \mathrm{dom}(\mathscr{B}))$ be their generators. If for all $f \in \mathrm{dom}(\mathscr{A})$ and $g \in \mathrm{dom}(\mathscr{B})$ it holds that*

$$\int_E (\mathscr{A} f)\, g\, d\nu = \int_E f\, (\mathscr{B} g)\, d\nu, \tag{A.2}$$

then (A.1) holds. It suffices if (A.2) holds for f, g in a core of the respective generators.

Proof The duality relation implies that for all $f \in \mathrm{dom}(\mathscr{A})$, $g \in \mathrm{dom}(\mathscr{B})$ and $\lambda \geq 0$ it holds that

$$\int_E f(g - \lambda \mathscr{L} g)\, d\nu = \int_E g(f - \lambda \hat{\mathscr{L}} f)\, d\nu.$$

Now we can replace f by $(\mathrm{Id}-\lambda\mathscr{L})^{-1} f$ and g by $(\mathrm{Id}-\lambda\hat{\mathscr{L}})^{-1} g$ to see that for all $f, g \in C(\Omega)$

$$\int_E [(\mathrm{Id} - \lambda\mathscr{A})^{-1} f] g\, d\nu = \int_E [(\mathrm{Id} - \lambda\mathscr{B})^{-1} g] f\, d\nu.$$

By iteratively applying this equality we obtain that for all $n \in \mathbb{N}$ and $\lambda \geq 0$ it holds that

$$\int_E [(\mathrm{Id} - \lambda\mathscr{A})^{-n} f] g\, d\nu = \int_E [(\mathrm{Id} - \lambda\mathscr{B})^{-n} g] f\, d\nu.$$

For fixed $t \geq 0$ we can replace λ by t/n and use Part *ii*. of Theorem A.3 to see that

$$\lim_{n\to\infty} \int_E \left[(\mathrm{Id} - \frac{t}{n}\mathscr{A})^{-n} f \right] g\, d\nu = \int_E g T_t f\, d\nu$$

and

$$\lim_{n\to\infty} \int_E \left[(\mathrm{Id} - \frac{t}{n}\mathscr{B})^{-n} g \right] f\, d\nu = \int_E f S_t g\, d\nu,$$

as desired. □

To sum this up, in order to show that a stationary Markov process Y is the time-reversal of a stationary Markov process X, it suffices to check that their generators satisfy the duality relation (A.2).

References

1. Banerjee, A., Merugu, S., Dhillon, I.S., Ghosh, J.: Clustering with Bregman divergences. J. Mach. Learn. Res. **6**(58), 1705–1749 (2005)
2. Chafaï, D.: Entropies, convexity, and functional inequalities: On Φ-entropies and Φ-Sobolev inequalities. Kyoto J. Math. **44**(2), 325–363 (2004)
3. Ethier, S.N., Kurtz, T.G. (eds.): Markov Processes, Wiley Series in Probability and Statistics. Wiley, Hoboken (1986)
4. Fontbona, J., Jourdain, B.: A trajectorial interpretation of the dissipations of entropy and Fisher information for stochastic differential equations. Ann. Probab. **44**(1), 131–170 (2016)
5. Friedli, S., Velenik, Y.: Statistical Mechanics of Lattice Systems: A Concrete Mathematical Introduction, 1st edn. Cambridge University Press, Cambridge (2017)

6. Georgii, H.-O.: Gibbs Measures and Phase Transitions. De Gruyter Studies in Mathematics, vol. 9, 2nd edn. De Gruyter, Berlin (2011)
7. Holley, R.A., Stroock, D.W.: L^2-theory for the stochastic Ising model. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **35**(2), 87–101 (1976)
8. Jahnel, B., Köppl, J.: Dynamical Gibbs variational principles for irreversible interacting particle systems with applications to attractor properties. (2022). arXiv:2205.02738
9. Karatzas, I., Schachermayer, W., Tschiderer, B.: Trajectorial Otto calculus. (2020). arXiv:1811.08686
10. Karatzas, I., Maas, J., Schachermayer, W.: Trajectorial dissipation and gradient flow for the relative entropy in Markov chains. Commun. Inf. Syst. **21**(4), 481–536 (2021)
11. Kim, D., Yeung, L.C.: A trajectorial approach to entropy dissipation for degenerate parabolic equations. (2022). arXiv: 2210.16158
12. Künsch, H.: Non reversible stationary measures for infinite interacting particle systems. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **66**(3), 407–424 (1984)
13. Liggett, T.M.: Interacting Particle Systems, Classics in Mathematics. Springer, Berlin (2005)
14. Rogers, L.C.G., Williams, D.: Diffusions, Markov Processes, and Martingales, 2nd edn. Cambridge University Press, Cambridge (2000)
15. Roldán, É., Neri, I., Chetrite, R., Gupta, S., Pigolotti, S., Jülicher, F., Sekimoto, K.: Martingales for Physicists. (2023). arXiv:2210.09983
16. Tschiderer, B., Yeung, L.C.: A trajectorial approach to relative entropy dissipation of McKean–Vlasov diffusions: gradient flows and HWBI inequalities. Bernoulli **29**(1), 725–756 (2023)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.