

# Ergodicity and large deviations in physical systems with stochastic dynamics

Robert L. Jack<sup>1,2,a</sup>

<sup>1</sup> Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK

<sup>2</sup> Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK

Received 17 December 2019 / Received in final form 29 February 2020

Published online 22 April 2020

© The Author(s) 2020. This article is published with open access at [Springerlink.com](https://www.springerlink.com)

**Abstract.** In ergodic physical systems, time-averaged quantities converge (for large times) to their ensemble-averaged values. Large deviation theory describes rare events where these time averages differ significantly from the corresponding ensemble averages. It allows estimation of the probabilities of these events, and their mechanisms. This theory has been applied to a range of physical systems, where it has yielded new insights into entropy production, current fluctuations, metastability, transport processes, and glassy behaviour. We review some of these developments, identifying general principles. We discuss a selection of dynamical phase transitions, and we highlight some connections between large-deviation theory and optimal control theory.

## 1 Introduction

In statistical mechanics, many properties of equilibrium systems can be calculated using free-energy methods, and the underlying Boltzmann distribution. However, this approach has two important restrictions – it only applies in equilibrium, and it is restricted to static properties. For example, the Boltzmann distribution has very little to say about dynamical quantities like viscosity and thermal conductivity, nor can it predict the time required for a protein to fold. Predicting such quantities requires some knowledge of the equations of motion of a system: the relevant statistical mechanical theories must include *dynamical information*. Such theories are useful in many contexts, which include non-equilibrium steady states [1–3] as well as dynamical aspects of the equilibrium state (for example in glassy materials [4]). Other physical phenomena also involve transient relaxation to equilibrium, for example nucleation [5] and self-assembly [6].

For complex systems (with many strongly-interacting components), dynamical theories often assume that the behaviour is *ergodic*. That is, the systems have steady states in which time-averaged measurements converge (for long times) to corresponding ensemble averages. Many important physical systems have this property, which motivates several questions. For example: (i) How long does it take for the time-averaged measurements to converge? (ii) What is the probability that the time-average

does not converge to the ensemble average, given some long time  $\tau$ ?

In systems with deterministic dynamics, there is a rich and complex mathematical structure that allows such questions to be addressed, but the resulting theory has many subtle features [7–9]. Here we focus on stochastic processes, where the situation is somewhat simpler. In particular, the mathematical theory of large deviations [10] can be used to analyse time-averaged quantities, as demonstrated by important work in the late 1990s and early 2000s [11–16]. The theory has been applied to a range of physical systems, where it has provided new insights. Examples include exclusion processes [2,13–16], glassy materials [17–21], models of heat transport [22–24], proteins [25,26], climate models [27], and non-equilibrium quantum systems [28].

This article outlines the application of large deviation theory as it applies to time-averaged quantities, and it describes some of the results and insights that have been obtained for physical systems. By considering a range of applications, the aim is to complement other papers that focus primarily on the general structure of the theory [29] or on specific classes of system [2,16]. The remainder of this section lays out some general principles and describes the theoretical context in more detail. Later sections are devoted to general aspects of the theory and to application areas including phase transitions, glassy systems, entropy production, and exclusion processes. A few examples are discussed in detail. The choice of applications and examples is biased towards the author's own work; they are presented within the broader context of the field.

<sup>a</sup> e-mail: [rlj22@cam.ac.uk](mailto:rlj22@cam.ac.uk)

### 1.1 Fluctuations of time-averaged quantities

This section introduces the main question that will be considered below. Consider a system with stochastic dynamics, whose configuration at time  $t$  is  $\mathcal{C}_t$ . Define an observable quantity  $b = b(\mathcal{C})$  and a time interval  $[0, \tau]$ ; then the time-average of  $b$  is

$$\bar{b}_\tau = \frac{1}{\tau} \int_0^\tau b(\mathcal{C}_t) dt. \quad (1)$$

As a simple example one may consider an Ising model with  $N$  spins, as in [30,31]. Then  $\mathcal{C} = (\sigma_1, \sigma_2, \dots, \sigma_N)$  where each spin  $\sigma_i = \pm 1$ . Take  $b(\mathcal{C})$  to be the energy of this configuration, so  $\bar{b}_\tau$  is the time-averaged energy. Clearly  $\bar{b}_\tau$  is a random variable: different trajectories of the system have different values for this quantity. However, in ergodic systems the typical situation is that  $\bar{b}_\tau$  obeys a central limit theorem at large times: its distribution is Gaussian with a variance that decays as  $\tau^{-1}$ . Motivated by questions (i) and (ii) above, this article considers fluctuations that are not covered by the central limit theorem: large deviation theory is used to characterize *rare events* where  $\bar{b}_\tau$  differs significantly from its mean value, even as  $\tau \rightarrow \infty$ . We will see below that these are *exponentially* rare, in the sense that their log-probability is negative and proportional to  $\tau$ .

Since these events are very rare, one might wonder what relevance they have for practical physical systems. In response to this question, we make two general points, which will be clarified below. First, large-deviation theory has a rich structure and enables sharp statements about the dynamical behaviour of complex systems. As such, it can be viewed as an idealised theoretical starting point for studies of dynamical behaviour in non-equilibrium systems, which enables general insight. An important example is the analysis of fluctuation theorems [12]. Second, the theory has already proven useful for understanding the behaviour of physical systems, for example through analysis of metastable states in glassy systems [18,32] and biomolecules [25], and through uncertainty bounds on fluctuations of the current [33], which are relevant for rare events and for typical fluctuations.

### 1.2 Theoretical context

The mathematical theory for large deviations of time-averaged quantities in stochastic processes was formulated by Donsker and Varadhan in the 1970s [34–37]. A clear presentation of the general (mathematical) theory of large deviations is given in the book of den Hollander [10]. An alternative mathematical approach to these problems is discussed in the book of Dupuis and Ellis [38], including a connection to ideas of optimal control theory, as discussed below. In physical studies of non-equilibrium systems, work by Derrida and Lebowitz [13] and Lebowitz and Spohn [12] laid the foundations for the work described here, building on earlier studies [11,39,40]. As mentioned in the introduction, theories of ergodicity and time-averages in deterministic systems also have a long history [7–9,41], and large deviation theory is also relevant

in these cases [8,9,42]. This article is restricted to stochastic systems, analysis of deterministic systems requires a different set of methods and assumptions.

A separate strand of mathematical work applied large deviation theory to hydrodynamic limits ([43], Ch. 10), and underlies the macroscopic fluctuation theory of Bertini, de Sole, Gabrielli, Jona-Lasinio and Landim [2,14], which can also be used to analyse fluctuations of time-averaged quantities. Yet another direction is the connection between large deviation theory and the theory of equilibrium statistical mechanics, as discussed by Ellis [44], see also [42,45].

A useful resource from the physics literature is the review of Touchette [29] which gives a clear presentation of large-deviation theory as it applies to equilibrium statistical mechanics and to time-averaged quantities, see also [16,46]. Two recent papers by Ch  trite and Touchette [47,48] provide a comprehensive summary of the large deviation theory of time-averaged quantities, as it applies to physical systems.

### 1.3 Outline

The remainder of this article is structured as follows. Section 2 gives an overview of the large deviation theory for time-averaged quantities. It focuses on finite systems, which simplifies the analysis. Section 3 discusses some of the dynamical phase transitions that can occur in infinite systems, including an example calculation for the 1d Glauber-Ising model and a discussion of dynamical phase coexistence. In Section 4 we discuss the behaviour of glassy systems, including dynamical phase transitions in kinetically constrained models. Section 5 discusses the role of time-reversal symmetries and large deviations of the entropy production, including an example from active matter. We give a short discussion of exclusion processes and hydrodynamic behaviour in Section 6 before ending in Section 7 with an outlook and a discussion of some possible future directions.

## 2 General theory

This section outlines the general theory of large deviations of time-averaged quantities. This presentation is not at all complete, the aim is to highlight useful facts, in order to provide physical insight and intuition. Nevertheless, some mathematical precision is required, in order to understand the scope and applicability of the theory; some technical details are provided in footnotes. A more comprehensive presentation of similar material is given by Ch  trite and Touchette [47,48].

### 2.1 Definitions

The central quantities that appear in large deviation theory are probability distributions, rate functions, and cumulant generating functions. These are introduced in a general context, some of the systems to which the theory can be applied are discussed in Section 2.2. We consider models that converge at long times to unique

steady states, and angled brackets  $\langle \cdot \rangle$  indicate steady-state averages.

Recalling (1), the probability density for  $\bar{b}_\tau$  is denoted by  $p(\bar{b}|\tau)$ . The *cumulant generating function* (CGF) for  $\bar{b}_\tau$  is

$$G(s, \tau) = \log \langle e^{-s\bar{b}_\tau} \rangle. \tag{2}$$

One sees that  $G(0, \tau) = 0$  and  $(\partial G/\partial s)_{s=0} = -\langle \bar{b}_\tau \rangle$ .<sup>1</sup>

To analyse large deviations, we consider the limit of large  $\tau$ , defining

$$I(\bar{b}) = - \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \log p(\bar{b}|\tau). \tag{3}$$

As anticipated in Section 1.1, the interesting case is where this limit is finite (and non-zero), so the relevant fluctuations occur with probabilities that decay exponentially with  $\tau$ . In this case we say that  $\bar{b}_\tau$  obeys a large deviation principle (LDP) and  $I$  is called the *rate function*.<sup>2</sup> The rate function is non-negative,  $I(b) \geq 0$  for all  $b$ . In cases where an LDP holds with rate function  $I$ , we write

$$p(\bar{b}|\tau) \asymp e^{-\tau I(\bar{b})}. \tag{4}$$

The meaning of the asymptotic equality symbol  $\asymp$  is that (4) is equivalent to (3), see [50]. It is a general property of LDPs that the argument of the exponential in (4) is the product of the rate function and a large parameter that is called the *speed* of the LDP. In (3,4) the speed is  $\tau$ , which is an assumption of the theory presented so far. There are physical systems where time-averaged quantities obey LDPs with other speeds (for example [51–54]) but we focus here on LDPs with speed  $\tau$ , which is the most common situation.

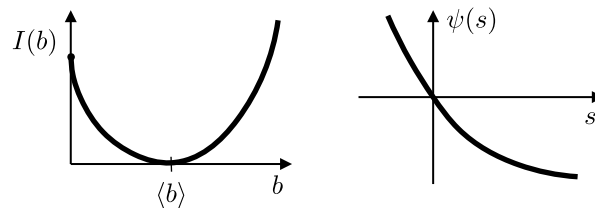
In simple cases (see Sect. 2.2 for examples), the rate function  $I$  is analytic and strictly convex, with a unique minimum at  $b = \langle b \rangle$ , and  $I(\langle b \rangle) = 0$ . In this case [10,29],  $\bar{b}_\tau$  obeys a central limit theorem (as in [55]), with a variance  $\sigma_b^2/\tau$  that is related to the curvature of the rate function as  $\sigma_b^2 = 1/I''(\langle b \rangle)$ . The next step is to define the *scaled cumulant generating function* (SCGF),

$$\psi(s) = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} G(s, \tau). \tag{5}$$

One sees from (2) that  $\psi'(0) = -\langle b \rangle$  and  $\psi''(0) = \sigma_b^2$ .

<sup>1</sup> Our definitions mostly follow [49], in particular we include a minus sign in the exponent of (2), which is natural for the thermodynamic analogy discussed in Section 2.3. However, analogous definitions of the CGF without any minus sign are also common in the literature.

<sup>2</sup> The mathematical theory of large deviations [10] expresses LDPs in a more general way that involves probabilities of events instead of probability densities, and also places some additional restrictions on rate functions. The details of the mathematical theory can be important in some physical situations, but we concentrate here on simple cases for which the presentation given here is adequate.



**Fig. 1.** Sketch of a rate function and an SCGF for a positive quantity  $\bar{b}$ . In this example, both functions are convex and related by Legendre transformation; the rate function has a single minimum at the ensemble average  $\langle b \rangle$ ; the SCGF has  $\psi(0) = 0$  and  $\psi'(0) = -\langle b \rangle$ , it is monotonic because  $\bar{b}$  is positive.

The rate function and the SCGF are related by a Legendre transformation:

$$\psi(s) = \sup_b [-sb - I(b)]. \tag{6}$$

This is (a particular case of) Varadhan’s lemma [10]. It can be motivated by writing (2) as

$$G(s, \tau) = \log \int p(\bar{b}|\tau) e^{-s\tau\bar{b}} d\bar{b}, \tag{7}$$

and substituting (4), then doing the integral by the saddle-point method. If the function  $\psi$  is analytic then one has also

$$I(b) = \sup_s [-sb - \psi(s)]. \tag{8}$$

An important result in large deviation theory is the Gärtner-Ellis theorem [10,29]: it allows large deviation results like (4) to be proved, as long as the SCGF  $\psi$  obeys certain conditions. In such cases the rate function can then be derived from (8).

A schematic illustration of two functions  $I$  and  $\psi$  related by Legendre transform are shown in Figure 1. Note, the minus sign in (2) means that the behaviour of the SCGF for  $s > 0$  is relevant for the rate function for  $b < \langle b \rangle$ , and vice versa.

## 2.2 Applicability of the theory

The theory of Section 2.1 can be applied to a wide range of models, but some assumptions are required in order to ensure that the limit in (3) is finite and non-zero. Some results including (8) also rely on analytic properties of  $\psi$ . In this article, much of our analysis is based on two main classes of system, which are Markov chains and diffusion processes. We make several assumptions, which ensure that the models are ergodic, the limit (3) is well-behaved, and the functions  $I$  and  $\psi$  are analytic and strictly convex, as discussed in [47]. These cases are useful to illustrate the theory. However, the tools of large-deviation theory are not at all restricted to these cases; this will become clear in later sections.

### 2.2.1 Finite-state Markov chains

We consider finite Markov chains, so the configurations  $\mathcal{C}$  come from a finite set. They may evolve in either continuous- or discrete-time. In continuous time, a model is defined by specifying the transition rates between the configurations, which are denoted by  $W(\mathcal{C} \rightarrow \mathcal{C}')$ . Prominent examples in this case include exclusion processes and Ising-like models on finite lattices. In this case  $\bar{b}$  may be defined as a time integral as in (1), or one may take the more general form [49]

$$\bar{b}_\tau = \frac{1}{\tau} \sum_{\text{jumps } \mathcal{C} \rightarrow \mathcal{C}'} \alpha(\mathcal{C}, \mathcal{C}') + \frac{1}{\tau} \int_0^\tau h(\mathcal{C}_t) dt \quad (9)$$

where the functions  $\alpha, h$  correspond to observable quantities similar to  $b$  in (1), and the sum runs over the transitions that take place in the trajectory. This type of time-averaged quantity is particularly useful when considering time-averaged currents: for example, if the model involves particles hopping on a  $1d$  lattice with periodic boundaries, one may take  $\alpha = 1$  for jumps to the right and  $\alpha = -1$  for jumps to the left [16], with  $h = 0$ .

If the continuous-time Markov chain is finite and irreducible and  $\alpha, h$  are finite then the limits in (3,5) certainly exist, and the functions  $I$  and  $\psi$  are analytic and strictly convex [47]. In discrete time the situation is the same, as long as the Markov chain is also aperiodic.

### 2.2.2 Diffusion processes

We also consider models defined by stochastic differential equations (or Langevin equations). In this case the configurations  $\mathcal{C}$  are vectors in  $d$ -dimensional space and they evolve by

$$d\mathcal{C}_t = v(\mathcal{C}_t)dt + \sqrt{2}\sigma(\mathcal{C}_t) \circ dW_t \quad (10)$$

where the circle indicates a Stratonovich product. Here,  $v$  is a vector-valued drift,  $\sigma$  is a matrix-valued noise strength, and  $W_t$  is a  $d$ -dimensional standard Brownian motion. For models in this class, some technical restrictions are needed on the functions  $v, \sigma$ , in order to establish existence of the limits in (3,5) and convexity of the rate function. For simplicity, we restrict to systems defined on finite domains, with periodic boundary conditions. In this case it is sufficient that  $v, \sigma$  should be finite and the matrix  $\sigma\sigma^\dagger$  should not have any zero eigenvalues. In this case  $\bar{b}$  may again be defined as in (1), or one may consider [47]

$$\bar{b}_\tau = \frac{1}{\tau} \int_0^\tau a(\mathcal{C}_t) \circ d\mathcal{C}_t + \frac{1}{\tau} \int_0^\tau h(\mathcal{C}_t) dt \quad (11)$$

where now  $a, h$  are vector-valued and scalar functions respectively. A simple case takes  $a$  to be constant and  $h = 0$  in which case  $\bar{b}_\tau$  is a time-averaged current in the direction  $a$ . (In systems with closed boundaries then such time-averaged currents must vanish as  $\tau \rightarrow \infty$  but periodic systems can support trajectories with sustained non-zero current.) The functions  $b, a, h$  are all assumed to be finite.

### 2.3 Analogy between $\tau \rightarrow \infty$ and thermodynamic limit

This article focuses on large deviations of time-averaged quantities, but there are other situations where large deviation theory is relevant in physics. The most prominent example is the theory of the thermodynamic limit [42,44]. We briefly outline the analysis of this limit within large deviation theory, which motivates an analogy between large-time limits and thermodynamic limits. A more detailed discussion of this analogy is given in the review of Touchette [29], see also [9,46,49]. The analogy is useful for two reasons. First, it provides valuable intuition about dynamical large deviations, since thermodynamic theories may be more familiar than dynamical ones. Second, it provides a route whereby established methods from thermodynamics can be generalised, in order to address dynamical problems.

Within the analogy, the CGF in (2) corresponds to a *difference in free-energy* between two states. Specifically, consider a thermodynamic system of volume  $V$ , where the energy of configuration  $\mathcal{C}$  is  $\mathcal{E}(\mathcal{C})$ . Define  $\beta = (k_B T)^{-1}$  where  $k_B$  is Boltzmann's constant, so the Boltzmann distribution of this system is  $p_\beta(\mathcal{C}) = e^{-\beta\mathcal{E}(\mathcal{C})}/Z_\beta$ , where  $Z_\beta$  is the partition function. We denote averages with respect to  $p_\beta$  by  $\langle \cdot \rangle_{\text{Boltz}}$ .

Now consider a perturbation to this system where the (extensive) energy is modified by  $\Delta\mathcal{E}(\mathcal{C}) = -hVm_V(\mathcal{C})$ . For example,  $m_V$  might be the (intensive) magnetisation of an Ising model, and  $h$  its conjugate (magnetic) field. The free energy difference between the original system and this new state is  $\Delta F$ . It satisfies

$$-\beta\Delta F(\beta, h, V) = \log \langle e^{\beta h V m_V(\mathcal{C})} \rangle_{\text{Boltz}} \quad (12)$$

which is analogous to (2) with  $(-\beta h, V, m_V) \rightarrow (s, \tau, \bar{b}_\tau)$  and  $(-\beta\Delta F) \rightarrow G$ . For the purposes of this analogy we consider  $\beta$  to be a fixed number, it is the field  $h$  that is the analogue of the parameter  $s$  from Sect. 2.1. In that section we considered the limit  $\tau \rightarrow \infty$ , here we consider  $V \rightarrow \infty$ .

Application of large-deviation theory to the fluctuations of  $m_V$  requires that this quantity is intensive, which means that it can be expressed as an average over the (large) system, analogous to the time-average in (1). Then standard thermodynamic arguments for large systems imply that  $\Delta F$  is extensive:  $\Delta F(\beta, h, V) \approx V\Delta f(\beta, h)$ , where  $\Delta f(\beta, h)$  is a difference in *free-energy density*. Comparing with (5), the dynamical SCGF  $\psi(s)$  is analogous to  $-\beta\Delta f(\beta, h)$ . Continuing the analogy shows that  $m_V$  for the unperturbed system has a probability distribution

$$p(m|V, \beta) \asymp e^{-VI(m, \beta)} \quad (13)$$

with  $I(m, \beta) = \sup_h [\beta h m + \beta\Delta f(\beta, h)]$ , similar to (4,8). Just like the dynamical case, some care is required with this analysis in cases where  $\Delta f(\beta, h)$  is not analytic. These cases correspond to thermodynamic phase transitions, for which there is a well-developed theory: see for example [44,45]. Section 3 discusses some ways that the thermodynamic theory of phase transitions can be generalised to the dynamical context.



### 2.4 Biased ensembles of trajectories (the $s$ -ensemble)

In the thermodynamic setting, it is natural to consider a family of Boltzmann distributions, parameterised by  $h$ . We now introduce corresponding distributions for trajectories, which we refer to as  $s$ -ensembles [49] or biased ensembles [47]. Let  $\mathcal{C}$  indicate a trajectory of the system of interest, where the time  $t$  runs from 0 to  $\tau$ . This trajectory has a probability density  $P_\tau(\mathcal{C})$ , which has the property that  $\langle F \rangle = \int F(\mathcal{C})P_\tau(\mathcal{C})d\mathcal{C}$ .<sup>3</sup> Note that the probability of the initial state  $\mathcal{C}_0$  is included in  $P_\tau(\mathcal{C})$ .

The probability density for trajectory  $\mathcal{C}$  in the biased ensemble is

$$P_\tau^s(\mathcal{C}) = P_\tau(\mathcal{C})e^{-s\tau\bar{b}_\tau(\mathcal{C})-G(s,\tau)} \quad (14)$$

which is normalised, by (2). The average of any trajectory-dependent observable  $F$  within this ensemble is

$$\langle F \rangle_s = \frac{\langle Fe^{-s\tau\bar{b}_\tau} \rangle}{\langle e^{-s\tau\bar{b}_\tau} \rangle}. \quad (15)$$

Note that these averages depend implicitly on the trajectory length  $\tau$ .

In the analogy with thermodynamics, (14) corresponds to a Boltzmann distribution, in the canonical ensemble. As discussed in [49], standard thermodynamic arguments for equivalence of ensembles then indicate that typical trajectories of (14) should be similar to typical trajectories from an associated microcanonical ensemble, where the value of  $\bar{b}_\tau$  is constrained to a specific value. A precise characterisation of this ensemble-equivalence is given in references [47,56].

An important observation is that the initial and final conditions of the trajectory are analogous to boundaries of thermodynamic systems, where the behaviour may differ from the bulk. Thermodynamic equivalence of ensembles applies to observable quantities that are evaluated in finite regions, within the bulk of a large system. In biased ensembles, these correspond to observables that are well-separated (in time) from the initial and final conditions at  $t = 0, \tau$ .

Bearing this mind, it is useful to consider a one-time dynamical observable  $a(\mathcal{C}_t)$ , such as the instantaneous energy of the system  $E(\mathcal{C}_t)$ . This quantity is associated with two different probability distributions, depending on the time  $t$  [30,49,57]. The bulk is characterised by a distribution which we define by evaluating the observable at a randomly-chosen time:

$$P_{\text{ave}}(a|s) = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau \langle \delta[a - a(\mathcal{C}_t)] \rangle_s dt. \quad (16)$$

<sup>3</sup> It is not trivial to define the integration measure  $d\mathcal{C}$ , but see [49] for an explicit construction for finite Markov chains. A more rigorous mathematical approach would sidestep this problem by working directly with probability measures for trajectories. The analysis of this work can be reformulated in that way: one should replace integration measures  $P_\tau(\mathcal{C})d\mathcal{C}$  by  $dP_\tau(\mathcal{C})$  and ratios of probability densities  $P(\mathcal{C})/Q(\mathcal{C})$  by Radon-Nikodym derivatives  $dP/dQ$ . All conclusions remain unchanged.

Alternatively one may evaluate the same observable at the final time  $\tau$  to obtain

$$P_{\text{end}}(a|s) = \lim_{\tau \rightarrow \infty} \langle \delta[a - a(\mathcal{C}_\tau)] \rangle_s. \quad (17)$$

The presence of boundaries means that  $P_{\text{ave}} \neq P_{\text{end}}$  in general. In the cases that we consider, the bulk of the  $s$ -ensemble is time-translation invariant (similar to homogeneity of thermodynamic systems), which means that  $P_{\text{ave}}$  can also be evaluated as  $P_{\text{ave}}(a|s) = \lim_{\tau \rightarrow \infty} \langle \delta[a - a(\mathcal{C}_{u\tau})] \rangle_s$  for any  $u$  with  $0 < u < 1$ .<sup>4</sup>

### 2.5 Formulation as eigenproblem (operator approach)

We describe two types of theoretical approach by which results for large deviations can be obtained. This section describes the first method, which is to characterise  $\psi(s)$  as the largest eigenvalue of an operator (or matrix), which is called a tilted generator or a biased master operator. In the physical context, this was the approach applied (for stochastic models) in [12,13], see also [11,39,40]. To explain it, define

$$\rho(\mathcal{C}|s, \tau) = \left\langle e^{-s\tau\bar{b}_\tau} \delta(\mathcal{C} - \mathcal{C}_\tau) \right\rangle \quad (18)$$

where the delta function restricts the average to trajectories that end in state  $\mathcal{C}$ . Comparing with (2), one sees that  $G(s, t) = \log \int \rho(\mathcal{C}|s, \tau)d\mathcal{C}$ .

The time derivative of  $\rho$  behaves as

$$\frac{\partial}{\partial \tau} \rho(\mathcal{C}|s, \tau) = \mathcal{W}_s \rho(\mathcal{C}|s, \tau) \quad (19)$$

where  $\mathcal{W}_s$  is an  $s$ -dependent linear operator.<sup>5</sup> For example, in finite-state Markov chains (with  $n$  states) then

$$\mathcal{W}_s \rho(\mathcal{C}|s, \tau) = \sum_{\mathcal{C}'} M_s(\mathcal{C}, \mathcal{C}') \rho(\mathcal{C}'|s, \tau) \quad (20)$$

where  $M_s$  is a matrix of size  $n \times n$  that depends on the transition rates of the model and on the observable  $\bar{b}_\tau$  [16, 47,49]. For diffusion processes then  $\mathcal{W}_s$  is an operator that involves first and second derivatives with respect to  $\mathcal{C}$ , an example is given in (31), below.

The large-time behavior of the solution of (19) can be deduced by considering the largest eigenvalue of  $\mathcal{W}_s$ . [In the example of (20), this is simply the largest eigenvalue of  $M_s$ .] Anticipating the answer, we assume that this largest eigenvalue is unique and we denote it by  $\psi(s)$ . The associated eigenvector (or eigenfunction) is  $P_{\text{end}}(\mathcal{C}|s)$  which we define to be normalised as a probability distribution

<sup>4</sup> The inequalities are strict so  $u = 0, 1$  are excluded, in particular taking  $u = 1$  recovers  $P_{\text{end}}$ .

<sup>5</sup> This is a tilted version of what would be called in mathematics the *forward generator*, the ‘tilting’ refers to the effect of  $s$  and setting  $s = 0$  recovers the usual (forward) generator. The adjoint of  $\mathcal{W}_s$  is the (tilted) backward generator. Mathematical analyses are typically framed in terms of the backwards generator.

$\int P_{\text{end}}(\mathcal{C}|s)d\mathcal{C} = 1$ . So the eigenproblem is

$$\mathcal{W}_s P_{\text{end}} = \psi(s) P_{\text{end}} \quad (21)$$

and the solution of (18) is

$$\rho(\mathcal{C}|s, \tau) = e^{\psi(s)\tau} P_{\text{end}}(\mathcal{C}|s) [A_s + O(e^{-\tau\Delta})] \quad (22)$$

for some constant  $A_s$  (independent of  $\tau$ ). In the correction term,  $\Delta$  is the gap between the largest and second-largest eigenvalues of  $\mathcal{W}_s$ .<sup>6</sup> Integrating over  $\mathcal{C}$  one sees that  $G(s, \tau) \approx \tau\psi(s) + \log A_s$ , consistent with (5). By (15), we also identify  $\rho(\mathcal{C}|s, \tau)e^{-G(s, \tau)}$  with  $\langle \delta(\mathcal{C} - \mathcal{C}_\tau) \rangle_s$ . Taking  $\tau \rightarrow \infty$  one sees from (22,5) that it is consistent to identify the eigenvector of  $\mathcal{W}_s$  with  $P_{\text{end}}$  as defined in (17).

Note that the operator  $\mathcal{W}_s$  is not generally Hermitian (self-adjoint). The eigenvector that we identified here as  $P_{\text{end}}$  is the right eigenvector. The role of the left eigenvector will be discussed in the next section.

To summarise, large deviations of  $b_\tau$  can be characterised by analysing the properties of the tilted operator  $\mathcal{W}_s$ , as in [12,13]. This approach is valuable as a tool for explicit computations (especially in finite-state Markov chains where the matrix  $M$  is finite). In addition, it establishes a connection between large-deviation problems and eigenproblems that are familiar from quantum mechanics. Like the analogy with thermodynamics discussed above, this connection with quantum mechanics is useful in practice because it means that methods from that field can be generalised in order to analyse large deviations [30,39,58,59].

## 2.6 Control representation and auxiliary process

This section describes a second method for analysis of large deviations, based on optimal control theory [60]. One advantage of this method is that it is built on a variational formula, which can be very useful for deriving approximate results in situations where diagonalisation of  $\mathcal{W}_s$  is not possible. The method has a transparent physical interpretation which is that (rare) large deviation events can be characterised by deriving a new physical model whose *typical trajectories* resemble closely the rare events of interest. This new model is called here the *optimally controlled process*, following earlier work by Fleming [61] and (more generally) the book of Dupuis and Ellis [38]. In previous work it has been called a driven process [47,48] or an auxiliary process [30,62], see also [39,63–65].

Consider first a general controlled process (not necessarily optimal). Let  $\langle F \rangle_{\text{con}}$  denote the average of a path-dependent quantity  $F$ , in this process. The probability density for trajectories in the controlled process is  $P_\tau^{\text{con}}(\mathcal{C})$ . Then a useful general formula ([38], Prop. 1.4.2) is

$$G(s, \tau) \geq -s\tau \langle \bar{b}_\tau \rangle_{\text{con}} - \mathcal{D}(P_\tau^{\text{con}} || P_\tau) \quad (23)$$

<sup>6</sup>For the cases described in Section 2.2, the gap  $\Delta$  is strictly positive. Models (and limits) where  $\Delta$  vanishes are often associated with anomalous fluctuations, including dynamical phase transitions, see Section 3.

where

$$\mathcal{D}(P_\tau^{\text{con}} || P_\tau) = \left\langle \log \frac{P_\tau^{\text{con}}(\mathcal{C})}{P_\tau(\mathcal{C})} \right\rangle_{\text{con}} \quad (24)$$

is the Kullback-Leibler (KL) divergence between  $P^{\text{con}}$  and  $P$ .<sup>7</sup> The KL divergence is non-negative and is equal to zero only if  $P_\tau^{\text{con}} = P_\tau$ . In the thermodynamic setting of Section 2.3, equation (23) is the Gibbs-Bogoliubov inequality, see [66], in particular their equation (25). It is possible to find a controlled process where (23) becomes an equality. To see this, use (15,24) to rewrite the right hand side of (23):

$$-s\tau \langle \bar{b}_\tau \rangle_{\text{con}} - \mathcal{D}(P_\tau^{\text{con}} || P_\tau) = G(s, \tau) - \mathcal{D}(P_\tau^{\text{con}} || P_\tau^s). \quad (25)$$

Hence equality is possible in (23) only if  $P_\tau^{\text{con}} = P_\tau^s$ : the controlled process must reproduce the probability distribution of the  $s$ -ensemble.<sup>8</sup>

The bound (23) can be analysed using tools from stochastic optimal control theory [60], see also [48]. The general aim of this theory is to find (controlled) Markov processes that maximise (or minimise) quantities like the right hand side of (23), which are interpreted as cost functions. For example, the process  $P^{\text{con}}$  might consist of requests which arrive randomly in a queue, and a stochastic rule for dealing with these requests. In this case a suitable cost would be some combination of the mean waiting time in the queue and the resource required to implement the policy. One seeks the policy that minimises the cost. Such problems have been studied in detail, they are obviously applicable in practical settings and they are also mathematically tractable [60].

Returning to the large-deviation context, observe that computation of the large-deviation rate function does not require a full characterisation of  $G(s, \tau)$  but only of  $\psi(s)$ , which is related to  $G(s, \tau)$  by (5). Hence

$$\psi(s) \geq \lim_{\tau \rightarrow \infty} \left[ -s \langle \bar{b}_\tau \rangle_{\text{con}} - \frac{1}{\tau} \mathcal{D}(P_\tau^{\text{con}} || P_\tau) \right]. \quad (26)$$

A key observation is that for the standard cases of Section 2.2, equality can be achieved in this formula by an (optimally)-controlled process that is Markovian and stationary [30,47,67]. This is a very useful simplification. From a comparison with (6), one may expect that  $I(b) = \tau^{-1} \mathcal{D}(P_\tau^{\text{con}} || P_\tau)$ , where  $P^{\text{con}}$  is a controlled process with  $\langle \bar{b}_\tau \rangle_{\text{con}} = b$ . In this case,

$$I(b) = \inf \left[ \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \mathcal{D}(P_\tau^{\text{con}} || P_\tau) \right] \quad (27)$$

<sup>7</sup>In a more rigorous approach, the ratio of probability densities in this definition would be replaced by a Radon-Nikodym derivative.

<sup>8</sup>It is not trivial to construct a stochastic process whose probability distribution of trajectories achieves  $P_\tau^{\text{con}} = P_\tau^s$ , this is related to the theory of dynamic programming [60]. The construction of such a process is possible for all examples considered here, although the controlled process may be complicated. For example, its transition rates may depend on time, see for example [47].

where the minimisation is over stationary Markovian controlled processes for which  $\langle \bar{b}_\tau \rangle_{\text{con}} = b$ .<sup>9</sup>

The final result (27) has an intuitive interpretation, it states that the *least unlikely* mechanism for achieving a rare event with  $\bar{b}_\tau = b$  can be reproduced by a controlled process that minimises the KL divergence. A central idea of large-deviation theory [10] is that this least unlikely mechanism is sufficient to characterise the rare event. The variational principle means that the controlled process differs as little as possible from original process; the size of the difference is quantified via the KL divergence.

### 2.7 Equivalence of different large deviation problems

An interesting aspect of the theory presented here is that the same optimally-controlled process may appear as the solution to several different large deviation problems. In the operator formalism, this happens because the same operator  $\mathcal{W}_s$  may appear in several different contexts. In fact, this is a very common situation. To see the reason, we define

$$\bar{g}_\tau(\mathbf{C}) = \frac{1}{\tau} \log \frac{P_\tau^{\text{con}}(\mathbf{C})}{P_\tau(\mathbf{C})}. \tag{28}$$

For models in the scope of Section 2.2, the quantity  $\bar{g}_\tau$  has a representation as either (9) or (11). Hence one sees that the biased ensemble  $P_\tau^s$  of (14) can also be characterised as a biased ensemble for the controlled process:

$$P_\tau^s(\mathbf{C}) \propto P_\tau^{\text{con}}(\mathbf{C}) e^{-s\tau \bar{b}_\tau(\mathbf{C}) - \tau \bar{g}_\tau(\mathbf{C})}. \tag{29}$$

Given a biased ensemble of interest, one may choose the controlled process (and hence  $\bar{g}$ ) in order to transform the problem into a form that is more tractable. This is very useful for numerical work [68–72]. It also enables analytic progress. For example, in biased ensembles where  $\bar{b}_\tau$  is of the form given in (9) of (11), it is simple to construct a controlled process such that the quantity  $s\bar{b}_\tau(\mathbf{C}) + \bar{g}_\tau(\mathbf{C})$  that appears in (29) reduces to a simple time-integral as in (1). Hence biased ensembles  $P_s$  with  $\bar{b}_\tau$  as in (9) have alternative formulations where the dynamics is modified but the bias has the (simpler) form (1). This observation was used in [49] to relate large deviations of the dynamical activity in spin models to large deviations of the time-integrated escape rate, see also [73] which discusses some relationships between large deviations of currents and dynamical activities.

### 2.8 Connection of operator and optimal-control approaches

There is a deep connection between the optimal control approach of Section 2.6 and the operator approach of Section 2.5. A similar connection appears in quantum

<sup>9</sup> The results (23,26) are extremely general but (27) is similar to (8) in that it requires assumptions related to analyticity and convexity of  $\psi$  and  $I$ . These assumptions are valid for models within the scope of Section 2.2.

mechanics, where one may use either an operator approach or an approach based on path integrals.

A general method to connect operator equations and controlled processes is to maximise the right hand side of (26) over some class of controlled processes, in order to find an optimally-controlled model. This variational problem is equivalent to solving for the largest eigenvalue of an operator  $\mathcal{W}_s^\dagger$ , which is the Hermitian conjugate (adjoint) of the operator  $\mathcal{W}_s$  discussed above. (The eigenvalue appears as the value of a Lagrange multiplier.) We present an example calculation for a simple diffusion process, after which we summarise the resulting general picture.

Consider large deviations of  $\bar{b}_\tau$  as in (1), for a diffusion problem described by a stochastic differential equation with additive noise:

$$d\mathcal{C}_t = v(\mathcal{C}_t)dt + \sqrt{2} dW_t, \tag{30}$$

where  $\mathcal{C}_t$  is a  $d$ -dimensional vector and  $W_t$  a  $d$ -dimensional standard Brownian motion. Using the operator method, the SCGF can be obtained for this process by solving the eigenvalue problem

$$\begin{aligned} \psi(s)P_{\text{end}} &= \mathcal{W}_s P_{\text{end}} \\ &= \nabla^2 P_{\text{end}} - \nabla \cdot (v P_{\text{end}}) - s b P_{\text{end}}. \end{aligned} \tag{31}$$

(The second line is an explicit formula for  $\mathcal{W}_s$ , the derivatives are with respect to  $\mathcal{C}$ .)

The controlled process is obtained from (30) by replacing  $v$  with  $(v - \nabla\phi)$  where  $\phi$  is a control-potential, that is

$$d\mathcal{C}_t = [v(\mathcal{C}_t) - \nabla\phi(\mathcal{C}_t)]dt + \sqrt{2} dW_t. \tag{32}$$

Similarly to [48,67], we show in Appendix A that if this control potential is used with (26), maximising the resulting bound on  $\psi$  is equivalent to solving the eigenproblem (31). In particular, the optimal control may be expressed as  $\phi = -2 \log \mathcal{F}$  where  $\mathcal{F}$  solves the eigenproblem

$$\psi \mathcal{F} = \mathcal{W}_s^\dagger \mathcal{F}, \tag{33}$$

in which  $\mathcal{W}_s^\dagger$  is the Hermitian conjugate (adjoint) of the operator  $\mathcal{W}_s$  given in (31). Its form is given in (A.7). Equation (33) is an eigenproblem for  $\mathcal{W}_s^\dagger$ , whose largest eigenvalue was already shown to be the SCGF  $\psi$ .<sup>10</sup> Constructing the controlled process from the corresponding eigenfunction  $\mathcal{F}$  achieves equality in (26) – hence this is an optimally-controlled process.

The conclusion of this analysis is that solving the eigenvalue problem (A.3) is equivalent to optimising (26) over controlled processes of the form (32). Also, the optimal control potential and the eigenvector are related as  $\mathcal{F} = e^{-\phi/2}$ . So the same information is available by the operator and optimal-control approaches.

We have analysed the simple model (30) but this structure is very general, see also [38]. Analogous steps can

<sup>10</sup> Of course,  $\mathcal{W}_s^\dagger$  and  $\mathcal{W}_s$  have the same eigenvalues.

be applied to all the models of Section 2.2. Taking  $\bar{b}_\tau$  as in (1) it is sufficient in these cases to consider controlled processes that are obtained by adding conservative control forces, as the derivative of a potential. For Markov chains with transition rates  $W(\mathcal{C} \rightarrow \mathcal{C}')$ , the appropriate controlled dynamics is [30,63]

$$W^{\text{con}}(\mathcal{C} \rightarrow \mathcal{C}') = e^{\phi(\mathcal{C})/2} W(\mathcal{C} \rightarrow \mathcal{C}') e^{-\phi(\mathcal{C}')/2}. \quad (34)$$

For  $\bar{b}_\tau$  as in (9,11) one should first use the method of Section 2.7 to transform the problem to a form where  $\bar{b}_\tau$  has the form given in (1): this may require a non-conservative control force. One then adds an additional conservative control force, as the gradient of  $\phi$ . It is sufficient to optimise over this  $\phi$ .

We end this section by observing that for time-reversal symmetric systems, both the eigenvalue problem and the optimal-control problem can be simplified. If (30) represents an equilibrium (time-reversal symmetric) system then  $v = -\nabla U$  for some potential  $U$ , so the controlled system (32) is also time-reversal symmetric (with potential  $U + \phi$ ). In this case the steady state of the controlled system is a Boltzmann distribution  $\mu \propto e^{-(U+\phi)}$ . Then (23) yields a simple variational result

$$\psi(s) = \sup_{\mathcal{F}} \frac{\int e^{-U} (\mathcal{F} \mathcal{W}_s^\dagger \mathcal{F}) d\mathcal{C}}{\int e^{-U} \mathcal{F}^2 d\mathcal{C}}, \quad (35)$$

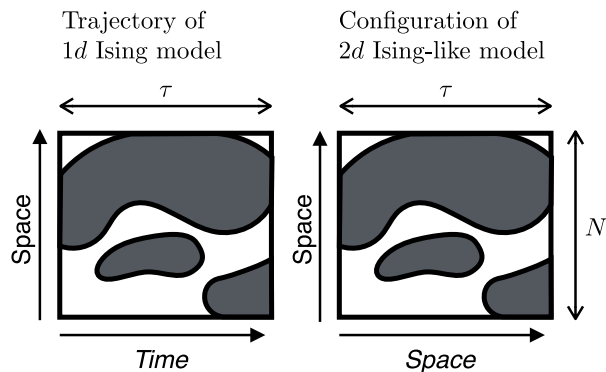
which is equivalent to the Rayleigh-Ritz formula for the largest eigenvalue of a self-adjoint operator, see also [11]. The physical origin of this simplification is the time-reversal symmetry of the biased ensemble (15). The  $\mathcal{F}$  that maximises the right hand side of (35) is the eigenfunction of  $\mathcal{W}_s^\dagger$  and gives the optimal control potential as  $\phi = -2 \log \mathcal{F}$ .

### 3 Dynamical phase transitions

We emphasised in Section 2.2 that finite systems are typically associated with analytic rate functions and SCGFs. However, there are many examples of rate functions that have singularities. For example, this can occur in Markov chains with infinite state spaces [17,30,74,75], which are not covered by Section 2.2. Motivated by the analogy with thermodynamics discussed in Section 2.3, these singularities can be identified as phase transitions.

Physically, the key feature is that singularities are (usually) associated with a qualitative difference in mechanism between rare events with different values of  $\bar{b}_\tau$ . There are several situations in which such behaviour can arise. We focus here on one broad class of phase transitions, which we describe as *space-time phase transitions*, sometimes called *trajectory phase transitions* [17], see also [76]. These occur in large systems where the observable  $b$  in (1) is an intensive variable in the spatial (thermodynamic) sense, see below.

Other kinds of dynamical phase transition have also been discussed in the context of dynamical large deviations [74,75,77–80]. Those results show that singular rate



**Fig. 2.** The idea of space-time thermodynamics [49,81,82] is that dynamical trajectories of  $d$ -dimensional models (left) can be analysed by mapping them to configurations of  $(d+1)$  thermodynamic models (right). The thermodynamic system has size  $N \times \tau$  and one analyses its behaviour in the joint limit  $N, \tau \rightarrow \infty$ .

functions can occur for a variety of different reasons. They also show that systems outside the scope of Section 2.2 cannot be assumed to have analytic rate functions, even if the models appear very simple.

#### 3.1 Thermodynamics in space-time

In Section 2.3 we described an analogy between large deviations of time-averaged quantities and the thermodynamic limit. In this section we are concerned with large deviations of time-averaged quantities in large systems. As a guiding example, we consider the one-dimensional Ising model with periodic boundaries, evolving by Glauber dynamics, as in [30]. There are  $N$  spins and the state of the  $i$ th spin at time  $t$  is  $\sigma_{i,t} = \pm 1$ . We consider a joint limit of large time  $\tau \rightarrow \infty$  and large system size  $N \rightarrow \infty$ .

To analyse this situation it is useful to make a mapping between trajectories of a  $d$ -dimensional model and configurations of a corresponding  $(d+1)$ -dimensional thermodynamic system [39,46,49,81–83]. The key idea is that the time  $t$  in the dynamical model is interpreted as an additional spatial co-ordinate in the thermodynamic system. Figure 2 illustrates this mapping for the 1d Glauber-Ising model, for which the corresponding thermodynamic system is a variant of the 2d Ising model.<sup>11</sup>

In the general case, we use the same symbol  $\mathcal{C}$  to indicate a trajectory of the dynamical model (as in Sect. 2.4) and the corresponding configuration of the  $(d+1)$ -dimensional thermodynamic model. We define a Boltzmann distribution for the  $(d+1)$ -dimensional model by assigning probability  $P_\tau(\mathcal{C})$  to configuration  $\mathcal{C}$ . This means that fluctuations in the dynamical model can (in principle) be analysed by applying methods of equilibrium statistical mechanics to the Boltzmann distribution of the  $(d+1)$ -dimensional system.

<sup>11</sup> This 2d model is somewhat unusual in that its vertical (space-like) dimension is defined in terms of a lattice while its horizontal (time-like) dimension is continuous. Nevertheless, it is a bona-fide model that can be analysed by equilibrium statistical mechanics.



Consider large deviations of some dynamical quantity  $\bar{u}$  that corresponds to an *intensive* variable in the  $(d + 1)$ -dimensional system.<sup>12</sup> For the example of the Ising model, we consider the time-averaged energy per spin:

$$\bar{u}(\mathbf{C}) = \frac{1}{\tau} \int_0^\tau \varepsilon(\mathbf{C}_t) dt, \quad \varepsilon(\mathbf{C}) = \frac{-1}{2N} \sum_i \sigma_i \sigma_{i+1}. \quad (36)$$

For a general dynamical model (with finite  $N$ ) that falls in the scope of Section 2.2, large deviations of  $\bar{u}$  can be analysed following Section 2. It is convenient to perform this analysis by setting  $\bar{b}_\tau = N\bar{u}$ . Then the biased ensemble of (15) is

$$P_\tau^s(\mathbf{C}) = P_\tau(\mathbf{C}) e^{-s\tau N\bar{u}(\mathbf{C}) - G_N(s, \tau)} \quad (37)$$

with

$$G_N(s, \tau) = \log \langle e^{-s\tau N\bar{u}(\mathbf{C})} \rangle, \quad (38)$$

analogous to (2). Recalling that  $P_\tau(\mathbf{C})$  is a Boltzmann distribution for the  $(d + 1)$ -dimensional model, we identify  $P_\tau^s(\mathbf{C})$  in (37) as a Boltzmann distribution where the energy has been perturbed by the extensive quantity  $sN\tau\bar{u}$ .<sup>13</sup> Also  $G_N$  is the difference in free energy between the perturbed and unperturbed models. Since  $\bar{u}$  was assumed to be intensive, this  $(d + 1)$ -dimensional system has an extensive energy function. On general thermodynamic grounds [45] one therefore expects for  $N, \tau \rightarrow \infty$  that

$$\frac{1}{N\tau} G_N(s, \tau) \rightarrow \mathcal{G}(s) \quad (39)$$

where  $\mathcal{G}$  is the bulk free-energy density. We recall from thermodynamics that there are no phase transitions in finite systems: in the present context this means that  $G_N(s, \tau)$  should always be an analytic function of  $s$ . However, the limiting function  $\mathcal{G}(s)$  may have singularities, which correspond to thermodynamic phase transitions in the  $(d + 1)$ -dimensional model. In the dynamical context, we refer to these as space-time phase transitions.

### 3.2 Space-time phase transitions

To analyse these phase transitions, it is convenient to first take  $\tau \rightarrow \infty$  at fixed  $N$ , and then later  $N \rightarrow \infty$ . At fixed  $N$ , we define a SCGF by analogy with (5)

$$\psi_N(s) = \frac{1}{N} \lim_{\tau \rightarrow \infty} \frac{1}{\tau} G_N(s, \tau) \quad (40)$$

<sup>12</sup> For our purposes, an extensive variable can be defined (loosely) as a quantity that is obtained by integrating a local quantity over a large system. The quantity  $\bar{u}$  is intensive if and only if  $\tau N\bar{u}$  is extensive.

<sup>13</sup> The temperature of the  $(d + 1)$ -dimensional model has been set to unity.

and a rate function by analogy with (3)

$$I_N(u) = \frac{1}{N} \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \log p_N(u|\tau). \quad (41)$$

(The factors of  $N$  are included for later convenience.) The assumptions of Section 2.2 are sufficient to ensure that  $\psi_N$  and  $I_N$  are analytic and strictly convex. The analogue of (6) is  $\psi_N(s) = \sup_u [-su - I_N(u)]$ . For large systems we are motivated by (39) to define

$$\mathcal{G}(s) = \lim_{N \rightarrow \infty} \psi_N(s) \quad (42)$$

and also

$$\mathcal{I}(b) = \lim_{N \rightarrow \infty} I_N(b). \quad (43)$$

These functions may not be analytic. However, the convexity of  $I_N$  means that

$$\mathcal{I}(u) = \sup_s [-su - \mathcal{G}(s)]. \quad (44)$$

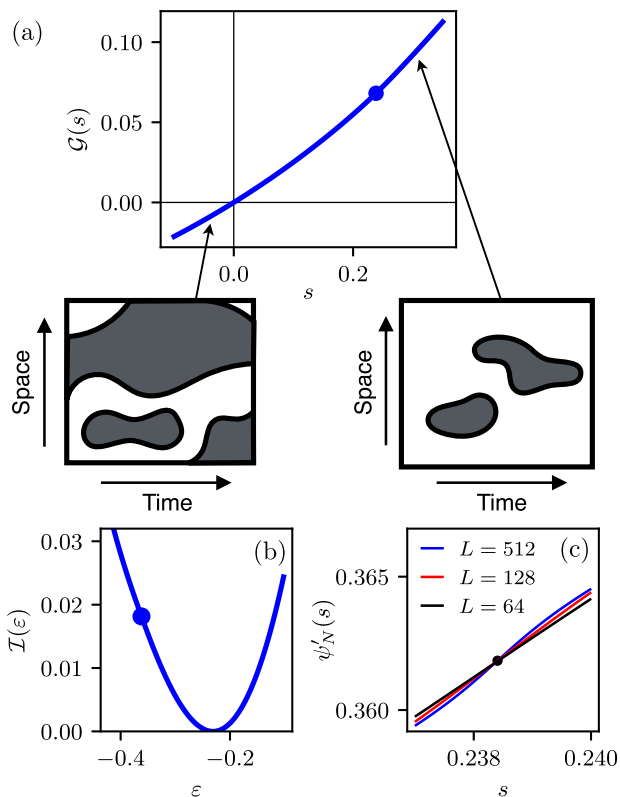
As an example of a dynamical phase transition, Figure 3 shows the large-deviation behaviour of the time-integrated energy in the 1d Glauber-Ising model. Exact results are available for this model, see [30] and also Appendix B. We show results at inverse temperature  $\beta = 1$  but the qualitative behaviour is the same for all positive  $\beta$  [30]. There is a critical point at  $s = s_c$  where  $\mathcal{G}$  is singular, and there is a corresponding singularity in  $\mathcal{I}$ . This critical point separates a paramagnetic regime for small  $s$  and a ferromagnetic regime for  $s > s_c$ , as might be anticipated by the correspondence with the 2d Ising-like model shown in Figure 2. The transition may also be analysed via a mapping to a quantum phase transition [84], see [30].

In finite systems the function  $\psi_N(s)$  is analytic, as is  $I_N(s)$ . However the second derivative  $\psi_N''(s_c)$  diverges logarithmically with  $N$ : this is the (weak) specific-heat singularity of the 2d Ising universality class [84]. The singularity is illustrated in Figure 3c by plots of  $\psi_N'(s)$  close to  $s_c$ ; its gradient  $\psi_N''(s_c)$  grows (slowly) with  $N$ .

### 3.3 First order phase transitions and dynamical phase coexistence

In thermodynamics, first-order phase transitions are associated with phase coexistence phenomena. The same situation holds at first-order space-time phase transitions. However, the manifestation of this phenomenon may differ between thermodynamic and dynamical transitions. This can be illustrated by the finite-size scaling behaviour at these transitions [17,58,69,82]. We summarise the associated behaviour, a more detailed analysis can be found in [69,85].

Applying the thermodynamic analogy of Section 3.2, note that the associated thermodynamic model is anisotropic because the horizontal (time-like) and vertical (space-like) axes in Figure 3 are not equivalent. To reflect this, consider a  $d$ -dimensional system with  $N = L^d$  so that  $G_N(s, \tau)$  depends separately on  $L$  and  $\tau$ . In the analogy

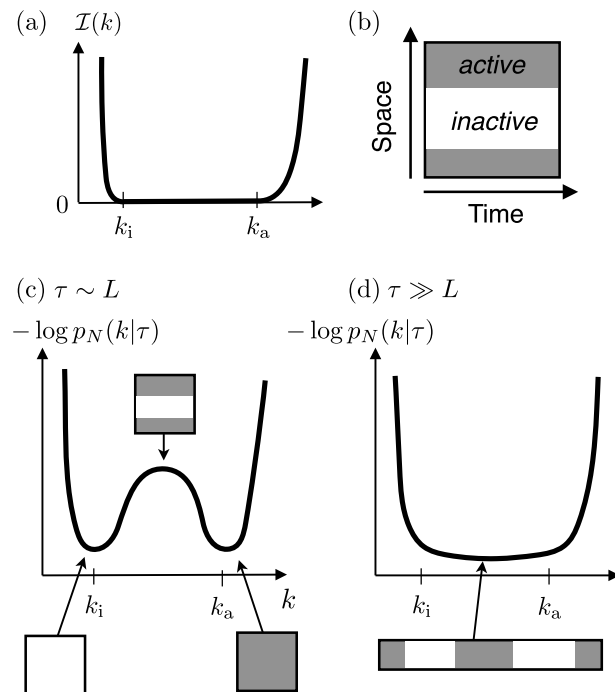


**Fig. 3.** Space-time phase transition for large deviations of the time-integrated energy in the 1d Ising model with Glauber dynamics at inverse temperature  $\beta = 1$ , using analytical results given in Appendix B, see [30]. (a) The limiting form of the SCGF as defined in (42). The filled circle indicates a critical point at  $s = s_c$ . Sketches of two representative trajectories are given, recall Figure 2. For  $s > s_c$  then trajectories exhibit long-ranged order in space and time, they resemble ferromagnetic configurations in a 2d Ising model. For  $s < s_c$  there is no long-ranged order (in the displayed range), so the behaviour resembles a paramagnet. (b) The limiting form of the rate function as defined in (43), the circle indicates the critical point. (c) The first derivative of the free energy in systems of three different sizes.

with thermodynamics,  $L^d \tau$  corresponds to the volume of the thermodynamic model, and  $\tau/L$  to its aspect ratio.

In large deviation analysis, a natural approach is to first take  $\tau \rightarrow \infty$  at fixed  $L$  as in (40), and then take  $L \rightarrow \infty$  as in (42). This means that the aspect ratio  $(\tau/L) \rightarrow \infty$ . However, in thermodynamic finite-size scaling analyses, it is more common to consider isotropic systems where the aspect ratio is fixed at unity [86], this corresponds to taking  $L, \tau \rightarrow \infty$  together. Nevertheless, thermodynamic systems with diverging aspect ratio have been analysed [87]: they provide a suitable comparison point for large-deviation analyses [69,85]. Limits where  $L, \tau \rightarrow \infty$  together have also been considered in numerical studies of large deviations [18,58].

The key fact is that physical behaviour at phase coexistence depends on the aspect ratio of the system. The situation is summarised in Figure 4. For  $\tau/L = O(1)$  one observes the familiar behaviour of thermodynamic



**Fig. 4.** Phase coexistence in space-time. (a) The limiting rate function defined as in (43) for a system exhibiting dynamical phase coexistence between two phases in which the order parameter has values  $k_i$  and  $k_a$ . (b) A trajectory exhibiting phase coexistence contains domains of both two dynamical phases, which are labelled as inactive ( $k \approx k_i$ ) and active ( $k \approx k_a$ ), see for example [81,82]. (c) Sketch of the probability distribution of  $k$  in a finite system where  $\tau$  is comparable to  $L$ , with sample trajectories that correspond to different values of  $k$ . Dynamical phase coexistence similar to (b) is associated with a local minimum of the probability, which is a local maximum in this plot. (d) Sketch of the same probability distribution in a finite system where  $\tau$  is very large, compared to  $L$ . In this case the distribution is unimodal and phase coexistence involves multiple domains arranged along the time-like axis. See [86,87] and [69,85].

phase coexistence, which means that the probability density  $p_N(b|\tau)$  is bimodal with two peaks corresponding to the coexisting phases, see Figure 4c and also [18,58]. The trough between the peaks corresponds to coexistence, where macroscopic domains of the phases are separated by an interface. On the other hand, if one takes instead a very large aspect ratio ( $\tau \rightarrow \infty$  before  $L \rightarrow \infty$ ) then  $I_N(b|\tau)$  in (41) is strictly convex so  $p_N$  is unimodal. In this case typical trajectories include many large domains of each phase, which are arranged along the time-like axis, see Figure 4d and also [69,85].

To summarise the central message of space-time thermodynamics: large-deviation theory can be applied to time-averages of (spatially) intensive quantities. The results can be understood by analogy with  $(d + 1)$ -dimensional thermodynamic systems. A natural approach to this limit is to consider the behaviour of  $G_N$  and  $I_N$  as  $N \rightarrow \infty$ , which means that we take a limit of large time *before* any limit of large  $N$ . In this case  $G_N$  and  $I_N$

are both analytic convex functions that converge to non-analytic limits as  $N \rightarrow \infty$ . This signals that a space-time phase transition is taking place.

## 4 Glassy systems and metastability

Interesting examples of space-time phase transitions appear in glassy systems, including supercooled liquids [18]. The dynamical behaviour of these systems continues to challenge theoretical understanding [88,89]. The structural relaxation time of a liquid is the time required for a molecule to diffuse a distance comparable with its (microscopic) diameter. In a simple liquid at a moderate temperature, this time might be a few picoseconds. On cooling through the glass transition, the structural relaxation time increases rapidly and eventually exceeds the (macroscopic) experimental time scale, which might be seconds or hours. For practical purposes, the system is no longer ergodic. The spatial correlations between molecules change only slightly as the system approaches its glass transition, but the system's dynamical properties change dramatically.

### 4.1 Dynamical phase transitions in glasses

Observing that the glass transition is a dynamical phenomenon, Merolle et al. [81] applied thermodynamic methods to the statistics of  $(d+1)$ -dimensional trajectories, similarly to Section 3.2 above, see also [82]. Their idea was that this methodology might capture information that is not available from standard thermodynamic methods. Early studies [81,82] focussed on simple kinetically-constrained lattice models (KCMs), which capture many of the dynamical features of glassy systems [89]. They considered fluctuations of the time-averaged dynamical activity, which in spin models is defined by counting the total number of configuration changes in a trajectory. This is a proxy for the extent to which molecules in a supercooled liquid are able to move around and explore their environment [18].

The connection of [81,82] to large deviation theory was realised shortly afterwards, and it was shown that dynamical phase transitions occur generically in KCMs [17,49]. This result is discussed in Section 4.2, below. It is notable because KCMs do not exhibit thermodynamic phase transitions, raising the possibility that the experimental glass transition might be related to an underlying dynamical phase transition, even in a system with simple thermodynamic properties [89].

Following this work on KCMs, numerical studies of atomistic models of liquids have shown evidence for dynamical phase transitions [18–20,90–92]. Large deviations have been analysed for a variety of time-averaged quantities including several different definitions of dynamical activity [19,90,91], and measures of liquid structure [20,92]. There is also evidence for dynamical phase transitions in experiments on glassy colloidal systems [21,93]. Some glassy spin models have thermodynamic glass transitions, and numerical and analytic arguments indicate that these models should also support dynamical

transitions [94]. Together, these works show that glassy systems generically exhibit large fluctuations, which can be probed by a variety of time-averaged quantities, and can be characterised via rate functions.

To explain the dynamical phase transition that takes place in KCMs, we discuss the prototypical example of the Fredrickson-Andersen (FA) model [95] in one dimension. This was one of the first glassy systems [81] for which large deviations were analysed. The existence of the phase transition can be proved by a very simple argument [17,49]. More recent work has characterised this transition in detail [59,69,96,97], as well as other large-deviation properties of this model [59,98,99].

### 4.2 Dynamical phase transition in the FA model

The FA model (in one dimension) consists of  $N$  spins in a linear chain with periodic boundaries. The state of the  $i$ th spin is  $n_i = 0, 1$  and a configuration of the system is  $\mathcal{C} = (n_1, n_2, \dots, n_N)$ . Spins with  $n_i = 1$  are active and indicate excitations, which are regions of a glassy system where particles are moving more than is typical. Spins with  $n_i = 0$  are inactive. The kinetic constraint is that spin  $i$  can change its state only if at least one of its neighbours  $n_{i\pm 1}$  are active. If this constraint is satisfied then spin  $i$  flips from state 0 to state 1 with rate  $c$ , while the reverse process happens with rate  $1 - c$ .

The behaviour of the model depends on the parameter  $c$ . In particular, for a system at equilibrium then the fraction of spins that are in state 1 is  $\langle n_i \rangle = c$ . The dependence of the model on temperature  $T$  is captured by identifying  $c = e^{-J/(k_B T)}$  where  $J$  is the characteristic energy of an active site (excitation).<sup>14</sup>

Now let  $k_{i,\tau}$  be the number of times that spin  $i$  changes its state, between time zero and time  $\tau$ . Summing over all spins, a time-averaged (intensive) measure of dynamical activity is

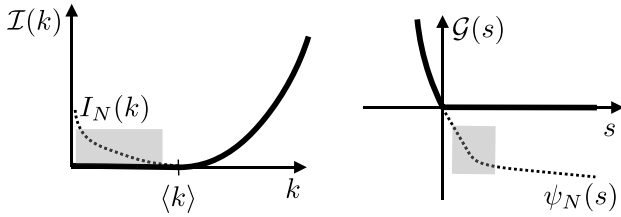
$$\bar{k}_\tau = \frac{1}{N\tau} \sum_i k_{i,\tau}. \quad (45)$$

This corresponds to (9) with  $\alpha(\mathcal{C}, \mathcal{C}') = (1/N)$  for all  $\mathcal{C}, \mathcal{C}'$ .

We analyse the large deviations of this activity by following Section 3.2 with  $\bar{u} \rightarrow \bar{k}_\tau$  (this is similar to Section 2.1, replacing  $\bar{b}_\tau \rightarrow N\bar{k}_\tau$ ). The following very simple argument shows that the functions  $\mathcal{G}$  and  $\mathcal{I}$  have singularities that correspond to first-order phase transitions. Consider the configuration with  $n_1 = 1$  and  $n_i = 0$  for all other sites. The rate of transitions out of this configuration is  $2c$ ; the probability that it occurs as initial condition is denoted by  $\pi_1$ .

Now define a very simplistic controlled process where the system begins in this configuration and never leaves

<sup>14</sup> Note also: if  $n_i = 0$  for all  $i$  then the configuration of the system can never change. For studies of large deviations it is therefore convenient to define the model on a configuration space that excludes this configuration [49]. In this case the model is irreducible and falls within the scope of Section 2.2.



**Fig. 5.** Dynamical phase transition for large deviations of the activity in the FA model. Heavy solid lines are sketches of the scaled rate function  $\mathcal{I}$  and the bulk free energy density  $\mathcal{G}$ . As discussed in the text,  $\mathcal{G}$  is singular at  $s = 0$  and  $\mathcal{G}(s) = 0$  for  $s \geq 0$ . Similarly  $\mathcal{I}(k) = 0$  for all  $k \leq \langle k \rangle$ . Dotted lines indicate the qualitative behaviour in finite systems, they are sketches of  $I_N(k)$  and  $\psi_N(s)$  as defined in (41,40). These functions are convex and analytic but they converge to the (non-analytic) limits  $\mathcal{I}$ ,  $\mathcal{G}$  as  $N \rightarrow \infty$ . The shaded boxes indicate the range over which the finite-size scaling analysis of [69] is relevant.

it. For this trajectory one has

$$\frac{P_\tau^{\text{con}}(\mathcal{C})}{P_\tau(\mathcal{C})} = \frac{1}{\pi_1 e^{-2c\tau}}. \quad (46)$$

Using this result with (23,39,42) and noting that  $\pi_1$  is independent of  $\tau$ , one obtains  $\psi_N > -2c/N$  and hence

$$\mathcal{G}(s) \geq 0. \quad (47)$$

Since the activity  $\bar{k}_\tau \geq 0$ , it follows from (2,47) that  $\mathcal{G}(s) = 0$  for  $s \geq 0$ .

Figure 5 illustrates the result: there is a discontinuity in the first derivative of  $\mathcal{G}$  at  $s = 0$ , which corresponds to a first-order space-time phase transition.<sup>15</sup> Applying (8), it follows that  $\mathcal{I}(k) = 0$  for all  $k < \langle k_\tau \rangle$ . This means that for large  $N, \tau$ , rare events where  $\bar{k}_\tau$  is smaller than its average have log-probabilities that do not scale as  $N\tau$ . In fact, these log-probabilities are much smaller: they are either proportional to  $N$  or  $\tau$ , depending on the relative magnitudes of these two quantities [82].

We note that the bound (47) is very general in KCMs, and establishes that these phase transitions occur in many different models [17,49]. However, it does rely on the existence of a “hard” kinetic constraint, which means that for a typical configuration  $\mathcal{C}$ , there are spins which cannot flip. This is a strong assumption and leaves open the question as to whether similar phase transitions are possible in models with softened constraints as in [58], where every spin flips with a non-zero rate. In fact similar (first-order) dynamical phase transitions still occur in the softened FA model [58], although in this case the singularity in  $\mathcal{G}(s)$  occurs at  $s^* > 0$ , and the only zero of  $\mathcal{I}(k)$  is at  $k = \langle k_\tau \rangle$ .

### 4.3 Large deviations and metastable states

These results for kinetically constrained models show that glassy systems with simple thermodynamic properties can still exhibit dynamical phase transitions. However, other

<sup>15</sup> In order to establish this one must show that  $\lim_{s \uparrow 0} \mathcal{G}'(s) > 0$ , this is straightforward [49].

theories of the glass transition assert that slow relaxation in liquids is linked to long-lived metastable states that can be analysed thermodynamically. This theoretical paradigm is certainly valid in a class of mean-field spin glasses,<sup>16</sup> while research continues into the question of whether it applies in physical (three-dimensional) liquids [100]. Some mean-field spin-glass models exhibit first-order dynamical phase transitions [94], similar to those in kinetically constrained models. The operator approach of Section 2.5 has been used to show that long-lived metastable states lead naturally to such transitions [94]. Here we give a brief explanation as to how the same conclusions can be reached (perhaps more intuitively) by an optimal-control argument.

Metastability is associated with a separation of time scales. The physical idea – which can be applied in non-equilibrium systems as well as in equilibrium [101–103] – is that if a system is initialised in a metastable state then it equilibrates quickly within that state, on a time scale  $\tau^f = O(1)$ , before eventually relaxing to some other state on a much longer time scale  $\tau^s \gg 1$ .

Consider a system with  $n \geq 2$  states, labelled by  $\alpha = 1, 2, \dots, n$ . This includes the case where one state is stable and the others are metastable (for example a mean-field ferromagnet in a field). It also includes systems at thermodynamic phase coexistence, which have two or more stable states.<sup>17</sup> Let  $\pi_\alpha$  be the probability that a steady-state configuration belongs to state  $\alpha$ . We analyse large deviations of an intensive observable  $\bar{u}$  that has different average values in each state: we denote these averages by  $\langle u \rangle_\alpha$ . In cases where the time scales are well-separated and the metastable states are well-defined then  $\sum_\alpha \pi_\alpha \simeq 1$  and the steady-state average of  $u$  is  $\langle u \rangle \simeq \sum_\alpha \pi_\alpha \langle u \rangle_\alpha$ . These approximate equalities are accurate if  $\tau_s \gg \tau_f$ .

Following (46) as well as [32,94] we consider a controlled process that starts in state  $\alpha$  and remains there for the entire trajectory. Its behaviour within state  $\alpha$  matches the natural dynamics of the model within that state. Since relaxation is fast within the metastable state, the time for the original (uncontrolled) model to leave this state is exponentially distributed with a mean that we denote by  $\tau_\alpha^s$ . By analogy with (46), we deduce that

$$\frac{P_\tau^{\text{con}}(\mathcal{C})}{P_\tau(\mathcal{C})} \simeq \frac{1}{\pi_\alpha e^{-\tau/\tau_\alpha^s}}. \quad (48)$$

As usual we consider large systems,  $N \rightarrow \infty$ . In idealised cases such as mean-field ferromagnets, the slow relaxation between states occurs on time scale  $\tau^s \sim e^{\kappa N}$  where  $N$  is the system size and  $\kappa = O(1)$ . If state  $\alpha$  is metastable then  $\pi_\alpha \sim e^{-N\Delta f}$  where  $\Delta f = O(1)$  is a difference in (intensive) free energy; if  $\alpha$  is stable then  $\pi_\alpha = O(1)$ . Using (48) with (23) and  $\bar{b} = N\bar{u}$  shows that

$$G_N(s, \tau) \geq -sN\tau \langle u \rangle_\alpha + \log \pi_\alpha - (\tau/\tau_\alpha^s). \quad (49)$$

<sup>16</sup> In this context, the “mean-field” nomenclature means that the strength of the interaction between spins is independent of the distance between them.

<sup>17</sup> For the purposes of this discussion, the difference between stable and metastable states is that metastable states have a vanishing probability in the steady state.



Taking  $\tau \rightarrow \infty$  at fixed  $N$  and using (40,42) gives

$$\mathcal{G}(s) \geq -s\langle u \rangle_\alpha - \lim_{N \rightarrow \infty} (N\tau_\alpha^s)^{-1}. \quad (50)$$

If  $\tau_\alpha^s \gg 1$  is a slow time scale then one sees that  $\mathcal{G}(s) \geq -s\langle u \rangle_\alpha$ . Using also that  $\psi'_N(0) = -\langle u \rangle$  and  $\langle u \rangle \neq \langle u \rangle_\alpha$ , this implies that  $\mathcal{G}'(s)$  has a discontinuity at  $s = 0$ , which corresponds to a first-order space-time phase transition, similar to the case of kinetically constrained models. A more detailed analysis of this case can be found in [94], using the operator approach.

We emphasise that such first-order transitions are generic for systems where  $(\tau^s/\tau^f) \rightarrow \infty$  which includes mean-field systems with metastable states, and finite-dimensional systems at phase coexistence. For finite-dimensional systems away from phase coexistence then all metastable states have finite lifetimes, and one expects  $(\tau_\alpha^s)^{-1} \propto N$ . (For example, recall that nucleation rates for systems close to phase coexistence are proportional to the system size  $N$  [5].) In such cases, (50) gives a bound on  $\mathcal{G}$  that is not sufficient to establish the existence of a phase transition, but can be used to relate crossovers in  $\mathcal{G}(s)$  and  $G_N(s, \tau)$  to properties of metastable states, particularly  $\langle u \rangle_\alpha$  and  $\tau_\alpha^s$  [32,94]. These arguments establish strong connections between metastability and large deviations, which (we argue) are very useful when interpreting large-deviation computations for glassy systems [18,20,92].

### 5 Fluctuation theorems and time's arrow

Glassy systems have slow dynamics but their equilibrium states are time-reversal symmetric. We now turn to models of non-equilibrium steady states. Early work in this area [8,12] demonstrated the usefulness of large deviation studies of time-averaged quantities in physics, by exploiting connections between dissipation and irreversibility. In this section we set Boltzmann's constant  $k_B = 1$ , so that entropy is a dimensionless quantity.

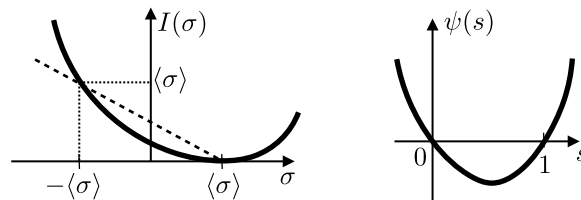
We write  $\mathcal{C}^R$  for the trajectory that is obtained by reversing the arrow of time in trajectory  $\mathcal{C}$ . In the simplest case, this means that  $\mathcal{C}_{\tau-t}^R = \mathcal{C}_t$ . More generally the time-reversal operation might involve a change in some system variables, such as reversal of molecular velocities, as in [104,105]. Then, a (time-integrated) measure of irreversibility for trajectory  $\mathcal{C}$  within a given model can be identified as

$$\Sigma_\tau(\mathcal{C}) = \log \frac{P_\tau(\mathcal{C})}{P_\tau(\mathcal{C}^R)}. \quad (51)$$

Recall that  $P$  includes the probability of the initial configuration of the system  $\mathcal{C}_0$  and that these initial conditions are taken from the steady state of the system. It follows that for equilibrium systems,  $\Sigma_\tau(\mathcal{C}) = 0$  exactly, for every time  $\tau$  and every trajectory  $\mathcal{C}$ .

It is useful to define

$$P_\tau^*(\mathcal{C}) = P_\tau(\mathcal{C}^R) \quad (52)$$



**Fig. 6.** Sketch of the rate function and SCGF for the entropy production in a generic system that obeys the fluctuation theorem. The dashed line in the left panel has gradient  $-1/2$ , it intersects the rate function at  $\sigma = \pm\langle\sigma\rangle$ , consistent with (61). The SCGF has reflection symmetry through  $s = \frac{1}{2}$ , consistent with (59).

and to identify this  $P^*$  as the probability distribution for trajectories under a particular controlled process which we refer to as the adjoint process, following [2]. One sees that

$$\langle \Sigma_\tau(\mathcal{C}) \rangle = \mathcal{D}(P||P^*) \geq 0. \quad (53)$$

The mean entropy production can never be negative; it is zero only for time-reversal symmetric (equilibrium) systems, since  $P = P^*$  in that case.

One drawback of the irreversibility measure  $\Sigma_\tau$  is that the quantity  $P_\tau(\mathcal{C})$  appearing in (51) cannot usually be evaluated, because it depends on the probability of the initial state of the trajectory, which is typically not known (except in equilibrium systems where it only depends on the energy). However, one may define a time-averaged rate of entropy production as

$$\bar{\sigma}_\tau = \frac{1}{\tau} \left[ \Sigma_\tau(\mathcal{C}) - \log \frac{\pi(\mathcal{C}_0)}{\pi(\mathcal{C}_0^R)} \right] \quad (54)$$

where we recall that  $\pi(\mathcal{C}_0)$  is the probability density for the initial condition of the trajectory.<sup>18</sup> In non-equilibrium systems, the usual situation is that  $\Sigma_\tau$  grows with  $\tau$  while  $\log \frac{\pi(\mathcal{C}_0)}{\pi(\mathcal{C}_0^R)}$  remains finite.<sup>19</sup> In this case, the large deviations of  $\bar{\sigma}_\tau$  are the same as the large deviations of  $\Sigma_\tau/\tau$ , even if these quantities have different values when  $\tau$  is finite.

In many physical systems, closed formulae for  $\bar{\sigma}_\tau$  are available. For example, consider a simple model for particle motion (in  $d$  dimensions)

$$d\mathcal{C}_t = \frac{f(\mathcal{C}_t)}{\gamma} dt + \sqrt{\frac{2T}{\gamma}} dW_t, \quad (55)$$

similar to (30). The natural physical interpretation of this model is that a particle moves through a viscous fluid with friction constant  $\gamma$  at temperature  $T$ , and feels a non-conservative external force  $f(\mathcal{C}_t)$ . Then it may be shown

<sup>18</sup> Note  $\mathcal{C}_0^R$  is the initial condition of the time-reversed trajectory which in the simplest case coincides with  $\mathcal{C}_\tau$ .

<sup>19</sup> This is certainly the case if  $\pi(\mathcal{C})$  is bounded, which holds for the finite systems of Section 2.2.

from (54) that

$$\bar{\sigma}_\tau = \frac{1}{\tau T} \int_0^\tau f(\mathcal{C}_t) \circ d\mathcal{C}_t. \quad (56)$$

We identify  $\tau T \bar{\sigma}_\tau$  as the total work done by the force  $f$  which coincides (in this simple situation) with the heat dissipated in the fluid. Dividing the dissipated heat by the temperature gives the entropy production, so the probabilistic definition of  $\bar{\sigma}_\tau$  in (54) coincides with the time-averaged rate of (physical) entropy production.

Note that heat and work coincide in this example system because all forces were assumed to be external: hence there is no internal energy. To separate the definitions of heat and work one should formulate the first law of thermodynamics by defining an internal energy  $U$  and a corresponding force  $-\nabla U$ . Then write  $f = F^{\text{ext}} - \nabla U$  in (55), where  $F^{\text{ext}}$  is an external force [3]. The work is then  $\int_0^\tau F^{\text{ext}}(\mathcal{C}_t) \circ d\mathcal{C}_t$ , and the heat transferred to the fluid is  $\int_0^\tau f(\mathcal{C}_t) \circ d\mathcal{C}_t$ , consistent with (54). The difference of these quantities is the change in internal energy: this is the first law of thermodynamics.

For Markov chains with jump rates  $W(\mathcal{C} \rightarrow \mathcal{C}')$ , the analogue of (56) is given by (9) with  $\alpha(\mathcal{C}, \mathcal{C}') = \log[W(\mathcal{C} \rightarrow \mathcal{C}')/W(\mathcal{C}' \rightarrow \mathcal{C})]$ , which requires the assumption that  $W(\mathcal{C} \rightarrow \mathcal{C}')$  is non-zero whenever  $W(\mathcal{C}' \rightarrow \mathcal{C})$  is non-zero. (This property is sometimes called weak reversibility.)

Returning to the main argument, it follows from the explicit formula (54) that large deviations of  $\bar{\sigma}_\tau$  can be analysed within the class of models discussed in Section 2.2. The connection of the entropy production  $\bar{\sigma}_\tau$  with the irreversibility measure  $\Sigma_\tau$  means that large deviations of  $\bar{\sigma}_\tau$  have interesting symmetry properties, as we now discuss.

### 5.1 Fluctuation theorem of Gallavotti-Cohen

We discuss fluctuation theorems for the entropy production in non-equilibrium steady states [3,8,12,83,106–108]. Consider first the CGF for  $\Sigma_\tau$ :

$$\begin{aligned} G^\Sigma(s, \tau) &= \log \langle e^{-s \Sigma_\tau} \rangle \\ &= \log \int \left( \frac{P_\tau(\mathcal{C}^R)}{P_\tau(\mathcal{C})} \right)^s P_\tau(\mathcal{C}) d\mathcal{C} \\ &= \log \int \left( \frac{P_\tau(\mathcal{C})}{P_\tau(\mathcal{C}^R)} \right)^{1-s} P_\tau(\mathcal{C}^R) d\mathcal{C} \end{aligned} \quad (57)$$

where the first line is the definition of  $G^\Sigma$ , the second line uses (51), and the third simply rearranges various terms. Changing integration variable from  $\mathcal{C}^R$  to  $\mathcal{C}$ , and using again (51) one finds the symmetry relation

$$G^\Sigma(s, \tau) = G^\Sigma(1-s, \tau). \quad (58)$$

See for example [12], where the quantity  $\Sigma_\tau$  was denoted by  $\bar{W}$ .

Now consider large deviations of the entropy production  $\bar{\sigma}_\tau$  whose SCGF is denoted here by  $\psi(s)$ . Recalling

(54), one may expect that the large deviations of  $\bar{\sigma}_\tau$  are the same as those of  $\Sigma_\tau/\tau$ , in which case one would have  $\psi(s) = \lim_{\tau \rightarrow \infty} \tau^{-1} G^\Sigma(s, \tau)$ . The relationship between  $\bar{\sigma}_\tau$  and  $\Sigma_\tau$  is discussed in [12], which showed (for several broad classes of stochastic model) that

$$\psi(s) = \psi(1-s). \quad (59)$$

This is the symmetry that was identified by Gallavotti and Cohen [8]. It is closely related to (58) but we note that (59) is a statement about large deviations of  $\bar{\sigma}_\tau$  as  $\tau \rightarrow \infty$ , in contrast to (58) which is a statement about  $\Sigma_\tau$  that is valid for all  $\tau$ . See also [3,108].

Now assume convexity of  $\psi$  and use (8) with (59) to write

$$I(-\sigma) = \sup_s [s\sigma - \psi(1-s)]. \quad (60)$$

Relabelling the dummy variable  $s = 1-x$  and using (8) one obtains a fluctuation theorem [12,83]

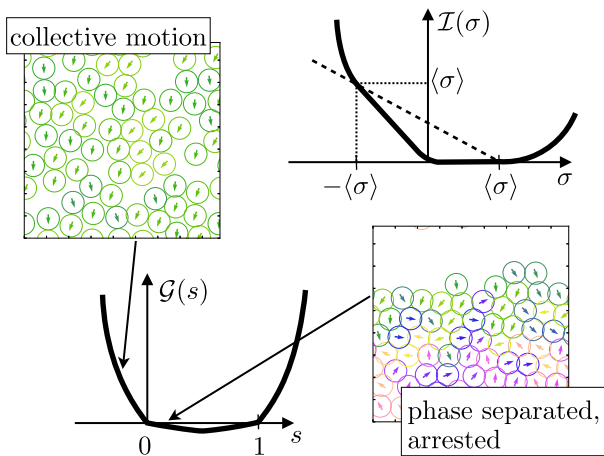
$$I(-\sigma) = I(\sigma) + \sigma. \quad (61)$$

Taking  $\sigma > 0$ , one sees that  $I(-\sigma)$  determines the log-probability of trajectories with negative entropy production. Since  $I(-\sigma) > I(\sigma)$ , these trajectories are exponentially rarer than trajectories with positive entropy production. (This can be interpreted as a statistical form of the second law of thermodynamics.) In addition, the difference in log-probability is given quantitatively by (61), so the fluctuation theorem (which is an equality) contains more information than the second law (which is an inequality).

For equilibrium systems we recall that  $\Sigma_\tau = 0$  exactly, so the methods of large-deviation theory are not relevant. At a formal level then  $I(\sigma) = \infty$  whenever  $\sigma \neq 0$ , and  $\psi(s) = 0$  for all  $s$ . For non-equilibrium systems, it is notable that the optimally-controlled process at  $s = \frac{1}{2}$  is time-reversal symmetric; also the optimally-controlled process at  $s = 1$  is the adjoint process, which corresponds to the original process running backwards in time, see for example [109].

### 5.2 Example: active Brownian particles

Fluctuation theorems such as (59) are very general results. However, the analysis of the fluctuations of the entropy production in specific systems can reveal additional rich structure. An interesting example is the behaviour of active-matter systems [72,110–112]. As an example we consider a system of active Brownian particles [113,114], as considered in [72]. It consists of  $N$  circular particles in a two-dimensional system of size  $L^2$ . Particle  $i$  has an orientation, which is represented by a unit vector  $\mathbf{e}_i$ . The particles interact by repulsive forces and they undergo thermal diffusion with diffusion constant  $D_0$ . In addition, they feel non-conservative propulsive forces of fixed strength which act along their orientation vectors. The propulsive forces are such that a single isolated particle moves with average speed  $v_0$ . Each orientation vector



**Fig. 7.** Illustration of the behaviour of the entropy production  $\sigma$  in a system of active Brownian particles, based on [72]. We show the limiting rate function  $\mathcal{I}$  and the corresponding free energy  $\mathcal{G}$ , which obey a Gallavotti-Cohen symmetry as in Figure 6. The free energy  $\mathcal{G}$  is singular at  $s = 0$ , and  $\mathcal{I}$  has two linear segments, see the text for a discussion. For  $s < 0$  (high entropy production), typical configurations involve spontaneous particle alignment and exhibit collective motion. For  $0 < s < 1$  (low entropy production) system enters a dynamically-arrested phase-separated state.

undergoes rotational diffusion, independent of all other co-ordinates.

Reference [72] considered large deviations of a quantity called the active work, which has a corresponding (intensive) measure of entropy production:

$$\bar{\sigma}_\tau = \frac{1}{N\tau} \sum_{i=1}^N \int_0^\tau \frac{v_0}{D_0} \mathbf{e}_i(t) \circ d\mathbf{r}_i(t) \quad (62)$$

where  $\mathbf{r}_i$  is the position of particle  $i$ , and the integral is evaluated using the Stratonovich convention. The physical interpretation of  $\bar{\sigma}_\tau$  is that there is a force on particle  $i$ , acting in direction  $\mathbf{e}_i$  (with constant magnitude). The integral in (62) is the work done by this force, normalised by the temperature. We refer to  $\bar{\sigma}_\tau$  as the entropy production, although other definitions of the entropy production are possible in such systems [115–118]. Large deviations of  $\bar{\sigma}_\tau$  obey the fluctuation theorems [59,61].

The resulting large-deviation phenomenology is illustrated in Figure 7, following [72]. We focus on large systems and we consider  $\mathcal{I}$  and  $\mathcal{G}$  as defined in (42,43). Note that  $\mathcal{I}$  obeys the fluctuation theorem (61) but its behaviour is quite different from the illustration in Figure 6. The reason is that this system exhibits several space-time phase transitions which appear in the limit of large system size. This corresponds to  $N \rightarrow \infty$  at a fixed overall density  $\rho_0 = N/L^2$ . In the following, it is sufficient to consider only  $\sigma > 0$ : the behaviour for  $\sigma < 0$  follows from the fluctuation theorem.

A first observation is that the behaviour for  $\sigma > 0$  and  $s < \frac{1}{2}$  in Figure 7 somewhat resembles Figure 5: there is a discontinuity in  $\mathcal{G}'(s)$  at  $s = 0$  and a range of  $\sigma$  over

which  $\mathcal{I}(\sigma) = 0$ . This was explained in [72] by an optimal-control argument: they proposed a controlled process that can be used with (27) to show that  $I_N(\sigma) \leq O(1/L)$  for a finite range of  $\sigma$  between 0 and  $\langle \sigma \rangle$ .<sup>20</sup> Hence  $\mathcal{I}(\sigma) = 0$  in this regime, by (43). The behaviour of the controlled system in this case is that the particles form a high-density cluster where particle motion is strongly reduced, and  $\sigma$  is small. Hence this state was called “phase-separated and arrested”. The associated reduction in particle motion is analogous to the transition to the inactive phase in Figure 5, which explains the similarity to that case, see also Section 6.

For large deviations with  $\sigma > \langle \sigma \rangle$ , the numerical results of [72] show spontaneous symmetry breaking, in that particles align their orientations with each other (Fig. 7). In this case they also move collectively through the system. For an intuitive understanding of this transition, it is useful to consider a controlled system where the particles’ orientation vectors feel forces (or torques) that tend to align them. Reference [72] considered a mean-field (infinite-ranged) interaction. If this interaction is strong enough to create long-ranged (ferromagnetic) order of the orientations, it clearly reduces the number of interparticle collisions, and this increases  $\sigma$ . This controlled process provides a bound on  $\mathcal{I}$  via (27) and numerical tests indicate that this bound is close to the true value of  $\mathcal{I}$ . The conclusion is that particle alignment is an effective mechanism for fluctuations of the entropy production.

The understanding of large deviations in this system is not yet complete, but it is clear from [72] that fluctuations with  $\sigma < \langle \sigma \rangle$  are strongly coupled to density fluctuations and the arrest of particle motion, while fluctuations with  $\sigma > \langle \sigma \rangle$  are associated with spontaneous symmetry breaking and particle alignment. This illustrates the rich large-deviation behaviour of these non-equilibrium systems.

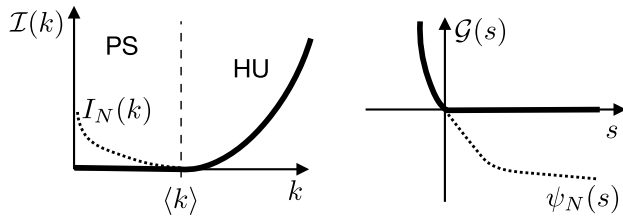
## 6 Exclusion processes and hydrodynamic behaviour

A very active area of large-deviation research is the behaviour of interacting-particle systems including exclusion processes and zero-range processes [2,13–16,74,119–121]. This section gives a brief overview of some of the relevant phenomena, focussing on the similarities and differences between these systems and those analysed in previous sections.

### 6.1 Activity fluctuations in the simple symmetric exclusion process

As a concrete example, we focus on the symmetric simple exclusion process (SSEP) with periodic boundaries, as considered in [122] as well as [71,73,123]. In this case,  $N$  particles move on the  $L$  sites of a one-dimensional periodic lattice, with at most one particle per site. Suppose that the particle hop rate is  $\gamma$  and the lattice spacing is

<sup>20</sup> An open question from that work is whether this range extends down to  $\sigma = 0$  or whether there it has a non-zero lower limit.



**Fig. 8.** Large deviations of the activity in the SSEP with periodic boundaries. There are some similarities with the results of Figure 5 for the FA model:  $\mathcal{G}$  is singular at  $s = 0$  and  $\mathcal{G}(s) = 0$  for  $s \geq 0$ . Also  $\mathcal{I}(k) = 0$  for all  $k \leq \langle k \rangle$ . Dotted lines indicate the qualitative behaviour in finite systems. In contrast to the FA model,  $\psi_N''(0)$  diverges with  $N$  and  $\mathcal{G}''(s)$  diverges as  $s \rightarrow 0^-$ . This is due to slow hydrodynamic fluctuations. As explained in the text, the system is macroscopically inhomogeneous (phase-separated, PS) for fluctuations with  $\bar{k}_\tau < \langle k \rangle$  [ $s > 0$ ] and hyperuniform (HU) for fluctuations with  $\bar{k}_\tau > \langle k \rangle$  [ $s < 0$ ]. The regime with  $s = O(L^{-2})$  can be characterised by macroscopic fluctuation theory, it includes the whole PS regime and corresponds to  $\mathcal{I}(k) = O(1/L)$ .

$a_0$  so that the diffusion constant for a single particle is  $D_0 = \gamma a_0^2/2$ .<sup>21</sup> In this system, exact results are available, for large deviations of the (time-averaged, intensive) particle current  $\bar{j}_\tau$  and the activity  $\bar{k}_\tau$  [122]. The current is defined using (9) with  $\alpha = 1/L$  when a particle hops to the right and  $\alpha = -1/L$  for hops to the left. Similarly the activity is defined by taking  $\alpha = 1/L$  for all hops.<sup>22</sup> We focus here on large deviations of the activity.

This process is a finite Markov chain, satisfying the conditions of Section 2.2. Hence the rate functions for finite systems are analytic and convex. The interesting behaviour occurs in the limit of large system size, which means that  $L \rightarrow \infty$  with a fixed mean density  $\bar{\rho} = N/L$ . This suggests that the space-time thermodynamic theory of Section 3.1 should be applicable. However, the number of particles is a conserved quantity in the SSEP, which means that the corresponding  $(d + 1)$ -dimensional thermodynamic model has some unusual features, from the thermodynamic perspective.

To understand dynamical large deviations, note first that if all the particles in the SSEP form a single cluster by occupying adjacent sites, then there are only two particle hops that are possible (at the edges of the cluster). In this case one may apply exactly the same argument as Section 4.2 to obtain

$$\psi_N(s) \geq -2\gamma/N \quad (63)$$

where  $\psi_N$  is the SCGF [as in (40)] for the activity  $\bar{k}_\tau$ . The activity  $\bar{k}_\tau \geq 0$  so  $\psi_N(s) \leq 0$  for positive  $s$ . This establishes that  $\psi_N(s) = O(1/N)$  for positive  $s$  (low activity).

<sup>21</sup> In the literature, one often measures time in units where  $D_0 = 1$  but we retain  $D_0$  here as a parameter.

<sup>22</sup> With this definition, the current and activity are intensive in the sense that  $\langle \bar{j}_\tau \rangle = O(1)$  as  $N \rightarrow \infty$ . In the literature it is more common to work with corresponding extensive observables, but we use the intensive versions here, to facilitate comparison with earlier sections.

On the other hand  $\psi_N(s)$  is of order unity for negative  $s$  (high activity). Defining  $\mathcal{G}$  as in (42) one arrives at a situation similar to Figure 5, with  $\mathcal{G}(s) = 0$  for  $s \geq 0$  while  $\mathcal{G}$  is of order unity for  $s < 0$ . The resulting situation is shown in Figure 8.

This result is correct but it misses some important properties of exclusion processes, for which one requires a more detailed analysis [122,123]. The SSEP has a slow diffusive time scale associated with large-scale density fluctuations  $\tau_L \sim L^2/D_0$ . These slow (hydrodynamic) fluctuations hinder ergodicity and tend to enhance the variance of time-averaged quantities. For example, it may be verified from [122] that the variance of  $\bar{k}_\tau$  behaves for large  $N, \tau$  as  $\text{Var}(\bar{k}_\tau) \propto 1/\tau$ , independent of  $N$ , and hence  $\psi_N''(0) = O(N)$ . This is in contrast to dynamical phase coexistence as it occurs in the FA model, where  $\psi_N''(0)$  is of order unity.

The activity fluctuations responsible for  $\psi_N''(0) \rightarrow \infty$  in the SSEP can be captured by macroscopic fluctuation theory [2]. It is convenient to rescale time by  $\tau_L$ : let

$$\tilde{t} = \frac{t}{\tau_L} = \frac{D_0 t}{L^2}. \quad (64)$$

Similarly  $\tilde{\tau} = D_0 \tau/L^2$ . The space-time thermodynamics approach of Section 3.1 focusses on large deviations with

$$-\log p_N(k|\tau) \simeq \tau N \mathcal{I}(k) \quad (65)$$

where  $\mathcal{I}(k)$  takes values of order unity. This is an LDP with speed  $\tau$ , where the rate function is proportional to  $N$ . By contrast, macroscopic fluctuation theory is a theory for large deviations with

$$\begin{aligned} -\log p_N(k|\tau) &\simeq \tilde{\tau} N \tilde{\mathcal{I}}(k) \\ &\simeq \tau \tilde{\mathcal{I}}(k) \frac{N D_0}{L^2} \end{aligned} \quad (66)$$

with  $\tilde{\mathcal{I}}(k)$  of order unity. From a physical perspective, the interpretation of this formula is that the log-probability of the large deviation is proportional to the system size  $N$  and to the time  $\tilde{\tau}$ , which is measured *on the hydrodynamic scale*. Just like (65), we interpret (66) as an LDP with speed  $\tau$ , but now with a rate function proportional to  $N/L^2$ . For this one-dimensional system then  $L \propto N$  as  $N \rightarrow \infty$ , so the rate function in (66) goes to zero with system size; this may be contrasted with (65), where the rate function diverges. In general, the question of whether (65) or (66) is applicable depends on whether the fluctuation of interest is governed by hydrodynamic (slow) variables or microscopic (fast) variables.

For the SSEP, the macroscopic fluctuation theory gives a quantitative description of fluctuations on the hydrodynamic scale. They can be analysed by considering a suitable SCGF, for small values of the biasing parameter  $s = O(N^{-2})$ , see [123] for details. The result is that (66) is applicable for large deviations throughout the range  $0 < \bar{k}_\tau < \langle k \rangle$ . In this case the fluctuation mechanism is that the SSEP becomes macroscopically inhomogeneous: it forms dense and dilute regions that suppress the activity.



However, for fluctuations where the intensive activity is significantly larger than  $\langle \bar{k}_\tau \rangle$ , the probability scales as in (65) and the macroscopic fluctuation theory is not applicable. Specifically, for small negative  $s$ , reference [122] gives

$$\mathcal{G}(s) \simeq -s\langle k \rangle + A(-s)^{3/2} \quad (67)$$

where the constant  $A = O(1)$  can be obtained by adapting ([122], Eq. 57) to the current notation. From (44) we see for  $k \geq \langle \bar{k}_\tau \rangle$  that

$$\mathcal{I}(k) \simeq \frac{4}{27A^2}(k - \langle \bar{k}_\tau \rangle)^3. \quad (68)$$

The second derivative  $\mathcal{G}''(s)$  diverges as  $s \rightarrow 0^-$ , while  $\mathcal{I}''(k)$  vanishes as  $k \rightarrow \langle \bar{k}_\tau \rangle$  from above. This is consistent with the scaling of the variance of  $\bar{k}_\tau$  (inversely proportional to  $\tau$  and independent of  $N$ ) and its link to hydrodynamic density fluctuations. In fact, these fluctuations have a strong dependence on  $s$ : for any  $s < 0$  the system is hyperuniform [73], which means that density fluctuations on large scales are very strongly suppressed [124].

## 6.2 General implications of hydrodynamic modes

We have explained that for the SSEP with periodic boundaries, high-activity fluctuations follow (65) and low-activity fluctuations follow (66). The low-activity regime may be analysed within macroscopic fluctuation theory, which can also be applied to large deviations in other interacting-particle systems, including (weakly) asymmetric exclusion processes and zero-range processes [2, 16, 75, 120]. Similar results can also be found in off-lattice models [125, 126].

An important general question in this area is whether slow (hydrodynamic) modes lead to fluctuations governed by (66). Macroscopic fluctuation theory provides a partial answer. We use the language of activity fluctuations but the argument is general. We introduce a notion of local equilibration within a spatial region of size  $\ell$ , with  $1 \ll \ell \ll L$ . A system is at local equilibrium [2] if the distribution of particles within that region resembles the natural (unbiased) system at the same (local) density. In this case the hydrodynamic behaviour can be analysed by considering the (smooth) density field, and an associated current.

Consider a system in  $d$  spatial dimensions so  $N \propto L^d$  and suppose that one can construct a macroscopically inhomogeneous state where the (total) activity differs from  $\langle k \rangle$  but the system is everywhere in local equilibrium. (Such states may be also be time-dependent, for example travelling waves [75], and there may be hydrodynamic flow of particles.) In this case, macroscopic fluctuation theory explains that the log-probability of fluctuations with this activity obeys (66). However, we now have  $N \propto L^d$  so the rate function scales as  $L^{d-2}$ . The physical interpretation is that local equilibrium states have densities that vary slowly in space: these smooth (hydrodynamic)

profiles relax slowly towards the steady state and can therefore be stabilised by adding very weak control forces to the system [125, 126]. This leads to small values of the KL divergence in (23, 27) and hence to small values of the rate function.

In fact, the nomenclature of local equilibrium may be slightly misleading in this context, in that the same argument may be applied to systems with non-equilibrium steady states, such as the active-matter system of Figure 7. In that case, the same hydrodynamic argument shows that  $\mathcal{I}(\sigma) = 0$  for a finite range of  $\sigma$  between 0 and  $\langle \sigma \rangle$ ; see [72]. However, this argument relies on the existence of a hydrodynamic theory for this active system where the only relevant field is the density – general conditions for this to hold in fluids with non-equilibrium steady states have not yet been established.

## 7 Outlook

This article has illustrated some aspects of the rich phenomenology of large deviations of time-averaged quantities. The focus has been on the behaviour in large systems, with many interacting degrees of freedom. In particular, on taking the system size  $N \rightarrow \infty$ , rate functions can develop singular behaviour. These singularities – which can be interpreted as dynamical phase transitions – happen when the mechanism for large fluctuations differs qualitatively from the typical behaviour. The main examples that we have considered are (i) the appearance of ferromagnetic order in a  $1d$  Ising model [30]; (ii) the existence of an inactive state in the FA model [17, 69, 96]; (iii) collective motion and arrested phases in an active matter system [72]; (iv) phase separation and hyperuniformity in the SSEP [73, 122].

We emphasise that these examples are illustrative and we have not attempted a comprehensive review. Among the things that have not been discussed are the recent development of large deviations at level-2.5 [127, 128], which can be interpreted as a more detailed fluctuation theory from which the main results of Section 2.1 can be derived by the contraction principle, see [47]. This theory also allows derivation of thermodynamic uncertainty principles, which are general bounds on the fluctuations of currents, including variances and large deviations [33, 129]. In a similar vein, there are some indications that large deviation principles are built on an underlying geometrical structure [63, 130, 131], which has consequences for optimally-controlled processes.

Looking forward, we mention a few directions of ongoing research. This review has concentrated on theoretical results and their implications for qualitative behaviour (such as how rate function scale with system size  $N$ ). However, numerical results have also contributed strongly to large-deviation research. Building on earlier studies [18, 81, 132, 133], recent years have seen renewed interest in efficient and accurate computation of rate functions and SCGFs [68–71, 134–137].

The theoretical ideas presented here are also being adapted to new settings. For example, large-deviations of time-averaged quantities are increasingly discussed

in *open quantum systems* [28], mostly using operator approaches applied to density matrices. Generalisation of the level-2.5 and optimal control approaches are also being explored in that context [138]. Another direction of interest is *non-Markovian* models [51,53,139,140], which can be even richer than the Markovian cases considered here [80,141]. Overall, the field has many interesting open questions, and new methods are becoming available, in order to address them. This makes us optimistic about future progress.

My understanding of large deviation theory and its applications has been shaped by many discussions and collaborations, and I am grateful to many people for their encouragement, advice, and patient explanations. Special thanks go to Juan P. Garrahan, Peter Sollich, Fred van Wijland, Vivien Lecomte, Hugo Touchette, Freddy Bouchet, Rosemary Harris, Lester Hedges, Mike Evans, Takahiro Nemoto, Paddy Royall, Todd Gingrich, David Limmer, Steve Whitelam, Johannes Zimmer, and Marcus Kaiser. I would also like to acknowledge fruitful discussions with David Chandler, whose advice and enthusiastic support were extremely important to me.

**Publisher's Note** The EPJ Publishers remain neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Appendix A: Optimal control calculation

### A.1 Equivalence of optimal-control problem and eigenproblems for a stochastic differential equation

For large deviations of  $\bar{b}_\tau$  in the model of (30), we show that using the controlled model (32) with (26) and maximising over  $\phi$  is equivalent to solving the eigenproblem (31). Using the theory of path integrals with Stratonovich convention, we write the path probability<sup>23</sup> for (30) as

$$P_\tau(\mathbf{C}) \propto \pi(\mathcal{C}_0) \exp \left( - \int_0^\tau \frac{|\partial_t \mathcal{C}_t - v(\mathcal{C}_t)|^2}{4} + \frac{\nabla \cdot v(\mathcal{C}_t)}{2} dt \right). \quad (\text{A.1})$$

where  $\pi$  is the probability of the initial condition. A similar expression holds for the controlled process (32): we assume for simplicity that this process has the same initial distribution  $\pi$  as the original process, although this

<sup>23</sup> The integral  $\int_0^\tau |\partial_t \mathcal{C}_t|^2 dt$  that appears in (A.1) is not mathematically well-defined for processes like (30). This may be resolved by defining  $P_\tau(\mathbf{C})$  as a density with respect to the path-measure for a Brownian motion (see for example [105]), or by considering  $P(\mathbf{C})$  to be the probability density of an explicitly time-discretised trajectory. In either case one arrives at the same result in (A.2), which is well-defined and unambiguous.

assumption is easily relaxed. In order to apply (26) we compute

$$\begin{aligned} \log \frac{P_\tau^{\text{con}}(\mathbf{C})}{P_\tau(\mathbf{C})} &= -\frac{1}{2} \int_0^\tau \nabla \phi(\mathcal{C}_t) \circ d\mathcal{C}_t \\ &+ \frac{1}{2} \int_0^\tau \nabla \phi(\mathcal{C}_t) \cdot v(\mathcal{C}_t) + \nabla^2 \phi(\mathcal{C}_t) - \frac{1}{2} |\nabla \phi(\mathcal{C}_t)|^2 dt \end{aligned} \quad (\text{A.2})$$

where the  $\circ$  indicates a Stratonovich product. This gives an explicit expression for  $\bar{g}_\tau$  in (28). The integral in the first line can be evaluated as  $[\phi(\mathcal{C}_\tau) - \phi(\mathcal{C}_0)]$ .

To apply (26) we require the average of (A.2), as  $\tau \rightarrow \infty$ . Ergodicity of the controlled process allows us to replace averages of time integrals by averages with respect to the steady-state distribution, which we denote by  $\mu$ . So (26) becomes

$$\psi(s) \geq \frac{1}{2} \int \left[ -2sb - \nabla \phi \cdot v - \nabla^2 \phi + \frac{1}{2} |\nabla \phi|^2 \right] \mu d\mathcal{C}. \quad (\text{A.3})$$

It is not necessary to compute  $\mu$  explicitly, one uses instead the Fokker-Planck equation for the controlled process to show that it solves

$$\nabla \cdot [v\mu - \mu \nabla \phi] = \nabla^2 \mu \quad (\text{A.4})$$

and one also has  $\int \mu(\mathcal{C}) d\mathcal{C} = 1$ . These are two constraints that can be implemented by Lagrange multipliers: we are left to find an extremum of

$$\begin{aligned} &\frac{1}{2} \int \left[ -2sb - \nabla \phi \cdot v - \nabla^2 \phi + \frac{1}{2} |\nabla \phi|^2 \right] \mu d\mathcal{C} \\ &+ \frac{1}{2} \int [\nabla \lambda \cdot (v\mu - \mu \nabla \phi - \nabla \mu) + 2\gamma \mu] d\mathcal{C} \end{aligned} \quad (\text{A.5})$$

where the functional Lagrange multiplier  $\lambda$  enforces (A.4) while  $\gamma$  enforces normalisation of  $\mu$ . A short calculation shows that the extremum occurs for  $\lambda = 0$ , and is characterised by

$$-\nabla \phi \cdot v - \nabla^2 \phi + \frac{1}{2} |\nabla \phi|^2 - 2sb = 2\gamma. \quad (\text{A.6})$$

This is an example of a Hamilton-Jacobi equation (or a Hamilton-Jacobi-Bellman equation). Using it with (A.3) shows that solutions of the variational problem have  $\gamma \leq \psi(s)$ . Moreover, writing  $\phi = -2 \log \mathcal{F}$  yields

$$\begin{aligned} \gamma \mathcal{F} &= \nabla^2 \mathcal{F} + v \cdot \nabla \mathcal{F} - sb \mathcal{F} \\ &= \mathcal{W}_s^\dagger \mathcal{F} \end{aligned} \quad (\text{A.7})$$

where the second line follows from the expression for  $\mathcal{W}_s$  given in (31). This is an eigenfunction equation for the operator  $\mathcal{W}_s^\dagger$ , and  $\gamma$  is the associated eigenvalue.

The optimal bound on  $\psi$  is obtained by taking the largest available solution for  $\gamma$ , which is therefore the largest eigenvalue of  $\mathcal{W}_s^\dagger$  – this is equal to  $\psi(s)$ , by (31). It

follows that (A.3) is an equality if one takes the (optimal) control potential  $\phi = -2\log \mathcal{F}$  where  $\mathcal{F}$  is the relevant eigenfunction of  $\mathcal{W}_s^\dagger$ .

### A.2 Example: large deviations of squared displacement in an Ornstein-Uhlenbeck process

To illustrate this general discussion, we analyse the specific case of a one-dimensional Ornstein-Uhlenbeck process, which is

$$dx_t = -\omega x_t dt + \sqrt{2}dW_t \tag{A.8}$$

where  $x_t$  is a real number. The force  $-\omega x$  is the gradient of a potential  $\frac{1}{2}\omega x^2$ . Large deviations for this process have been discussed previously in several contexts, for example [54,142,143]. We consider large deviations of  $\bar{b}_\tau = \tau^{-1} \int_0^\tau x_t^2 dt$  so  $b(x_t) = x_t^2$ . The tilted generator of (31) is  $\mathcal{W}_s$ , which acts on probability density functions  $P$  as

$$\mathcal{W}_s P = \nabla \cdot (\omega x P + \nabla P) - s x^2 P. \tag{A.9}$$

Its largest eigenvalue is the SCGF

$$\psi(s) = (\omega/2) - \sqrt{s + (\omega^2/4)}. \tag{A.10}$$

This result is valid for  $s > -\omega^2/4$ , otherwise the spectrum of  $\mathcal{W}_s$  is not bounded from below, this is linked to the behaviour of the rate function, as we discuss below. The associated eigenfunction is

$$P_{\text{end}}(x|s) = \sqrt{\frac{\omega - \psi(s)}{2\pi}} \exp\left[-\frac{x^2(\omega - \psi(s))}{2}\right]. \tag{A.11}$$

Note that  $\psi(s) \leq (\omega/2)$  so this is a normalised probability density function, for  $\psi(s) = 0$  it reduces to the steady-state distribution of (A.8). Using (A.10) with (8), the rate function for  $\bar{b}_\tau$  is

$$I(b) = \frac{(b\omega - 1)^2}{4b}. \tag{A.12}$$

For the original OU process one has  $\langle b \rangle = (1/\omega)$ , so  $I(\langle b \rangle) = 0$ , as required. The fact that the SCGF only exists for  $s > -\omega^2/4$  reflects the fact that the rate function grows linearly as  $b \rightarrow \infty$ , where it scales as  $I(b) \sim b\omega^2/4$ . Hence its Legendre transform  $\psi(s)$  only exists if  $s$  is sufficiently large. The solution of the backward Fokker-Planck equation (A.7) for this process is

$$\mathcal{F}(x|s) = \exp\left(\frac{x^2\psi(s)}{2}\right). \tag{A.13}$$

From this we infer (using  $\phi = -2\log \mathcal{F}$ ) that the optimal control potential  $\phi = -\psi(s)x^2$ . For  $s > 0$  then  $\psi(s) < 0$  and the control potential results in an additional confining force that reduces the typical value of  $x_t^2$ .

To verify that these results are consistent with the optimal-control theory, recall that the controlled process

in this case is

$$dx_t = -(\omega x_t + \nabla\phi(x_t))dt + \sqrt{2}dW_t. \tag{A.14}$$

Identifying  $v = -\omega x$  then (A.6) becomes

$$\omega x \nabla\phi - \nabla^2\phi + \frac{1}{2}|\nabla\phi|^2 - 2s x^2 = 2\gamma. \tag{A.15}$$

This equation can be solved by taking  $\gamma = \psi(s)$  [as given in (A.10)] and  $\phi = -\psi(s)x^2$ . This is indeed the optimal control potential  $\phi$  that maximises the bound in (A.3), which then becomes an equality.

To understand this result in a more intuitive way, it is useful to restrict to a controlled process with a quadratic control potential  $\phi(x) = \frac{1}{2}u x^2$ , where  $u$  is a variational parameter. Then (A.2) becomes

$$\log \frac{P_\tau^{\text{con}}(\mathcal{C})}{P_\tau(\mathcal{C})} = \frac{u}{4}[x_0^2 - x_\tau^2] + \frac{1}{2} \int_0^\tau \left[-\omega u x_t^2 + u - \frac{u^2 x_t^2}{2}\right] dt \tag{A.16}$$

The controlled process describes motion in a potential  $\frac{1}{2}(\omega + u)x^2$  so  $\langle x_t^2 \rangle_{\text{con}} = 1/(\omega + u)$  and one obtains by (24)

$$\lim_{\tau \rightarrow \infty} \frac{1}{\tau} \mathcal{D}(P_\tau^{\text{con}} || P_\tau) = \frac{u^2}{4(\omega + u)}. \tag{A.17}$$

This quantity measures how different is the controlled process from the original OU process. For  $u = 0$  the two processes coincide and the KL divergence is zero.

Moreover, this controlled process has  $\langle b \rangle_{\text{con}} = 1/(\omega + u)$  so it can be used with (27), as long as one takes  $b = \langle b \rangle_{\text{con}}$ , that is  $u = b^{-1} - \omega$ . The result is a bound on the rate function

$$I(b) \leq \frac{b}{4} (\omega - b^{-1})^2. \tag{A.18}$$

Comparison with (A.12) shows that this variational result holds as an equality. This occurs because the ansatz of a quadratic control potential is sufficient to capture the optimal control. It follows that the optimally-controlled dynamics for fluctuations with  $\bar{b}_\tau = b$  is simply

$$dx_t = -(x_t/b)dt + \sqrt{2}dW_t. \tag{A.19}$$

The SCGF can also be obtained by using (A.17) with (26), or equivalently by using (A.18) with (6).

The conclusion of this analysis for the OU process (A.8) is that large deviations of the time-average of  $x_t^2$  occur by trajectories that are representative of a similar (controlled) OU process, in which only the parameter  $\omega$  is modified. This parameter governs the size of the restoring force in (A.8) and hence the typical value of  $\bar{b}_\tau$ .

## Appendix B: Glauber-Ising model

This section summarises some results for large deviations of the time-integrated energy in the one-dimensional

Glauber-Ising model. This situation was analysed in [30]. Here we summarise the results and we also correct two small errors in that analysis. Details of the corrected analysis are given in [144].

The Glauber-Ising chain has  $N$  spins  $\sigma_i = \pm 1$ . The energy is  $E = -\frac{1}{2} \sum_i \sigma_i \sigma_{i+1}$  with periodic boundaries. Spin  $i$  flips with rate  $1/(1 + e^{\beta h_i \sigma_i})$  where  $\beta$  is the inverse temperature, and  $h_i = (\sigma_{i-1} + \sigma_{i+1})$  is the local field on site  $i$ , such that  $h_i \sigma_i$  is the change in energy on flipping spin  $i$ . The model can be analysed by writing  $\tau_i = \sigma_i \sigma_{i+1}$  so that  $\tau_i = -1$  if there is a domain wall between spin  $i$  and spin  $i + 1$ , and  $\tau_i = +1$  otherwise. Then the energy is  $-\sum_i \tau_i/2$ . The domain walls can be interpreted as particles that evolve according to a reaction-diffusion dynamics. Since the system has periodic boundaries then it is important to note that the number of these domain walls is always even.

In the domain-wall representation, the operator  $\mathcal{W}_s$  can be represented in terms of Pauli matrices. The dependence of the model on temperature is incorporated through a parameter  $\lambda = 2/(1 + e^{2\beta}) \leq 1$ . As stated in [73], the operator  $\mathcal{W}_s$  can be diagonalised using a Jordan-Wigner transformation. Some details of a similar computation are given in [145], where it is emphasised that the Jordan-Wigner transformation requires some care with the periodic boundary conditions. Using that the number of domain walls is always even the relevant operator can be diagonalised in a Fourier basis as

$$\mathcal{W}_s = \frac{1}{2} \sum_q [\Omega_q (\beta_q^\dagger \beta_q - \beta_q \beta_q^\dagger) - 1] \quad (\text{B.1})$$

where the sum runs over wavevectors  $q$  in the first Brillouin zone (see below),  $\beta_q$  and  $\beta_q^\dagger$  are fermionic creation and annihilation operators, and

$$\Omega_q = \sqrt{(1 - \cos q + s - \lambda)^2 + \lambda(2 - \lambda) \sin^2 q}. \quad (\text{B.2})$$

Compared with [30], we have corrected a factor of two in (B.1). In addition, careful treatment of the interplay between periodic boundaries and the Jordan-Wigner transformation requires that the wavevectors  $q$  in (B.1) are  $q = (2m + 1)\pi/N$  with  $m = 0, 1, 2, \dots, L - 1$  so that  $qN/\pi$  is an odd integer [145], contrary to [30]. Deriving this result uses explicitly that the number of domain walls is even. The largest eigenvalue of  $\mathcal{W}_s$  is then

$$\psi_N(s) = \frac{1}{2} \sum_q (\Omega_q - 1). \quad (\text{B.3})$$

The corresponding eigenvector  $|0\rangle$  obeys  $\beta_q^\dagger |0\rangle = 0$  for all  $q$ . For  $s = 0$  it can be checked that  $\Omega_q = [1 + (1 - \lambda) \cos q]$  so  $\psi(0) = 0$ , as required.

The function  $\psi_N(s)$  is analytic and convex in  $s$ , as it must be because the model is in the class of Section 2.2. Taking the large- $N$  limit, the free energy per

site of (42) is

$$\mathcal{G}(s) = \frac{1}{2} \int_{-\pi}^{\pi} (\Omega_q - 1) \frac{dq}{2\pi}. \quad (\text{B.4})$$

This function has a singularity in its second derivative at the phase transition point  $s = \lambda$ , at which  $\Omega_q \rightarrow 0$  for small- $q$ . This is a dynamical phase transition model and the system is ferromagnetic for  $s > \lambda$ .

## References

1. S.R. de Groot, P. Mazur, *Non-equilibrium thermodynamics* (Dover Publications, New York, 1984)
2. L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, C. Landim, *Rev. Mod. Phys.* **87**, 593 (2015)
3. U. Seifert, *Rep. Prog. Phys.* **75**, 126001 (2012)
4. A. Cavagna, *Phys. Rep.* **476**, 51 (2009)
5. S. Auer, D. Frenkel, *Nature* **409**, 1020 (2001)
6. S. Whitelam, R.L. Jack, *Ann. Rev. Phys. Chem.* **66**, 143 (2015)
7. D. Ruelle, *Phys. Today* **57**(5), 48 (2004)
8. G. Gallavotti, E.G.D. Cohen, *J. Stat. Phys.* **80**, 931 (1995)
9. P. Gaspard, J.R. Dorfman, *Phys. Rev. E* **52**, 3525 (1995)
10. F. den Hollander, *Large deviations* (American Mathematical Society, Providence, RI, 2000)
11. G.L. Eyink, *Phys. Rev. E* **54**, 3419 (1996)
12. J. L. Lebowitz, H. Spohn, *J. Stat. Phys.* **95**, 333 (1999)
13. B. Derrida, J.L. Lebowitz, *Phys. Rev. Lett.* **80**, 209 (1998)
14. L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, C. Landim, *J. Stat. Phys.* **107**, 635 (2002)
15. T. Bodineau, B. Derrida, *Phys. Rev. Lett.* **92**, 180601 (2004)
16. B. Derrida, *J. Stat. Mech.* **2007**, P07023 (2007)
17. J.P. Garrahan, R.L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, F. van Wijland, *Phys. Rev. Lett.* **98**, 195702 (2007)
18. L.O. Hedges, R.L. Jack, J.P. Garrahan, D. Chandler, *Science* **323**, 1309 (2009)
19. T. Speck, D. Chandler, *J. Chem. Phys.* **136**, 184509 (2012)
20. T. Speck, A. Malins, C.P. Royall, *Phys. Rev. Lett.* **109**, 195703 (2012)
21. R. Pinchaipat, M. Campo, F. Turci, J.E. Hallett, T. Speck, C.P. Royall, *Phys. Rev. Lett.* **119**, 028004 (2017)
22. P.I. Hurtado, P.L. Garrido, *Phys. Rev. Lett.* **102**, 250601 (2009)
23. V. Lecomte, A. Imparato, F. van Wijland, *Prog. Theor. Phys. Suppl.* **184**, 276 (2010)
24. U. Ray, D.T. Limmer, *Phys. Rev. B* **100**, 241409 (2019)
25. J.K. Weber, R.L. Jack, C.R. Schwantes, V.S. Pande, *Biophys. J.* **107**, 974 (2014)
26. A.S.J.S. Mey, P.L. Geissler, J.P. Garrahan, *Phys. Rev. E* **89**, 032109 (2014)
27. F. Ragone, J. Wouters, F. Bouchet, *Proc. Natl. Acad. Sci. USA* **115**, 24 (2018)
28. J.P. Garrahan, I. Lesanovsky, *Phys. Rev. Lett.* **104**, 160601 (2010)
29. H. Touchette, *Phys. Rep.* **478**, 1 (2009)



30. R.L. Jack, P. Sollich, Prog. Theor. Phys. Suppl. **184**, 304 (2010)
31. E.S. Loscar, A.S. J.S. Mey, J.P. Garrahan, J. Stat. Mech. **2011**, P12011 (2011)
32. R.L. Jack, L.O. Hedges, J.P. Garrahan, D. Chandler, Phys. Rev. Lett. **107**, 275702 (2011)
33. T.R. Gingrich, J.M. Horowitz, N. Perunov, J.L. England, Phys. Rev. Lett. **116**, 120601 (2016)
34. M.D. Donsker, S.R.S. Varadhan, Comm. Pure Appl. Math **28**, 1 (1975)
35. M.D. Donsker, S.R.S. Varadhan, Comm. Pure Appl. Math **28**, 279 (1975)
36. M.D. Donsker, S.R.S. Varadhan, Comm. Pure Appl. Math **29**, 389 (1976)
37. M.D. Donsker, S.R.S. Varadhan, Comm. Pure Appl. Math **36**, 183 (1983)
38. P. Dupuis, R.S. Ellis, *A weak convergence Approach to the Theory of Large Deviations* (Wiley, Hoboken, New Jersey, 1997)
39. L.-H. Gwa, H. Spohn, Phys. Rev. Lett. **68**, 725 (1992)
40. P. Gaspard, *Chaos, Scattering and Statistical Mechanics* (Cambridge University Press, Cambridge, 1998)
41. J.P. Eckmann, D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985)
42. D. Ruelle, *Thermodynamic Formalism: The Mathematical Structure of Equilibrium Statistical Mechanics*, 2nd edn., (Cambridge University Press, Cambridge, 2004)
43. C. Kipnis, C. Landim, *Scaling Limits of Interacting Particle System* (Springer, Berlin/Heidelberg, 1999)
44. R.S. Ellis, *Entropy, large deviations and statistical mechanics* (Springer, New York, 1985)
45. A.C.D. van Enter, R. Fernandez, A.D. Sokal, J. Stat. Phys. **72**, 879 (1993)
46. V. Lecomte, C. Appert-Rolland, F. van Wijland, J. Stat. Phys. **127**, 51 (2007)
47. R. Ch  trite, H. Touchette, Ann. Henri Poincar   **16**, 2005 (2015)
48. R. Ch  trite, H. Touchette, J. Stat. Mech. **2015**, P12001 (2015)
49. J.P. Garrahan, R.L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, F. van Wijland, J. Phys. A **42**, 075007 (2009)
50. R.S. Ellis, Scand. Actuarial J. **1995**, 97 (1995)
51. R.J. Harris, H. Touchette, J. Phys. A **42**, 342001 (2009)
52. P.L. Krapivsky, K. Mallick, T. Sadhu, Phys. Rev. Lett. **113**, 078101 (2014)
53. R.J. Harris, J. Stat. Mech. **2015**, P07021 (2015)
54. D. Nickelsen, H. Touchette, Phys. Rev. Lett. **121**, 090602 (2018)
55. A. Dechant, E. Lutz, D.A. Kessler, E. Barkai, Phys. Rev. Lett. **107**, 240603 (2011)
56. R. Ch  trite, H. Touchette, Phys. Rev. Lett. **111**, 120601 (2013)
57. T. Bodineau, B. Derrida, V. Lecomte, F. van Wijland, J. Stat. Phys. **133**, 1013 (2008)
58. Y.S. Elmatad, R.L. Jack, D. Chandler, J.P. Garrahan, Proc. Natl. Acad. Sci. USA **107**, 12793 (2010)
59. M.C. Ba  nuls, J.P. Garrahan, Phys. Rev. Lett. **123**, 200601 (2019)
60. D.P. Bertsekas, in *Dynamic programming and optimal control*, 3rd edn. (Athena Scientific, New Delhi, India, 2005), Vol. 1
61. W.H. Fleming, in *Recent Mathematical Methods in Dynamic Programming*, edited by I.C. Dolcetta, W.H. Fleming, T. Zolezzi (Springer Berlin Heidelberg, Berlin, Heidelberg, 1985) pp. 52–66
62. R.L. Jack, P. Sollich, Eur. Phys. J.: Special Topics **224**, 2351 (2015)
63. C. Maes, K. Netocny, Europhys. Lett. **82**, 30003 (2008)
64. A. Simha, R.M.L. Evans, A. Baule, Phys. Rev. E **77**, 031117 (2008)
65. D. Simon, J. Stat. Mech. **2009**, P07017 (2009)
66. H. Falk, Am. J. Phys. **38**, 858 (1970)
67. T. Nemoto, S.-i. Sasa, Phys. Rev. E **84**, 061113 (2011)
68. T. Nemoto, F. Bouchet, R.L. Jack, V. Lecomte, Phys. Rev. E **93**, 062123 (2016)
69. T. Nemoto, R.L. Jack, V. Lecomte, Phys. Rev. Lett. **118**, 115702 (2017)
70. U. Ray, G.K.-L. Chan, D.T. Limmer, Phys. Rev. Lett. **120**, 210602 (2018)
71. T. Brewer, S.R. Clark, R. Bradford, R.L. Jack, J. Stat. Mech. **2018**, 053204 (2018)
72. T. Nemoto, E. Fodor, M.E. Cates, R.L. Jack, J. Tailleur, Phys. Rev. E **99**, 022605 (2019)
73. R.L. Jack, I.R. Thompson, P. Sollich, Phys. Rev. Lett. **114**, 060601 (2015)
74. R.J. Harris, A. R  kos, G.M. Sch  tz, J. Stat. Mech. **2005**, P08003 (2005)
75. T. Bodineau, B. Derrida, Phys. Rev. E **72**, 066110 (2005)
76. R.S. MacKay, Nonlinearity **21**, T273 (2008)
77. G. Bunin, Y. Kafri, D. Podolsky, J. Stat. Mech. **2012**, L10001 (2012)
78. S. Vaikuntanathan, T.R. Gingrich, P.L. Geissler, Phys. Rev. E **89**, 062108 (2014)
79. P.T. Nyawo, H. Touchette, EPL **116**, 50009 (2016)
80. R.L. Jack, Phys. Rev. E **100**, 012140 (2019)
81. M. Merolle, J.P. Garrahan, D. Chandler, Proc. Natl. Acad. Sci. USA **102**, 10837 (2005)
82. R.L. Jack, J.P. Garrahan, D. Chandler, J. Chem. Phys. **125**, 184509 (2006)
83. C. Maes, J. Stat. Phys. **95**, 367 (1999)
84. S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, Cambridge, 2000)
85. R.L. Jack, T. Nemoto, V. Lecomte, J. Stat. Mech. [in press] (2020), [arXiv:1912.02465](https://arxiv.org/abs/1912.02465)
86. C. Borgs, R. Kotecky, J. Stat. Phys. **61**, 79 (1990)
87. V. Privman, M.E. Fisher, J. Stat. Phys. **33**, 385 (1983)
88. L. Berthier, G. Biroli, Rev. Mod. Phys. **83**, 587 (2011)
89. D. Chandler, J.P. Garrahan, Ann. Rev. Phys. Chem. **61**, 191 (2010)
90. E. Pitard, V. Lecomte, F. van Wijland, Europhys. Lett. **96**, 56002 (2011)
91. C.J. Fullerton, R.L. Jack, J. Chem. Phys. **138**, 224506 (2013)
92. F. Turci, C.P. Royall, T. Speck, Phys. Rev. X **7**, 031028 (2017)
93. B. Abou, R. Colin, V. Lecomte, E. Pitard, F. van Wijland, J. Chem. Phys. **148**, 164502 (2018)
94. R.L. Jack, J.P. Garrahan, Phys. Rev. E **81**, 011111 (2010)
95. G.H. Fredrickson, H.C. Andersen, Phys. Rev. Lett. **53**, 1244 (1984)
96. T. Bodineau, C. Toninelli, Commun. Math. Phys. **311**, 357 (2012)

97. T. Bodineau, V. Lecomte, C. Toninelli, *J. Stat. Phys.* **147**, 1 (2012)
98. R.L. Jack, P. Sollich, *J. Phys. A* **47**, 015003 (2014)
99. T. Nemoto, V. Lecomte, S.-i. Sasa, F. van Wijland, *J. Stat. Mech.* **2014**, P10001 (2014)
100. L. Berthier, M.D. Ediger, *Phys. Today* **69**(1), 40 (2016)
101. B. Gaveau, L.S. Schulman, *J. Math. Phys.* **37**, 3897 (1996)
102. B. Gaveau, L.S. Schulman, *J. Math. Phys.* **39**, 1517 (1998)
103. G. Biroli, J. Kurchan, *Phys. Rev. E* **64**, 016101 (2001)
104. G.E. Crooks, *J. Stat. Phys.* **90**, 1481 (1998)
105. R.L. Jack, M. Kaiser, J. Zimmer, *Entropy* **19**, 562 (2017)
106. D.J. Evans, E.G.D. Cohen, G.P. Morriss, *Phys. Rev. Lett.* **71**, 2401 (1993)
107. K. Sekimoto, *Prog. Theor. Phys. Suppl.* **130**, 17 (1998)
108. G.E. Crooks, *Phys. Rev. E* **61**, 2361 (2000)
109. M.V.S. Bonança, C. Jarzynski, *Phys. Rev. E* **93**, 022101 (2016)
110. F. Cagnetta, F. Corberi, G. Gonnella, A. Suma, *Phys. Rev. Lett.* **119**, 158002 (2017)
111. T. GrandPre, D.T. Limmer, *Phys. Rev. E* **98**, 060601 (2018)
112. É. Fodor, T. Nemoto, S. Vaikuntanathan, *New J. Phys.* **22**, 013052 (2020)
113. Y. Fily, M.C. Marchetti, *Phys. Rev. Lett.* **108**, 235702 (2012)
114. G.S. Redner, M.F. Hagan, A. Baskaran, *Phys. Rev. Lett.* **110**, 055701 (2013)
115. E. Fodor, C. Nardini, M.E. Cates, J. Tailleur, P. Visco, F. van Wijland, *Phys. Rev. Lett.* **117**, 038103 (2016)
116. P. Pietzonka, U. Seifert, *J. Phys. A* **51**, 01LT01 (2017)
117. D. Mandal, K. Klymko, M.R. DeWeese, *Phys. Rev. Lett.* **119**, 258001 (2017)
118. S. Shankar, M.C. Marchetti, *Phys. Rev. E* **98**, 020604 (2018)
119. P.I. Hurtado, C.P. Espigares, J.J. del Pozo, P.L. Garrido, *J. Stat. Phys.* **154**, 214 (2014)
120. Y. Baek, Y. Kafri, V. Lecomte, *Phys. Rev. Lett.* **118**, 030604 (2017)
121. N. Tizón-Escamilla, C. Pérez-Espigares, P.L. Garrido, P.I. Hurtado, *Phys. Rev. Lett.* **119**, 090602 (2017)
122. C. Appert-Rolland, B. Derrida, V. Lecomte, F. van Wijland, *Phys. Rev. E* **78**, 021122 (2008)
123. V. Lecomte, J.P. Garrahan, F. van Wijland, *J. Phys. A* **45**, 175001 (2012)
124. S. Torquato, F.H. Stillinger, *Phys. Rev. E* **68**, 041113 (2003)
125. J. Dolezal, R.L. Jack, *J. Stat. Mech.* **2019**, 123208 (2019)
126. A. Das, D.T. Limmer, *J. Chem. Phys.* **151**, 244123 (2019)
127. A.C. Barato, R. Chetrite, *J. Stat. Phys.* **160**, 1154 (2015)
128. L. Bertini, R. Chetrite, A. Faggionato, D. Gabrielli, *Ann. Henri Poincaré* **19**, 3197 (2018)
129. A.C. Barato, R. Chetrite, A. Faggionato, D. Gabrielli, *New J. Phys.* **20**, 103023 (2018)
130. A. Mielke, M.A. Peletier, D.R.M. Renger, *Potential Analysis* **41**, 1293 (2014)
131. M. Kaiser, R.L. Jack, J. Zimmer, *J. Stat. Phys.* **170**, 1019 (2018)
132. C. Giardinà, J. Kurchan, L. Peliti, *Phys. Rev. Lett.* **96**, 120603 (2006)
133. J. Tailleur, J. Kurchan, *Nature Phys.* **3**, 203 (2007)
134. U. Ray, G.K.-L. Chan, D.T. Limmer, *J. Chem. Phys.* **148**, 124120 (2018)
135. G. Ferré, H. Touchette, *J. Stat. Phys.* **172**, 1525 (2018)
136. D. Jacobson, S. Whitelam, *Phys. Rev. E* **100**, 052139 (2019)
137. C. Pérez-Espigares, P.I. Hurtado, *Chaos* **29**, 083106 (2019)
138. F. Carollo, R.L. Jack, J.P. Garrahan, *Phys. Rev. Lett.* **122**, 130605 (2019)
139. C. Maes, K. Netocny, B. Wynants, *J. Phys. A* **42**, 365002 (2009)
140. A. Faggionato, [arXiv:1709.05653](https://arxiv.org/abs/1709.05653) (2017)
141. S. Franchini, *Stoch. Process. Appl.* **127**, 3372 (2017)
142. S.N. Majumdar, A.J. Bray, *Phys. Rev. E* **65**, 051112 (2002)
143. H. Touchette, *Physica A* **504**, 5 (2018)
144. J. Guioth, R.L. Jack, [arXiv:2002.00905](https://arxiv.org/abs/2002.00905) (2020)
145. M.D. Grynberg, T.J. Newman, R.B. Stinchcombe, *Phys. Rev. E* **50**, 957 (1994)