

Large deviations and gradient flows

Stefan Adams* Nicolas Dirr† Mark Peletier‡ Johannes Zimmer§

January 24, 2012

Abstract

In recent work [1] we uncovered intriguing connections between Otto's characterisation of diffusion as entropic gradient flow [16] on one hand and large-deviation principles describing the microscopic picture (Brownian motion) on the other. In this paper, we sketch this connection, show how it generalises to a wider class of systems, and comment on consequences and implications.

Specifically, we connect macroscopic gradient flows with large deviation principles, and point out the potential of a bigger picture emerging: we indicate that in some non-equilibrium situations, entropies and thermodynamic free energies can be derived via large deviation principles. The approach advocated here is different from the established hydrodynamic limit passage but extends a link that is well known in the equilibrium situation.

1 Introduction

For systems in equilibrium, it is well known that the roles of energy and entropy can be understood rigorously in terms of large-deviation principles. We describe two examples below. Recently, we showed how large-deviation principles also allow us to understand the role of entropy in a specific *non-equilibrium* system [1]: the large-deviation behaviour of a system of independent Brownian particles connects rigorously to the entropy gradient-flow structure of the diffusion equation. We explain this connection in Section 3.1.

The aim of this paper is to take this connection two steps further. The first step is to extend the connection of [1], which was studied in a discrete-time context, to the case of continuous time. The second step is to discuss a variety of examples that illustrates the breadth of this phenomenon, and suggest a general principle that might hold across a wide range of systems.

In equilibrium systems, the connection is as follows. Let X_i ($i = 1, 2, \dots$) be independent and identically distributed stochastic variables with distribution μ on a state space \mathcal{X} . We think of the X_i as positions of particles in the space \mathcal{X} , so that their concentration is given by the *empirical measure* $\rho_n := \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$. Sanov's theorem (e.g., [7, Sec. 6.2]) states that the random measure ρ_n satisfies the *large-deviation principle*

$$\text{Prob}(\rho_n \approx \rho) \sim \exp[-nI(\rho)], \quad \text{as } n \rightarrow \infty, \quad (1)$$

*Mathematics Institute, University of Warwick

†Cardiff School of Mathematics, Cardiff University

‡Department of Mathematics and Computer Sciences and Institute for Complex Molecular Systems, Technische Universiteit Eindhoven

§Department of Mathematical Sciences, University of Bath

where the *rate function* $I \geq 0$ is the *relative entropy* of ρ with respect to μ , which is

$$I(\rho) = H(\rho|\mu) := \begin{cases} \int f \log f d\mu & \text{if } \rho \ll \mu \text{ and } \rho = f\mu, \\ +\infty & \text{otherwise,} \end{cases}$$

This property illustrates how the relative entropy $H(\rho|\mu)$ characterises the probability of observing a state ρ : higher relative entropy means smaller probability, as described by (1). It also provides a rigorous version of the well-known thermodynamic principle that a system aims to maximise its entropy (which corresponds to minimising $H(\rho|\mu)$, since the physical entropy carries the opposite sign). For in the limit of large n , the characterisation (1) gives vanishing probability to all states ρ except those for which $I(\rho) = 0$; in other words, only the minimisers of I have non-vanishing probability.

This connection between entropy and large-deviation principles extends to systems involving energy. In the appendix we show, for instance, how coupling a system with energy E to a heat bath with temperature θ changes the rate functional I to the *free energy* $\mathcal{F}(\rho) := H(\rho|\mu) + (k\theta)^{-1}E(\rho) + \text{constant}$:

$$\text{Prob}(\rho_n \approx \rho) \sim \exp[-n\mathcal{F}(\rho)], \quad \text{as } n \rightarrow \infty. \quad (2)$$

In the same way as (1) explains why relative entropy is minimized, (2) explains why systems coupled to a heat bath minimize their free energy: when n is large, only states ρ with near-minimal free energy $\mathcal{F}(\rho)$ will have finite probability.

As mentioned above, the central aim of this paper is to show how this connection between entropy and free energies on one hand and large-deviation principles on the other extends into the realm of non-equilibrium systems. We restrict our focus to the important class of *gradient flows*, where this connection explains many aspects of these systems. Since the entropy appears as the driving force of the process, we will occasionally call this functional “energy” to conform with the standard terminology for gradient flows.

The general philosophy is illustrated by the diagram below.

$$\begin{array}{ccc} \begin{array}{c} \text{dynamic rate functional} \\ I \text{ or } I_h \end{array} & \xleftrightarrow{\text{this paper}} & \begin{array}{c} \text{gradient-flow structure} \\ J \text{ or } I_h \end{array} \\ \text{large-deviation principle} \uparrow & & \uparrow \\ n \rightarrow \infty & & \\ \text{stochastic } n\text{-particle system} & \xrightarrow[n \rightarrow \infty]{\text{continuum limit}} & \text{continuum evolution equation} \end{array} \quad (3)$$

The bottom row in this diagram is the classical connection between a stochastic n -particle system and its hydrodynamic limit: the typical case is that as $n \rightarrow \infty$, the particle system becomes deterministic, and the empirical measure of the particle system converges to the solution of the (deterministic) continuum equation. Note that this statement concerns only the typical behaviour of the particle system; large deviations are not captured.

In the left-hand column, a large-deviation principle characterises the behaviour in the limit $n \rightarrow \infty$ in a different manner, in terms of a functional I or I_h of the *time-dependent* system, as we shall see below. The right-hand column is the connection between an evolution equation and the corresponding gradient-flow structure, when it exists.

The central statement of this paper is the double-headed arrow at the top. It provides a connection between representations with more information on both sides: on the left-hand side, the rate functional contains more information than just the most probable behaviour,

and on the right-hand side, the gradient-flow structure is an additional structure on top of the equation itself.

In the following sections, we illustrate the double-headed arrow in a number of concrete examples, first in the discrete-time approximation (Section 3) and then in continuous time (Section 4). Section 5 generalises the argument to non-quadratic dissipations. Since the implications of this connection are best appreciated once one has an overview of the breadth of the phenomenon, we postpone most of the discussion of the consequences to Section 6.

The mathematical results described in this paper are not new, and mostly due to other authors, such as Freidlin & Wentzell [10], Dawson & Gärtner [5, 6], Feng & Kurtz [9], Kipnis, Olla, & Varadhan [12] and others. Instead, we see the novelty of this paper in extracting from these results the suggestion of a general principle connecting the broad class of gradient flows with large deviations of stochastic processes. A particularly interesting aspect of this connection is that thermodynamic quantities are derived in a non-equilibrium context.

2 The Wasserstein metric

Much of this paper centres on the Wasserstein metric and Wasserstein gradient flows. The (quadratic) *Wasserstein distance* between two probability measures ρ_0 and ρ_1 with finite second moments is [18]

$$d(\rho_0, \rho_1)^2 = \inf_q \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 q(dx dy), \quad (4)$$

where the infimum is taken over all q with marginals ρ_0 and ρ_1 , i.e., over all q satisfying

$$\text{for any } A \subset \mathbb{R}^d, \quad q(A \times \mathbb{R}^d) = \rho_0(A) \quad \text{and} \quad q(\mathbb{R}^d \times A) = \rho_1(A).$$

We also need an incremental version of the Wasserstein distance. The Brenier-Benamou formula [3] gives an alternative formulation of d as an infimum of curves of measures $t \mapsto \rho(t)$ such that $\rho(0) = \rho_0$ and $\rho(1) = \rho_1$:

$$d(\rho_0, \rho_1)^2 = \inf_{\rho: [0,1] \rightarrow \mathcal{M}_1(\mathbb{R}^d)} \int_0^1 \|\partial_t \rho(t)\|_{\rho(t),*}^2 dt. \quad (5)$$

Here the local norm $\|\cdot\|_{\rho,*}$ at a given point ρ is derived from an inner product (a local metric tensor) formally given by

$$(s_1, s_2)_{\rho,*} := \int_{\mathbb{R}^d} \rho(x) \nabla p_1(x) \cdot \nabla p_2(x) dx, \quad (6)$$

where ∇ is the usual gradient in \mathbb{R}^d , and the p_i solve the equation $\text{div}(\rho \nabla p_i) = s_i$ in \mathbb{R}^d (see [5, 13] or [9, Sec. 9.4] for a rigorous definition).

A *Wasserstein gradient flow* is a gradient flow of an energy \mathcal{E} with respect to the Wasserstein metric structure. A curve of measures $t \mapsto \rho(t)$ is a solution of such a gradient-flow equation if its time derivative $\partial_t \rho$, in the sense of distributions, satisfies

$$(\partial_t \rho(t), s_2)_{\rho(t),*} = - \int_{\mathbb{R}^d} \frac{\delta \mathcal{E}}{\delta \rho}(\rho(t)) s_2 dx \quad \text{for all } s_2 \text{ and all } t > 0, \quad (7)$$

where $\delta\mathcal{E}/\delta\rho$ is the variational derivative of \mathcal{E} . A straightforward calculation shows that this is equivalent to the equation

$$\partial_t\rho = \operatorname{div} \rho \nabla \left(\frac{\delta\mathcal{E}}{\delta\rho} \right). \quad (8)$$

By analogy with gradients in Riemannian geometry, this suggests to define the *Wasserstein gradient* of a functional \mathcal{E} as

$$\operatorname{grad}_W \mathcal{E}(\rho) := -\operatorname{div} \rho \nabla \left(\frac{\delta\mathcal{E}}{\delta\rho} \right). \quad (9)$$

Below we shall also use more general versions of this structure. Replacing ρ above by a general diffusion matrix $D(\rho)$, we define

$$(s_1, s_2)_{D(\rho),*} := \int_{\mathbb{R}^d} D(\rho(x)) \nabla p_1(x) \cdot \nabla p_2(x) dx, \quad \text{where} \quad s_i = \operatorname{div} D(\rho) \nabla p_i. \quad (10)$$

Repeating the construction above, it follows that the *D-Wasserstein gradient* of a functional \mathcal{E} is characterised by the equation

$$\partial_t\rho = \operatorname{div} D(\rho) \nabla \left(\frac{\delta\mathcal{E}}{\delta\rho} \right). \quad (11)$$

Gradient flows have natural time-discrete approximations, constructed in an iterative manner:

*For given approximation ρ_{k-1} at time $(k-1)h$, choose ρ_k at time kh
as minimiser of the functional $\rho \mapsto \frac{1}{2h} d(\rho, \rho^{k-1})^2 + \mathcal{E}(\rho)$.*

$$\rho \mapsto \frac{1}{2h} d(\rho, \rho^{k-1})^2 + \mathcal{E}(\rho). \quad (12)$$

This is essentially a backward-Euler discretisation, as can be recognised by comparing it with the \mathbb{R}^d -gradient-flow $\dot{x} = -\nabla E(x)$. For this equation the backward-Euler discretisation is constructed by solving

$$\frac{1}{h}(x_k - x_{k-1}) = -\nabla E(x_k),$$

for x_k , which is equivalent to minimising

$$x \mapsto \frac{1}{2h} |x - x_{k-1}|^2 + E(x). \quad (13)$$

Note the similarity between (13) and (12): in both expressions the first term measures the distance between old and new states, while the second term favours a reduction of the functional \mathcal{E} respectively E .

3 Discrete time

We can now formulate the first example.

3.1 A system of independent Brownian particles

We consider n independent Brownian particles $X_{n,i}(t)$ in \mathbb{R}^d , with deterministic initial positions $X_{n,i}(0) = x_{n,i}$, each hopping to a new position $X_{n,i}(h)$ at time $h > 0$ with a Gaussian probability with mean $x_{n,i}$ and variance¹ $2h$.

As in the equilibrium case discussed above, we describe this system by the empirical measure $\rho_n(t) := \frac{1}{n} \sum_{i=1}^n \delta_{X_{n,i}(t)}$ at a given time t , and we assume that the initial measure $\rho_n(0)$ converges to a given measure ρ^0 as $n \rightarrow \infty$. In the limit of large n , the probability of this jump process attaining any ρ^1 at time $t = h$ is again characterised in terms of a large-deviation principle,

$$\text{Prob}(\rho_n(h) \approx \rho^1) \approx \exp[-nI_h(\rho^1)], \quad (14)$$

where the rate functional I_h has an explicit expression that can be derived from Stirling's formula (see [1] for the expression; in [1], I_h is only the limit of a sequence of rate functionals, but can be shown to be a rate functional in its own right [14, 17]).

The main result of [1] is that

$$I_h \approx K_h \quad \text{as } h \rightarrow 0, \quad (15)$$

where

$$K_h(\rho^1; \rho^0) := \frac{1}{4h} d(\rho^0, \rho^1)^2 + \frac{1}{2} \text{Ent}(\rho^1) - \frac{1}{2} \text{Ent}(\rho^0). \quad (16)$$

Here d is the Wasserstein distance defined above, and

$$\text{Ent}(\rho) := H(\rho|\mathcal{L}) = \begin{cases} \int f \log f \, dx & \text{if } \rho \ll \mathcal{L} \text{ and } \rho = f\mathcal{L}, \\ +\infty & \text{otherwise,} \end{cases}$$

is the relative entropy of ρ with respect to the Lebesgue measure \mathcal{L} . The rigorous formulation of (15) is a Gamma-convergence result of I_h to K_h after both have been desingularised.

The functional K_h has the same form as the functional in (12), since the term $\text{Ent}(\rho^0)/2$ does not influence the minimisation with respect to ρ^1 . Therefore the time-discrete approximation that one constructs with this K_h is an approximation of the Wasserstein gradient flow of the entropy Ent , which is the diffusion equation [11]

$$\partial_t \rho = \Delta \rho \quad \text{in } \mathbb{R}^d. \quad (17)$$

This is the connection referred to above: the large-deviation behaviour of the system of particles is represented by the rate functional I_h , and this functional is asymptotically equal to the functional K_h that defines the gradient-flow formulation of the diffusion equation. The approximation result (15) therefore creates a link between the gradient-flow structure of the deterministic limit equation on one hand and the large-deviation behaviour of the system of particles on the other. The same result can be shown for Gaussian measures on the real line [8]. In the rest of this paper we shall see many more versions of such connections.

¹In this paper, we consider Brownian particles with generator Δ , rather than $(1/2)\Delta$, and therefore the transition kernel is $(4\pi h)^{-d/2} \exp -|x - y|^2/4h$.

Consequences While most of the discussion is deferred to Section 6, we mention here a few consequences of the fact (15) that the large-deviation rate functional I_h and the constructing functional K_h of the gradient flow are equal in the limit $h \rightarrow 0$.

First, the construction of a time-discrete approximation (12) to the diffusion equation (17) was motivated in [11] by analogy with the backward-Euler discretisation (13). This is an indirect and purely mathematical motivation, which explains neither the reason for the appearance of the entropy and the Wasserstein distance in K_h , nor the reason for minimising just this combination.

The connection between K_h and I_h , however, gives a direct motivation. By (14)–(15), $K_h(\rho; \rho^0)$ is a measure of the likelihood of observing a state ρ after time h . For large n , the characterisation (14) implies that only the global minimiser of I_h , and therefore of K_h , is observed with non-vanishing probability. The stochastic minimisation (14) of I_h thus becomes converted into an absolute minimisation of K_h .

Secondly, in the limit $h \rightarrow 0$, the *proof* that $I_h \approx K_h$ explains the origin of the two terms of K_h . The entropy arises from the indistinguishability of the particles after transforming to an empirical measure. The origin of the Wasserstein cost functional $|x - y|^2$ in (4) can be traced back to the exponent of the term $e^{-|x-y|^2/4h}$ in the Gaussian transition probability of the Brownian particles. We return to this issue in Section 6.

4 Continuous time

The construction in the previous section is discrete in time: the rate function I_h describes the probability distribution of the state $\rho_n(h)$ at time $h > 0$. A continuous-time large-deviation principle, where one considers deviations from a whole path of empirical measures for a fixed terminal time, provides a different kind of insight, and may be even closer to the gradient-flow formulation. We start with some preliminaries.

4.1 An alternative formulation of the gradient-flow structure

In a formal sense, Wasserstein gradient flows and many others can be written in the form

$$\partial_t \rho = -M_\rho \frac{\delta \mathcal{E}}{\delta \rho}, \quad (18)$$

where \mathcal{E} is the ‘energy’ functional driving the evolution, and M_ρ a ρ -dependent symmetric mapping². In the case of Wasserstein gradient flows, for instance,

$$M_\rho \xi = -\operatorname{div} \rho \nabla \xi,$$

as follows by comparing (8) with (18). Taking this case of Wasserstein gradient flow as an example, we shall encounter the equation (18) in a different form, connected to the functional J given by

$$J(\rho) := \mathcal{E}(\rho(T)) - \mathcal{E}(\rho(0)) + \frac{1}{2} \int_0^T \left[\|\partial_t \rho\|_{\rho,*}^2 + \left\| -\frac{\delta \mathcal{E}}{\delta \rho} \right\|_\rho^2 \right] dt, \quad (19)$$

²This way of writing the gradient flow highlights the fact that a gradient flow is an instance of a GENERIC evolution, in which the conservative evolution term is absent [15].

where

$$\|\xi\|_\rho^2 := \int_{\mathbb{R}^d} \xi M_\rho \xi \, dx = \int_{\mathbb{R}^d} \rho |\nabla \xi|^2$$

and the norm $\|\cdot\|_{\rho,*}^2$ is the norm defined in (6). The norms $\|\cdot\|_\rho$ and $\|\cdot\|_{\rho,*}$ are dual norms, and $\|\cdot\|_{\rho,*}$ has the alternative characterisation

$$\|s\|_{\rho,*} := \sup_{\xi \neq 0} \frac{\int_{\mathbb{R}^d} s \xi \, dx}{\|\xi\|_\rho}.$$

By writing the energy difference $\mathcal{E}(\rho(T)) - \mathcal{E}(\rho(0))$ as

$$\mathcal{E}(\rho(T)) - \mathcal{E}(\rho(0)) = \int_0^T \int_{\mathbb{R}^d} \frac{\delta \mathcal{E}}{\delta \rho} \partial_t \rho \, dx dt = \int_0^T \left(M_\rho \frac{\delta \mathcal{E}}{\delta \rho}, \partial_t \rho \right)_{\rho,*} dt,$$

using the inner product defined in (7), the functional J in (19) can now be written as

$$J(\rho) = \frac{1}{2} \int_0^T \left\| \partial_t \rho + M_\rho \frac{\delta \mathcal{E}}{\delta \rho} \right\|_{\rho,*}^2 dt.$$

This expression shows that J is non-negative. It also implies that if ρ satisfies $J(\rho) = 0$, then equation (18) holds at almost each time $0 < t < T$; therefore

$$\rho \text{ is a Wasserstein gradient flow of } \mathcal{E} \iff J(\rho) = 0. \quad (20)$$

In the examples of this paper, J is a large-deviation rate functional, and this equivalence is the connection between the large-deviation behaviour, given by J , and the gradient-flow structure of the limiting equation.

If we take for the operator M_ρ in (18) not the Wasserstein operator but a general operator, then we find a similar statement:

$$\rho \text{ is a solution of the } (\mathcal{E}, M_\rho)\text{-gradient-flow (18)} \iff J_M(\rho) = 0, \quad (21)$$

where

$$J_M(\rho) := \mathcal{E}(\rho(T)) - \mathcal{E}(\rho(0)) + \frac{1}{2} \int_0^T \left[\|\partial_t \rho\|_{M_\rho^{-1}}^2 + \left\| -\frac{\delta \mathcal{E}}{\delta \rho} \right\|_{M_\rho}^2 \right] dt, \quad (22)$$

and the two norms are defined, at least formally, by

$$\begin{aligned} \|\xi\|_{M_\rho}^2 &:= \int_{\mathbb{R}^d} \xi M_\rho \xi \, dx, \\ \|s\|_{M_\rho^{-1}} &:= \sup_{\xi \neq 0} \frac{\int_{\mathbb{R}^d} s \xi \, dx}{\|\xi\|_{M_\rho}} = \int_{\mathbb{R}^d} s M_\rho^{-1} s \, dx = \|M_\rho^{-1} s\|_{M_\rho}^2. \end{aligned}$$

We now discuss a number of examples.

4.2 Continuous-time large deviations for the diffusion equation

Taking the same system of particles as in Section 3.1, the continuous-time large-deviation principle for that system of Brownian particles is as follows. Fix a terminal time $T > 0$ and consider the whole path $[0, T] \rightarrow \mathcal{M}_1(\mathbb{R}^d)$ of empirical measures $[0, T] \ni t \mapsto \rho_n(t)$. Then the probability that the *entire curve* $\rho_n(\cdot)$ is close to some other $\rho(\cdot)$ is characterised as [5, 13] as a *pathwise large-deviation principle*,

$$\text{Prob}(\rho_n \approx \rho) \sim \exp[-nI(\rho)],$$

where now

$$I(\rho) := \frac{1}{2} \int_0^T \|\partial_t \rho - \Delta \rho\|_{\rho(t),*}^2 dt. \quad (23)$$

This rate function I has the structure of J in (19). Using the fact that

$$\Delta \rho = \text{div } \rho \nabla \left(\frac{\delta \text{Ent}}{\delta \rho} \right),$$

we find that

$$I(\rho) = \text{Ent}(\rho(T)) - \text{Ent}(\rho(0)) + \frac{1}{2} \int_0^T \left[\|\partial_t \rho\|_{\rho,*}^2 + \left\| -\frac{\delta \text{Ent}}{\delta \rho} \right\|_{\rho}^2 \right] dt.$$

Therefore the Entropy-Wasserstein gradient flow is connected to the large-deviation behaviour of a system of stochastic particles, in the sense of (20). We discuss this further in Section 6.

4.3 Diffusive particles with interactions

We extend the previous example by including interaction of the particles with a background potential Ψ and with each other via an interaction potential Φ , and modelled by Itô stochastic differential equations. Specifically, we take the microscopic system of n particles to be described by

$$dX_i(t) = -\nabla \Psi(X_i(t)) dt - \frac{1}{n} \sum_{j=1}^n \nabla \Phi(X_i(t) - X_j(t)) dt + \sqrt{2} dW_i(t), \quad (24)$$

where for each i , W_i is a Brownian motion in \mathbb{R}^d . The hydrodynamic limit of this system is the equation

$$\partial_t \rho = \Delta \rho + \text{div } \rho \nabla [\Psi + \rho * \Phi]. \quad (25)$$

The large-deviation rate functional describing fluctuations of the system is given by (see [9, Theorem 13.37], and also [5] for weakly interacting diffusive particle systems)

$$I(\rho) := \frac{1}{2} \int_0^T \left\| \partial_t \rho - \Delta \rho - \text{div } \rho \nabla [\Psi + \rho * \Phi] \right\|_{\rho,*}^2 dt, \quad (26)$$

which again can be written as

$$I(\rho) = \mathcal{F}(\rho(T)) - \mathcal{F}(\rho(0)) + \frac{1}{2} \int_0^T \left[\|\partial_t \rho\|_{\rho,*}^2 + \left\| \frac{\delta \mathcal{F}}{\delta \rho} \right\|_{\rho}^2 \right] dt,$$

where the free energy \mathcal{F} is given by the sum of entropy and potential energy,

$$\mathcal{F}(\rho) := \text{Ent}(\rho) + \int_{\mathbb{R}^d} \left[\rho \Psi + \frac{1}{2} \rho(\rho * \Phi) \right]. \quad (27)$$

Indeed equation (25) is the Wasserstein gradient flow of the functional \mathcal{F} .

4.4 The Symmetric Simple Exclusion Process

The diffusion equation (17) is the continuum limit for various stochastic processes, one of which is the system of Brownian particles described above. Here we briefly describe the symmetric simple exclusion process, which has the same limiting equation in a parabolic scaling. However, it has a different large-deviation behaviour, which gives rise to a different gradient flow.

Consider a periodic lattice $T_n = \{0, 1/n, 2/n, \dots, (n-1)/n\}$ and its continuum limit, the flat torus $T = \mathbb{R}/\mathbb{Z}$. Each lattice site contains zero or one particle; each particle attempts to jump from to a neighbouring site with rate $n^2/2$, and they succeed if the target site is empty. We define the configuration $\rho_n: T_n \rightarrow \{0, 1\}$ such that $\rho_n(k/n) = 1$ if there is a particle at site k/n , and zero otherwise. For this system the large deviations are characterised by the rate function [12]

$$I(\rho) := \frac{1}{2} \int_0^T \|\partial_t \rho - \partial_{xx} \rho\|_{\rho(1-\rho),*}^2 dt, \quad (28)$$

where the norm $\|\cdot\|_{\rho(1-\rho),*}$ is given by (10) with $D(\rho) = \rho(1-\rho)$. This functional can be written as

$$I(\rho) = \text{Ent}_{\text{mix}}(\rho(T)) - \text{Ent}_{\text{mix}}(\rho(0)) + \frac{1}{2} \int_0^T \left[\|\partial_t \rho\|_{\rho(1-\rho),*}^2 + \left\| -\frac{\delta \text{Ent}_{\text{mix}}}{\delta \rho} \right\|_{\rho(1-\rho)}^2 \right] dt,$$

where the mixing entropy Ent_{mix} is defined as

$$\text{Ent}_{\text{mix}}(\rho) := \int_{\mathbb{R}^d} [\rho \log \rho + (1-\rho) \log(1-\rho)].$$

This is true since $-\partial_{xx} \rho$ is the ' $\rho(1-\rho)$ '-Wasserstein gradient of Ent_{mix} , by

$$-\partial_{xx} \rho = -\partial_x \left(\rho(1-\rho) \partial_x \log \frac{\rho}{1-\rho} \right) = -\partial_x \left(\rho(1-\rho) \partial_x \frac{\delta \text{Ent}_{\text{mix}}}{\delta \rho}(\rho) \right)$$

(compare this to (11)). Therefore I is of the form (22), with operator

$$M_\rho \xi := \text{div } \rho(1-\rho) \nabla \xi,$$

and the equation $\partial_t \rho = \partial_{xx} \rho$ is (also) the gradient flow of Ent_{mix} with respect to this ' $\rho(1-\rho)$ '-Wasserstein structure $\|\cdot\|_{\rho(1-\rho),*}$.

5 Further generalisations

The arguments of the integrals in (5), (23), (26), and (28) are quadratic. This arises from a parabolic rescaling and the central limit theorem, and it leads to a gradient flow with a (formal) inner-product structure, or equivalently, to a linear operator M_ρ in (18). Other types of randomness lead to non-quadratic gradient-flow structures, as we now describe.

A close inspection of the arguments of Section 4.1 shows that they hinge on the inequality

$$\partial_t \mathcal{E}(\rho) \geq -\frac{1}{2} \|\partial_t \rho\|_{M_\rho^{-1}}^2 - \frac{1}{2} \left\| -\frac{\delta \mathcal{E}}{\delta \rho} \right\|_{M_\rho}^2,$$

together with the observation that equality holds if and only if $\partial_t \rho = -M_\rho \delta \mathcal{E} / \delta \rho$. This can be generalised by introducing a Legendre pair of convex functions ψ_ρ and ψ_ρ^* , where the subscript ρ serves to indicate that they may depend on ρ , in the same way as the operator M_ρ does; in this context, ψ_ρ , is often called dissipation potential. In terms of this pair we then derive that

$$\partial_t \mathcal{E}(\rho(t)) = \int \frac{\delta \mathcal{E}}{\delta \rho} \partial_t \rho \geq -\psi_\rho^*(\partial_t \rho) - \psi_\rho\left(\frac{\delta \mathcal{E}}{\delta \rho}\right),$$

and equality holds if and only if

$$\partial_t \rho \in \partial \psi_\rho\left(-\frac{\delta \mathcal{E}}{\delta \rho}\right). \quad (29)$$

The case of the M -gradient flow (29) corresponds to

$$\psi_\rho^*(\xi) := \frac{1}{2} \|\xi\|_{M_\rho}^2 \quad \text{and} \quad \psi_\rho(s) := \frac{1}{2} \|s\|_{M_\rho^{-1}}^2.$$

The obvious generalisation of (20) then is

$$\rho \text{ is a solution of the } (\mathcal{E}, \psi)\text{-gradient-flow (29)} \iff J_\psi(\rho) = 0, \quad (30)$$

where J_ψ is given by

$$J_\psi(\rho) := \mathcal{E}(\rho(T)) - \mathcal{E}(\rho(0)) + \int_0^T \left[\psi_\rho^*(\partial_t \rho) + \psi_\rho\left(-\frac{\delta \mathcal{E}}{\delta \rho}\right) \right] dt. \quad (31)$$

5.1 Birth-death processes

A simple example of a stochastic process with non-quadratic dissipation ψ and a corresponding generalised gradient flow is a birth-death process, which is a continuous-time jump process on \mathbb{Z} . The system may only jump to neighbours, from position k with rate a_k to $k+1$ and with rate b_k to $k-1$. We construct a continuum limit by defining the new stochastic variable U_n by rescaling time t and position $k(t)$ with n :

$$U_n(t) := \frac{k(nt)}{n}.$$

A standard argument gives the large-deviation behaviour for U_n in terms of the rate functional (see [4] for a finite-lattice proof of the claims made below). If we choose the jump rates so that

$$a_k = \alpha e^{-\mathcal{E}'(k/n)} \quad \text{and} \quad b_k = \alpha e^{+\mathcal{E}'(k/n)}$$

for $\alpha > 0$ and some smooth function $\mathcal{E}: \mathbb{R} \rightarrow \mathbb{R}$, then the rate functional is

$$I(u) = \int_0^T L(u(t), u'(t)) dt,$$

with

$$L(u, v) = v \log \frac{v + \sqrt{v^2 + 4\alpha^2}}{2\alpha \exp(-\mathcal{E}'(u))} - \sqrt{v^2 + 4\alpha^2} + \alpha e^{-\mathcal{E}'(u)} + \alpha e^{+\mathcal{E}'(u)}.$$

Writing

$$\psi^*(v) = v \log \frac{v + \sqrt{v^2 + 4\alpha^2}}{2\alpha} - \sqrt{v^2 + 4\alpha^2},$$

it follows that $\psi(\xi) = \alpha(e^\xi + e^{-\xi})$, and I can be written in the form (31).

The corresponding generalised gradient flow in \mathbb{R} , given by (29), reads

$$\dot{u} = 2 \sinh(-\mathcal{E}'(u)).$$

Observe how this differs from the standard (quadratic-dissipation) gradient flow, which is $\dot{u} = -\mathcal{E}'(u)$; the non-quadratic dissipation preserves the sign of the velocity, but not its amplitude. Because of the preservation of sign, the energy \mathcal{E} is monotonic along a solution:

$$\frac{d}{dt} \mathcal{E}(u(t)) = \mathcal{E}'(u(t)) \dot{u}(t) = 2\mathcal{E}'(u) \sinh(-\mathcal{E}'(u)) \leq 0.$$

This example shows how the connection between large-deviation principles and (generalised) gradient flows extends to the case of non-quadratic dissipations. Note that here the large deviations refer to a single process and henceforth are not due to an averaging process as in the empirical measure case.

5.2 Spin-flip processes

For $n \in \mathbb{N}$, let \mathbb{T}_n be the one-dimensional n -torus ($\mathbb{Z}/n\mathbb{Z}$). An Ising spin at sites of \mathbb{T}_n takes values in $\{-1, +1\}$ and is subject to a rate-1 independent spin-flip dynamics. We consider the trajectory of the magnetisation, i.e., $t \mapsto m_n(t) = \frac{1}{n} \sum_{i \in \mathbb{T}_n} \sigma_i(t)$, where $\sigma_i(t)$ is the spin at site $i \in \mathbb{T}_n$ at time t . The generator for the process $(m_n(t))_{t \geq 0}$ is given by

$$(A_n f)(m) = \frac{(1+m)}{2} n [f(m - 2n^{-1}) - f(m)] + \frac{(1-m)}{2} n [f(m + 2n^{-1}) - f(m)]$$

for $m \in \{-1, -1 + 2n^{-1}, \dots, 1\}$. The trajectory of the magnetisation satisfies a large deviation principle, i.e., for every trajectory $\gamma = (\gamma_t)_{t \in [0, T]}$,

$$\text{Prob}((m_n(t))_{t \in [0, T]} \approx (\gamma_t)_{t \in [0, T]}) \approx \exp \left[-n \int_0^T L(\gamma_t, \dot{\gamma}_t) dt \right],$$

where the Lagrangian L can be computed following the scheme of Feng and Kurtz [9, Example 1.5.]. We obtain

$$L(m, q) = \frac{q}{2} \log \left(\frac{q + \sqrt{q^2 + 4(1-m^2)}}{2(1-m)} \right) - \frac{1}{2} \sqrt{q^2 + 4(1-m^2)} + 1.$$

This can similarly be written as $\psi^*(q) + \psi(-\mathcal{E}'(m)) + q\mathcal{E}'(m)$, where

$$\psi^*(q) = \frac{q}{2} \log \frac{q + \sqrt{q^2 + 4(1-m^2)}}{2\sqrt{1-m^2}} - \frac{1}{2} \sqrt{q^2 + 4(1-m^2)}$$

and

$$\psi(\xi) = \frac{1}{2} \sqrt{1-m^2} (\exp(2\xi) + \exp(-2\xi));$$

the involved energy is

$$\mathcal{E}(m) = \frac{1}{4}(1+m) \log(1+m) + \frac{1}{4}(1-m) \log(1-m).$$

Then the limiting equation (29) can be written as $\dot{m} = -2m$. This is consistent with the optimal trajectory via the Euler-Lagrange equation, $m(t) = m_0 e^{-2t}$.

6 Discussion

In the sections above we have described a number of pairs of systems, each consisting of a stochastic process and its continuum limit. Each pair has the property that the large deviations of the stochastic process are closely linked to a gradient-flow structure of the limit equation. These links are time-dynamic versions of the equilibrium connection mentioned in the introduction. We now describe how this provides us insight into the properties of the gradient-flow structures for each pair.

6.1 Wasserstein gradient flows

We claim that the Wasserstein metric characterises the mobility of the empirical measure of a large number of Brownian particles. Indeed, this claim can be made meaningful in a number of different ways:

1. In discrete time, letting ρ_n be the empirical measure of a system of Brownian particles, we have

$$\text{Prob}(\rho_n(h) \approx \rho^1 | \rho_n(0) \approx \rho^0) \sim e^{-nh \cdot d(\rho^0, \rho^1)^2/4} \quad \text{as } n \rightarrow \infty,$$

which follows from (16) and was proved independently in [14].

2. In continuous time, for the whole path $\rho_n: [0, T] \rightarrow \mathcal{M}_1(\mathbb{R}^d)$ of empirical measures up to a fixed terminal time T , we have

$$\text{Prob}(\rho_n \approx \rho) \sim e^{-nI(\rho)}, \quad \text{as } n \rightarrow \infty,$$

where I , defined in (23), measures the size of the deviation by the Wasserstein metric tensor $\|\cdot\|_\rho$.

3. When the particles also undergo a deterministic drift, the same statement holds with I defined by (26), where again the size of the deviation is measured by the norm $\|\cdot\|_\rho$.

The origin of this role of the Wasserstein metric as the mobility of Brownian particles can be understood by considering the geometric relationship between $(\mathbb{R}^d)^n$ and the space of measures endowed with the Wasserstein distance. Consider the embedding

$$e: (\mathbb{R}^d)^n \rightarrow (\mathcal{M}_1(\mathbb{R}^d), d), \quad (x_1, \dots, x_n) \mapsto \frac{1}{n} \sum_{i=1}^n \delta_{x_i}.$$

Note that e is not one-to-one, since the numbering of the particles is lost: the particles have become indistinguishable. Indeed, one can identify the set of empirical measures of the form $n^{-1} \sum_i \delta_{x_i}$ with the space obtained by identifying all elements in $(\mathbb{R}^d)^n$ that are rearrangements of each other, i.e., the quotient space $(\mathbb{R}^d)^n/S_n$, where S_n is the set of all permutations of n elements.

Now the Wasserstein metric on \mathcal{M}_1 makes the embedding of $(\mathbb{R}^d)^n/S_n$ in $\mathcal{M}_1(\mathbb{R}^d)$ *isometric*. This follows from the simple property that

$$d\left(\frac{1}{n} \sum_{i=1}^n \delta_{x_i}, \frac{1}{n} \sum_{j=1}^n \delta_{y_j}\right)^2 = \frac{1}{n} \inf_{\sigma \in S_n} \sum_{i=1}^n |x_i - y_{\sigma(i)}|^2. \quad (32)$$

With this property the role of the Wasserstein distance can be fully explained. The Freidlin-Wentzell theory for Brownian particles [10] shows how the mobility of a *vector* $X = (X_1, \dots, X_n)$ of n Brownian particles has a stochastic mobility given by the Euclidean norm $(x, y) \mapsto \sum_i |x_i - y_i|^2$, in the sense that

$$\text{Prob}(X(h) \approx y | X(0) \approx x) \sim \exp -\frac{1}{4h} \sum_{i=1}^n |x_i - y_i|^2 \quad \text{for small } h.$$

The loss of information upon introducing indistinguishability, or equivalently upon transforming to empirical measures, implies by the contraction principle (e.g., [7, Sec. 4.2.1]) that the exponent $(1/4) \sum_i |x_i - y_i|^2$ becomes replaced by its minimum under rearrangement,

$$\frac{1}{4h} \inf_{\sigma \in S_n} \sum_{i=1}^n |x_i - y_{\sigma(i)}|^2.$$

This expression is equal to $n/(4h)$ times (32). If we gloss over the approximations in different limits ($h \rightarrow 0$ and $n \rightarrow \infty$), this explains how the Wasserstein distance is the natural measure of the mobility of an *empirical measure* of Brownian particles, through transformation of the original mobility of a single Brownian particle.

6.2 Consequences for modelling

Gradient flows can be thought of as overdamped systems, in the sense that any inertial effects are damped out quickly by the effects of viscous, frictional, or other damping forces, and can therefore be neglected. One way of modelling such overdamped systems is therefore by assuming an abstract gradient-flow structure from the start and making it concrete by postulating an energy \mathcal{E} and a dissipation potential ψ . These choices should be motivated, and in the case of Wasserstein and Wasserstein-like dissipations this motivation is non-trivial.

One area where this is particularly visible is in the modelling of lower-dimensional structures, such as threads and surfaces, moving through a viscous fluid. The biology of sub-cell structures knows many such examples, including microtubules and lipid bilayers. The assumption of overdampedness is reasonable in this viscosity-dominated situation, but the interplay of geometry and mechanics makes the direct formulation of evolution equations complicated and error-prone (see, e.g., [2]). In this context, the construction of evolution equations through the postulation of energy and dissipation is often simpler and allows for clearer separation of the various assumptions. However, it remains necessary to motivate the choices made for the energy and the dissipation.

To take the Wasserstein metric as an example, its interpretation as the measure of mobility of empirical measures of Brownian particles provides such a motivation, and because of the connection to the Brownian mobility of the particles it also allows for generalisation to other situations.

But similar arguments apply to other dissipations, coupled to other underlying stochastic processes. For instance, the symmetric simple exclusion process leads to $\rho(1 - \rho)$ mobility, implying that if such an exclusion process is one's idea of the underlying system, then the $\rho(1 - \rho)$ -dissipation is the natural choice.

One might go even further. The diffusion equation (17) is known to be a gradient flow in many different ways; in addition to the two mentioned above, also as the L^2 -gradient flow

of the Dirichlet integral $\int |\nabla \rho|^2$, for instance, as the H^{-1} -gradient flow of the L^2 -norm, and even as the H^{s-1} -gradient flow of the H^s -seminorm for each $s \in \mathbb{R}$. For the two structures that we have discussed, the different underlying stochastic processes provide clear reasons for the differing dissipations and energies. Here we formulate the

Conjecture 1 *Each gradient-flow structure can be connected to an appropriate stochastic process via a large-deviation principle.*

To the extent that this conjecture turns out to be true, it provides an explanation for the occurrence of multiple gradient-flow formulations of the same differential equation.

6.3 Geometry and reversibility

There are interesting connections between the geometry of the Brownian noise, the reversibility of the stochastic process, and the question whether the resulting evolution equation is a gradient flow or not.

This becomes apparent when we modify the system of Section 4.3 by introducing a diffusion matrix $A \in \mathbb{R}^{d \times d}$ and replacing the scalar σ by a mobility matrix $\sigma \in \mathbb{R}^{d \times d}$, thus obtaining

$$dX_i(t) = -A \nabla \Psi(X_i(t)) dt - \frac{1}{n} \sum_{j=1}^n A \nabla \Phi(X_i(t) - X_j(t)) dt + \sqrt{2} \sigma dW_i(t). \quad (33)$$

The large-deviation rate functional of the system is similarly given by

$$I(\rho) := \frac{1}{2} \int_0^T \left\| \partial_t \rho - \operatorname{div} \sigma \sigma^T \nabla \rho - \operatorname{div} \rho A \nabla [\Psi + \rho * \Phi] \right\|_{D(\rho),*}^2 dt, \quad (34)$$

where the norm $\|\cdot\|_{D(\rho),*}$ is induced by (10) with $D(\rho) = \rho \sigma \sigma^T$. The formula (34) implies that the hydrodynamic limit of this system is the minimiser of I , satisfying

$$\partial_t \rho = \operatorname{div} \sigma \sigma^T \nabla \rho + \operatorname{div} \rho A \nabla [\Psi + \rho * \Phi]. \quad (35)$$

With this additional parameter freedom, it is not always possible to write (34) in the form (22). This depends on whether the cross term in (34) is an exact differential, i.e., whether there exists a functional \mathcal{E} such that

$$\left(\partial_t \rho, -\operatorname{div} \sigma \sigma^T \nabla \rho - \operatorname{div} \rho A \nabla [\Psi + \rho * \Phi] \right)_{D(\rho),*} = \partial_t \mathcal{E}(\rho).$$

This is the case if and only if $\sigma \sigma^T$ is a positive multiple of A , a condition that is familiar from the fluctuation-dissipation theorem. In that case, and writing $\sigma \sigma^T = kT A$ for some ‘temperature’ $T > 0$ and the Boltzmann constant k ,

$$-\operatorname{div} \sigma \sigma^T \nabla \rho - \operatorname{div} \rho A \nabla [\Psi + \rho * \Phi] = M_\rho \frac{\delta \mathcal{F}}{\delta \rho},$$

where $M_\rho \xi$ is defined as $-\operatorname{div} D(\rho) \nabla \xi$ and the free energy \mathcal{F} is a modification of (27),

$$\mathcal{F}(\rho) := \operatorname{Ent}(\rho) + \frac{1}{kT} \int_{\mathbb{R}^d} \left[\rho \Psi + \frac{1}{2} \rho (\rho * \Phi) \right].$$

Then the rate functional I can be written in the form (19) as

$$I(\rho) = \mathcal{F}(\rho(T)) - \mathcal{F}(\rho(0)) + \frac{1}{2} \int_0^T \left[\|\partial_t \rho\|_{D(\rho),*}^2 + \left\| \frac{\delta \mathcal{F}}{\delta \rho} \right\|_{D(\rho)}^2 \right] dt$$

and the evolution equation (35) is the (modified, D -) Wasserstein gradient flow of \mathcal{F} .

Our freedom to choose A and σ separately gives us the insight that for this system the following four statements are equivalent:

1. $\sigma \sigma^T = kTA$ for some $T > 0$;
2. The evolution (35) is a $D(\rho)$ -Wasserstein gradient flow of \mathcal{F} ;
3. The rate functional I can be written in the form (19);
4. For any finite number n of particles, the system (33) is reversible.

We expect that such an equivalence property, including the reversibility of the microscopic system, might hold more generally.

6.4 Diffusion with decay

Yet another generalisation concerns systems with decay, which is implemented as a jump process. In [17], Peletier and Renger have derived a similar connection for the case of diffusing particles that are convected and may also decay, given by the equation (in one space dimension)

$$\partial_t \rho = \partial_{xx} \rho + \partial_x(\rho \partial_x \Psi) - \lambda \rho, \quad (36)$$

with $\Psi \in C_b^2(\mathbb{R})$ and $\lambda \geq 0$.

In [17], the particles perform a Brownian motion in the spatial dimension, augmented by a deterministic drift given by $-\partial_x \Psi$. This part of the process gives rise to the two terms $\partial_{xx} \rho + \partial_x(\rho \partial_x \Psi)$. In addition, the particles change their state from ‘normal’ to ‘decayed’, after an exponentially distributed time; this part gives rise to the term $-\lambda \rho$. The opposite transition is not allowed: decay is irreversible.

An analysis similar to Section 3.1 then connects the large-deviation rate functional for this stochastic particle system to a corresponding minimisation problem describing the time-discrete evolution, i.e., the equivalent of (12). In this case the time-discrete minimisation problem is

$$\rho^k \in \arg \min_{\rho} \inf_{\rho_{ND}: |\rho + \rho_{ND}| = |\rho^{k-1}|} -\frac{1}{2} \mathcal{F}(\rho + \rho_{ND}) - \frac{1}{2} \mathcal{F}(\rho^{k-1}) + \frac{1}{4h} d(\rho + \rho_{ND}, \rho^{k-1})^2 + \mathcal{F}(\rho) + \mathcal{F}(\rho_{ND}) - |\rho| \log e^{-\lambda h} - |\rho_{ND}| \log(1 - e^{-\lambda h}), \quad (37)$$

where $|\rho| := \int \rho$ and the free energy \mathcal{F} is defined as

$$\mathcal{F}(\rho) = \text{Ent}(\rho) + \int \Psi d\rho.$$

In [17], the authors explain how the structure of (37) can be understood: if we define

$$\begin{aligned} K_{\Psi}^h(\bar{\rho}; \rho^{k-1}) &:= \frac{1}{2} \mathcal{F}(\bar{\rho}) - \frac{1}{2} \mathcal{F}(\rho^{k-1}) + \frac{1}{4h} d(\bar{\rho}, \rho^{k-1})^2, \\ K_{Dec}^h(\rho; \bar{\rho}) &:= \mathcal{F}(\rho) - \mathcal{F}(\bar{\rho}) + \mathcal{F}(\bar{\rho} - \rho) - |\rho| \log e^{-\lambda h} + |\bar{\rho} - \rho| \log(1 - e^{-\lambda h}), \end{aligned}$$

then the terms inside the infimum in (37) can be written as $K^h(\rho + \rho_{ND}; \rho^{k-1}) + K_{Dec}^h(\rho; \rho + \rho_{ND})$. In this decomposition, the first term describes diffusion and convection by Ψ of the joint measure $\rho + \rho_{ND}$ starting from the previous state ρ^{k-1} , similar to (16) and (12). The second term describes the decay process, in which the joint diffused-and-convected measure $\rho + \rho_{ND}$ is split into a part ρ that remains ‘normal’ and the remainder ρ_{ND} that becomes decayed.

While the structure of (37) is not the same as (12), and (37) does not represent a time discretisation of a gradient flow, both are minimisation problems that define the next step in the iteration, and in both cases one can identify a driving force (the free energy \mathcal{F} , in the case of (37)) and a mechanism that acts as a brake. In K_{Ψ}^h the ‘brake’ is the Wasserstein metric $d(\bar{\rho}, \rho^{k-1})^2/4h$, and in K_{Dec}^h it is the two terms $-|\rho| \log e^{-\lambda h} + |\bar{\rho} - \rho| \log(1 - e^{-\lambda h})$. In both cases these terms restrict the movement of $\bar{\rho}$ respectively ρ , and this restriction becomes more and more severe as $h \rightarrow 0$.

6.5 General remarks on interacting particle systems

Section 5 explained how, once a large deviation principle for the interacting particle system with rate functional $I(\rho)$ is established, different Wasserstein-type metrics occur in a natural way. Such large deviation results are stronger than results on limit equations. Indeed, a part of the standard proof of a large deviation result involves modifying the process by adding a forcing such that a given path which does not solve the original limit equation solves the limit equation of the modified process. So the question arises whether the point of view advocated in this paper has the potential of deriving limit equations without using large deviation results which contain limit results derived in the classical way. This open question is of particular importance because limit points of the implicit time discretisation provide a weak notion of solution of the limit equation in cases where distributional solutions are not appropriate, e.g., for problems with a sharp interface like the mean curvature flow. In situations such as (24), where a particle interacts with the average of many others, the distribution of a family of initially independent particles stays close to a product measure (propagation of chaos), so a modification of the techniques for independent particles seems promising.

7 Conclusion

The examples of this paper illustrate how the two concepts of large-deviation principles for stochastic particle systems and gradient flows are closely entwined. Further examples are currently under study, such as Brownian particles with inertia, which lead to the Kramers’ equation, and rate-independent systems such as friction and fracture. We expect that many more examples of this kind will be uncovered.

A Free energy and the Boltzmann distribution

In this appendix we show how the *free energy*

$$\mathcal{F}(\rho) := H(\rho|\mu) + \frac{1}{k\theta}E(\rho) \quad (38)$$

arises from the coupling of a system of particles with a *heat bath*. Here $\theta > 0$ (in Joules) is the temperature of the heat bath, and the *Boltzmann constant* k has the value $1.4 \cdot 10^{-23}$ J/K.

The measure $\mu \in \mathcal{P}(\mathcal{X})$ is the probability distribution of the particles in a state space \mathcal{X} , and E is the average energy of the particles:

$$E(\rho) = \int_{\mathcal{X}} e(x) \rho(dx),$$

where $e: \mathcal{X} \rightarrow \mathbb{R}$ is a fixed function that we call the *energy* of a state $x \in \mathcal{X}$. We now construct an explicit system in which \mathcal{F} arises as the large-deviation rate functional. This will allow us to interpret all these concepts in the context of large deviations.

We start by choosing a system S and its connection to a heat bath called S_B . Both are probabilistic systems of particles; S consists of n independent particles $X_i \in \mathcal{X}$, with probability law $\mu \in \mathcal{P}(\mathcal{X})$; similarly S_B consists of m independent particles $Y_j \in \mathcal{Y}$, with law $\nu \in \mathcal{P}(\mathcal{Y})$. The total state space of the system is therefore $\mathcal{X}^n \times \mathcal{Y}^m$.

The *coupling* between these systems is done via an *energy constraint*. We assume that there are energy functions $e: \mathcal{X} \rightarrow \mathbb{R}$ and $e_B: \mathcal{Y} \rightarrow \mathbb{R}$, and we will constrain the joint system to be in a state of fixed total energy, i.e., we will only allow states in $\mathcal{X}^n \times \mathcal{Y}^m$ that satisfy

$$\sum_{i=1}^n e(X_i) + \sum_{j=1}^m e_B(Y_j) = \text{constant}. \quad (39)$$

The physical interpretation of this is that energy (in the form of heat) may flow freely from one system to the other, but no other form of interaction is allowed.

Similar to the example in the Introduction, we describe the total states of systems S and S_B by empirical measures $\rho_n = \frac{1}{n} \sum_i \delta_{X_i}$ and $\zeta_m = \frac{1}{m} \sum_j \delta_{Y_j}$. We define the average energies $E(\rho_n) := \frac{1}{n} \sum_i e(X_i) = \int_{\mathcal{X}} e d\rho_n$ and $E_B(\zeta_m) := \int_{\mathcal{Y}} e_B d\zeta_m$, so that the energy constraint (39) reads $nE(\rho_n) + mE_B(\zeta_m) = \text{constant}$.

By Sanov's theorem each of the systems *separately* satisfies a large-deviation principle with rate functions $I(\rho) = \mathcal{H}(\rho|\mu)$ and $I_B(\zeta) = \mathcal{H}(\zeta|\nu)$. However, instead of using the explicit formula for I_B , we are going to assume that I_B can be written as a function of the energy E_B of the heat bath alone, i.e., $I_B(\zeta) = \tilde{I}_B(E_B(\zeta))$. For the coupled system we derive a joint large-deviation principle by choosing that (a) $m = nN$ for some large $N > 0$, and (b) the constant in (39) scales as nN , i.e.,

$$nE(\rho_n) + nNE_B(\zeta_{nN}) = nN\bar{E} \quad \text{for some } \bar{E}.$$

Formally, the joint system then satisfies a large-deviation principle

$$\text{Prob}\left((\rho_n, \zeta_{nN}) \approx (\rho, \zeta) \mid E(\rho_n) + NE_B(\zeta_{nN}) = \bar{E}\right) \sim \exp(-nJ(\rho, \zeta)),$$

with rate functional

$$J(\rho, \zeta) := \begin{cases} \mathcal{H}(\rho|\mu) + N\tilde{I}_B(E_B(\zeta)) + \text{constant} & \text{if } E(\rho) + NE_B(\zeta) = N\bar{E}, \\ +\infty & \text{otherwise.} \end{cases}$$

Here the constant is chosen to ensure that $\inf J = 0$.

The functional J can be reduced to a functional of ρ alone,

$$J(\rho) = \mathcal{H}(\rho|\mu) + N\tilde{I}_B\left(\bar{E} - \frac{E(\rho)}{N}\right) + \text{constant}.$$

In the limit of large N , one might approximate

$$N\tilde{I}_B\left(\bar{E} - \frac{E(\rho)}{N}\right) \approx N\tilde{I}_B(\bar{E}) - \tilde{I}'_B(\bar{E})E(\rho).$$

The first term above is absorbed in the constant, and we find

$$J(\rho) \approx \mathcal{H}(\rho|\mu) - E(\rho)\tilde{I}'_B(\bar{E}) + \text{constant}.$$

We expect that \tilde{I}'_B is negative, since larger energies typically lead to higher probabilities and therefore smaller values of I_B . Now we simply define $k\theta := -1/\tilde{I}'_B(\bar{E})$, and we find

$$J(\rho) \approx \mathcal{H}(\rho|\mu) + \frac{1}{k\theta}E(\rho) + \text{constant}.$$

This is the same expression as (38). Note that the right-hand side can be written as $\mathcal{H}(\rho|\tilde{\mu})$, where $\tilde{\mu}$ is the tilted distribution

$$\tilde{\mu}(A) = \frac{\int_A e^{-e(x)/k\theta} \mu(dx)}{\int_{\mathcal{X}} e^{-e(x)/k\theta} \mu(dx)}.$$

This derivation shows that the effect of the heat bath is to *tilt* the system S : a state ρ of S with larger energy $E(\rho)$ implies a smaller energy E_B of S_B , which in turn reduces the probability of ρ . This is reflected in the approximation $\tilde{I}'_B(\bar{E})E(\rho)$ of $I_B(\zeta)$. The role of temperature θ is that of an exchange rate, since it characterises the change in probability (as measured by the rate function I_B) per unit of energy. When θ is large, the exchange rate is low, and then larger energies incur only a small probabilistic penalty. When temperature is low, then higher energies are very expensive, and therefore more rare. From this point of view, the Boltzmann constant k is simply the conversion factor that converts our Kelvin temperature scale for θ into the appropriate ‘exchange rate’ scale.

In thermodynamics one often encounters the identity (or definition) $\theta = dS/dE$. This is formally the same as our definition of $k\theta$ as $-dI_B/dE$, if one interprets I_B as an entropy and adopts the convention to multiply the non-dimensional quantity I_B with $-k$.

Acknowledgement. The authors wish to thank Dejan Slepčev and Rob Jack for various interesting discussions. The research of M. A. Peletier and J. Zimmer has received funding from the Initial Training Network “FIRST” of the Seventh Framework Programme of the European Community (grant agreement number 238702). The research of S. Adams was supported by EPSRC grant number EP/I003746/1.

References

- [1] Stefan Adams, Nicolas Dirr, Mark A. Peletier, and Johannes Zimmer. From a large-deviations principle to the Wasserstein gradient flow: A new micro-macro passage. *Communications in Mathematical Physics*, 307:791–815, 2011.
- [2] Marino Arroyo and Antonio DeSimone. Relaxation dynamics of fluid membranes. *Phys. Rev. E* (3), 79(3):031915, 17, 2009.

- [3] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3):375–393, 2000.
- [4] Terence Chan. Large deviations and quasi-stationarity for density-dependent birth-death processes. *J. Austral. Math. Soc. Ser. B*, 40(2):238–256, 1998.
- [5] D. A. Dawson and J. Gärtner. Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics*, 20(4):247–308, 1987.
- [6] D. A. Dawson and J. Gärtner. Large deviations, free energy functional and quasi-potential for a mean field model of interacting diffusions. *Mem. Amer. Math. Soc.*, 78(398):iv+94, 1989.
- [7] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.
- [8] Nicolas Dirr, Vaios Laschos, and Johannes Zimmer. Upscaling from particle models to entropic gradient flows. Submitted, 2010.
- [9] J. Feng and Th. G. Kurtz. *Large deviations for stochastic processes*. American Mathematical Society, 2006.
- [10] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, New York, second edition, 1998. Translated from the 1979 Russian original by Joseph Szücs.
- [11] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17 (electronic), 1998.
- [12] C. Kipnis, S. Olla, and S. R. S. Varadhan. Hydrodynamics and large deviation for simple exclusion processes. *Comm. Pure Appl. Math.*, 42(2):115–137, 1989.
- [13] Claude Kipnis and Stefano Olla. Large deviations from the hydrodynamical limit for a system of independent Brownian particles. *Stochastics Stochastics Rep.*, 33(1-2):17–25, 1990.
- [14] C. Léonard. A large deviation approach to optimal transport. <http://arxiv.org/abs/0710.1461>, 2007.
- [15] H.C. Öttinger. *Beyond equilibrium thermodynamics*. Wiley Online Library, 2005.
- [16] Felix Otto. The geometry of dissipative evolution equations: The porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.
- [17] M. A. Peletier and M. Renger. Variational formulation of the Fokker-Planck equation with decay: A particle approach. *Arxiv preprint arXiv:1108.3181*, 2011.
- [18] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.