

The Gradient Flow Approach to Hydrodynamic Limits for the Simple Exclusion Process

Max Fathi and Marielle Simon

Abstract We present a new approach to prove the macroscopic hydrodynamic behaviour for interacting particle systems, and as an example we treat the well-known case of the symmetric simple exclusion process (SSEP). More precisely, we characterize any possible limit of its empirical density measures as solutions to the heat equation by passing to the limit in the gradient flow structure of the particle system.

Keywords Hydrodynamic limits · Gradient flows structure · Diffusive heat equation · Reversible markov chains

1 Introduction

The aim of this work is to show how one can use gradient flow structures to prove convergence to the hydrodynamic limit for interacting particle systems. The exposition is focused on the case of the symmetric simple exclusion process on the discrete d -dimensional torus, but the strategy can be adapted to other reversible particle systems, such as zero-range processes (see [11] for the definitions of these models). Gradient flows are ordinary differential equations of the form

$$\dot{x}(t) = -\nabla V(x(t)),$$

where ∇V denotes the gradient of the function V . De Giorgi and his collaborators showed in [5] how to give a meaning to solutions to such equations in the setting of

M. Fathi
Department of Mathematics, University of California, Berkeley, USA
e-mail: mfathi@phare.normalesup.org

M. Simon (✉)
Équipe MEPHYSTO, INRIA Lille Nord Europe, 40 avenue du Halley,
59650 Villeneuve d'Ascq, Lille, France
e-mail: marielle.simon@inria.fr

metric spaces: these solutions are called minimizing-movement solutions, or curves of maximal slope. When considering the case of spaces of probability measures, one can use this notion to rewrite the partial differential equations governing the time evolution of the laws of diffusion processes, such as the heat equation, as gradient flows, for which V is the entropy with respect to the optimal transport (or Wasserstein) distance. We refer to [3] for more details. This framework can be adapted to the case of reversible Markov chains on finite spaces. This was proven independently by Maas [13] and Mielke [16], who both developed a discrete counterpart to the Lott-Sturm-Villani theory of lower bounds on Ricci curvature for metric spaces.

These gradient flow structures are a powerful tool to study convergence of sequences of dynamics to some limit. Two main strategies have already been developed. One of them consists in using the discrete (in time) approximation schemes suggested by the gradient flow structure (see for example [4]). The second one, which we shall use here, consists in characterizing gradient flows in terms of a relation between the energy function and its variations, and passing to the limit in this characterization. It was first developed by Sandier and Serfaty in [18], and then generalized in [19]. This strategy can be combined with the gradient flow structure of [13, 16] to prove convergence to some scaling limit for interacting particle systems. This was recently done for chemical reaction equations in [14] and mean-field interacting particle systems on graphs in [6].

Gradient flow structures are also related to large deviations, at least when considering diffusion processes, see [1, 8]. While we only present here the case of the SSEP, the technique is fairly general, and can be adapted to other reversible interacting particle systems. For example, the adaptation of the proof to the case of a zero-range process on the lattice (with nice rates) is quite straightforward. It would be very interesting to apply this method to obtain other PDEs, more degenerate than the heat equation, as hydrodynamic limits of some interacting particle system: for instance, porous medium and fast diffusion equations (see [17, 20]) also have a gradient flow structure, and are not directly solvable by standard techniques.

The plan of the sequel is as follows: in Sect. 2, we present the gradient flow framework for Markov chains on discrete spaces developed in [13, 16]. In Sect. 3, we expose the setup for proving convergence of gradient flows. Finally, in Sect. 4, we investigate the symmetric simple exclusion process, and reprove the convergence to its hydrodynamic limit.

2 Gradient Flow Structure for Reversible Markov Chains

2.1 Framework

We start by describing the gradient flow framework for Markov chains on discrete spaces. The presentation we use here is the one of [13]. We consider an irreducible continuous time reversible Markov chain on a finite space \mathcal{X} with kernel $K : \mathcal{X} \times$

$\mathcal{X} \rightarrow \mathbb{R}_+$ and invariant probability measure ν . Let $\mathcal{P}(\mathcal{X})$ (resp. $\mathcal{P}_+(\mathcal{X})$) be the set of probability densities (resp. positive) with respect to ν . The probability law $\rho_t \nu$ of the Markov chain at time t satisfies the evolution equation

$$\dot{\rho}_t(x) + \sum_{y \in \mathcal{X}} (\rho_t(x) - \rho_t(y))K(x, y) = 0, \quad \text{for all } x \in \mathcal{X}. \quad (1)$$

Hereafter we denote by $\dot{\rho}_t(x)$ the derivative with respect to time of the function $(t, x) \mapsto \rho_t(x)$. Given a function $\psi : \mathcal{X} \rightarrow \mathbb{R}$, we define $\nabla \psi(x, y) := \psi(y) - \psi(x)$. The discrete divergence of a function $\Phi : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ is defined as

$$\text{div}(\Phi)(x) := \frac{1}{2} \sum_{y \in \mathcal{X}} (\Phi(x, y) - \Phi(y, x))K(x, y).$$

With these definitions, we have the integration by parts formula

$$\sum_{x, y \in \mathcal{X}} \nabla \psi(x, y)\Phi(x, y)K(x, y)\nu(x) = - \sum_{x \in \mathcal{X}} \psi(x) \text{div}(\Phi)(x)\nu(x). \quad (2)$$

Let us introduce three notions we shall use to define the gradient flow structure:

Definition 2.1 1. The *relative entropy with respect to ν* is defined as

$$\text{Ent}_\nu(\rho) := \sum_{x \in \mathcal{X}} \nu(x)\rho(x) \log \rho(x), \quad \text{for } \rho \in \mathcal{P}(\mathcal{X}),$$

with the convention that $\rho(x) \log \rho(x) = 0$ if $\rho(x) = 0$. We sometimes denote $\mathcal{H}(\rho) := \text{Ent}_\nu(\rho)$, whenever ν is fixed and no confusion arises.

2. The *symmetric Dirichlet form* is given for two real-valued functions ϕ, ψ by

$$\mathcal{E}(\phi, \psi) := \frac{1}{2} \sum_{x, y \in \mathcal{X}} (\phi(y) - \phi(x))(\psi(y) - \psi(x))K(x, y)\nu(x),$$

3. The *Fisher information* (or entropy production) writes as $\mathcal{I}(\rho) := \mathcal{E}(\rho, \log \rho)$.

Notice that Ent_ν is the *mathematical* entropy, and not the physical entropy. It decreases along solutions of (1), so in physical terms it plays the role of a free energy. We call \mathcal{I} the entropy production since along solutions of (1) we have

$$\frac{d}{dt} \text{Ent}_\nu(\rho_t) = -\mathcal{I}(\rho_t).$$

2.2 Continuity Equation

We introduce the *logarithmic mean* $\Lambda(a, b)$ of two non-negative numbers a, b as

$$\Lambda(a, b) = \int_0^1 a^s b^{1-s} ds = \frac{b - a}{\log(b) - \log(a)} \quad \text{if } a \neq b, a > 0, b > 0,$$

and $\Lambda(a, a) = a$, and also $\Lambda(a, b) = 0$ if $a = 0$ or $b = 0$. The mean Λ satisfies:

$$\sqrt{ab} \leq \Lambda(a, b) \leq (a + b)/2, \quad \text{for all } a \geq 0, b \geq 0. \quad (3)$$

Let us now define, for $\rho \in \mathcal{P}_+(\mathcal{X})$, its *logarithmic mean* $\hat{\rho}$ defined on $\mathcal{X} \times \mathcal{X}$ as

$$\hat{\rho}(x, y) := \Lambda(\rho(x), \rho(y)). \quad (4)$$

In order to define a suitable metric on $\mathcal{P}(\mathcal{X})$, we need a representation of curves as solving a continuity equation:

Lemma 2.1 *Given a smooth flow of positive probability densities $\{\rho_t\}_{t \geq 0}$ on \mathcal{X} , there exists a function $(t, x) \mapsto \psi_t(x)$ such that the following continuity equation holds for any $t \geq 0$ and $x \in \mathcal{X}$:*

$$\dot{\rho}_t(x) + \sum_{y \in \mathcal{X}} (\psi_t(y) - \psi_t(x)) K(x, y) \hat{\rho}_t(x, y) = 0. \quad (5)$$

Moreover, for any $t \geq 0$, $\psi_t(\cdot)$ is unique up to an additive constant.

We refer to [13, Sect. 3] for the proof.

Definition 2.2 Given $\rho \in \mathcal{P}(\mathcal{X})$ and $\psi : \mathcal{X} \rightarrow \mathbb{R}$, we define *the action*

$$\mathcal{A}(\rho, \psi) := \frac{1}{2} \sum_{x, y \in \mathcal{X}} (\psi(y) - \psi(x))^2 \hat{\rho}(x, y) K(x, y) \nu(x) \geq 0.$$

A distance between two probability densities (ρ_0, ρ_1) could then be defined as the infimum of the action of all curves $\{\rho_t, \psi_t\}_{t \in [0, 1]}$ linking these densities, as was done in [13]. However, we do not need to introduce that metric here, since we shall only use the formulation of gradient flows as minimizing-movement curves, as follows:

Proposition 2.2 *Let $\{\rho_t\}_{t \geq 0}$ be a smooth flow of probability densities on \mathcal{X} , and let $\{\psi_t\}_{t \geq 0}$ be such that the continuity equation (5) holds. Then, for any $T > 0$,*

$$\text{Ent}_\nu(\rho_T) - \text{Ent}_\nu(\rho_0) + \frac{1}{2} \int_0^T \mathcal{A}(\rho_t) dt + \frac{1}{2} \int_0^T \mathcal{A}(\rho_t, \psi_t) dt \geq 0, \quad (6)$$

with equality if and only if $\{\rho_t\}_{t \geq 0}$ is the flow of the Markov process on \mathcal{X} with kernel K and invariant measure ν , solution to (1).

This is the analogue of the characterization of solutions to $\dot{x}_t = -\nabla V(x_t)$ on \mathbb{R}^d as the only curves for which the non-negative functional

$$V(x_T) - V(x_0) + \frac{1}{2} \int_0^T |\nabla V(x_t)|^2 dt + \frac{1}{2} \int_0^T |\dot{x}_t|^2 dt$$

cancels. Hence in the framework of Markov chains, Ent_ν plays the role of V , and the entropy production \mathcal{I} plays the role of $|\nabla V|^2$.

Proof (of Proposition 2.2) Denote $\mathcal{H}(\rho) = \text{Ent}_\nu(\rho)$. We have

$$\begin{aligned} \mathcal{H}(\rho_T) - \mathcal{H}(\rho_0) &= \int_0^T \frac{d}{dt} (\mathcal{H}(\rho_t)) dt = \int_0^T \sum_{x \in \mathcal{X}} \nu(x) \frac{d}{dt} (\rho_t(x) \log \rho_t(x)) dt \\ &= \int_0^T \sum_{x \in \mathcal{X}} \nu(x) \dot{\rho}_t(x) \log \rho_t(x) dt. \end{aligned}$$

Using the reversibility of the invariant measure ν , we write

$$\begin{aligned} \mathcal{H}(\rho_T) - \mathcal{H}(\rho_0) &= - \int_0^T \sum_{x, y \in \mathcal{X}} \nu(x) (\psi_t(y) - \psi_t(x)) K(x, y) \hat{\rho}_t(x, y) \log \rho_t(x) dt \\ &= \int_0^T \frac{1}{2} \sum_{x, y \in \mathcal{X}} (\psi_t(y) - \psi_t(x)) (\log \rho_t(y) - \log \rho_t(x)) \hat{\rho}_t(x, y) K(x, y) \nu(x) dt \\ &\geq -\frac{1}{4} \int_0^T \sum_{x, y \in \mathcal{X}} (\psi_t(x) - \psi_t(y))^2 \hat{\rho}_t(x, y) K(x, y) \nu(x) dt \\ &\quad - \frac{1}{4} \int_0^T \sum_{x, y \in \mathcal{X}} (\log \rho_t(x) - \log \rho_t(y))^2 \hat{\rho}_t(x, y) K(x, y) \nu(x) dt \\ &= -\frac{1}{2} \int_0^T \mathcal{A}(\rho_t, \psi_t) dt - \frac{1}{2} \int_0^T \mathcal{I}(\rho_t) dt \end{aligned}$$

with equality if and only if, for all $x, y \in \mathcal{X}$ and almost every $t \in [0, T]$, we have

$$\psi_t(x) - \psi_t(y) = \log \rho_t(y) - \log \rho_t(x)$$

which is equivalent to saying that for almost every t and for every x we have

$$\dot{\rho}_t(x) + \sum_{y \in \mathcal{X}} (\log \rho_t(x) - \log \rho_t(y)) \hat{\rho}_t(x, y) K(x, y) = 0.$$

3 Scaling Limits and Gradient Flows

With the formulation of Proposition 2.2, we can use the approach of Sandier and Serfaty [19] to study convergence of sequences of Markov chains to a scaling limit. Let (K_n) be a sequence of reversible Markov kernels on finite spaces \mathcal{X}_n , and let (ν_n) be the sequence of invariant measures on \mathcal{X}_n . Since we wish to investigate the asymptotic behaviour of the sequence of random processes, it is much more convenient to work in a single space \mathcal{X} that contains all the \mathcal{X}_n . Hence we shall assume that we are given a space \mathcal{X} and a collection of embeddings $\mathbf{p}_n : \mathcal{X}_n \rightarrow \mathcal{X}$. In practice, the choice of \mathcal{X} and \mathbf{p}_n is suggested by the model under investigation. In the next section, which is focused on the simple exclusion process on the torus, the embeddings will map a configuration η onto the associated empirical measure $\pi^n(\eta)$ (see (12), Sect. 4.1). Such embeddings immediately define embeddings of $\mathcal{P}(\mathcal{X}_n)$ into $\mathcal{P}(\mathcal{X})$.

In order to simplify the exposition below, we adopt the following convention: whenever we say that a sequence (x_n) of elements of \mathcal{X}_n converges to $x \in \mathcal{X}$, we shall mean that $\mathbf{p}_n(x_n) \rightarrow x$ as n goes to infinity. In particular, the topology used for convergence is implicitly the topology of \mathcal{X} , which is assumed to be a separable complete metric space. The strategy is to characterize possible candidates for the limit as gradient flows. For that purpose we give a definition of minimizing-movement curves in the metric setting:

Definition 3.1 Let (\mathcal{X}, d) be a complete metric space. The *gradient flows* of an energy functional $\mathcal{H} : \mathcal{X} \rightarrow \mathbb{R}$ with respect to the metric d are the curves $\{m_t\}$ s.t.

$$\mathcal{H}(m_T) - \mathcal{H}(m_0) + \frac{1}{2} \int_0^T g(m_t) dt + \frac{1}{2} \int_0^T |\dot{m}_t|^2 dt = 0,$$

where g is the local slope for \mathcal{H} , defined as

$$g(m) := \limsup_{\tilde{m} \rightarrow m} \frac{\mathcal{H}(m) - \mathcal{H}(\tilde{m})}{d(m, \tilde{m})} \quad \text{and} \quad |\dot{m}_t| = \limsup_{h \rightarrow 0} \frac{d(m_t, m_{t+h})}{h}.$$

Remark 3.1 This is not a complete definition. To make it correct, we should introduce the notion of *absolutely continuous curves*, whose slopes are well defined. This is not a real issue here, since we shall only use it for reversible Markov chains (for which the notions have already been well defined previously for curves of strictly positive probability measures) and the heat equation, for which smooth curves of strictly positive functions do not cause any issue (see the next example). We refer to [3, 19] for a more rigorous discussion of the issues in the metric setting.

Example 3.1 (Heat equation) Let us consider the parabolic PDE

$$\frac{\partial m}{\partial t} = \frac{\partial^2 m}{\partial \theta^2}, \quad t \geq 0, \theta \in (0, 1). \quad (7)$$

We know from [2] that (7) is associated to a gradient flow, since we have:

$$\int_0^1 h(m(T, \theta))d\theta - \int_0^1 h(m(0, \theta))d\theta + \frac{1}{2} \int_0^T \int_0^1 m(1-m) \left(\frac{\partial(h'(m))}{\partial\theta} \right)^2 d\theta dt + \frac{1}{2} \int_0^T \left\| \frac{\partial m}{\partial t} \right\|_{-1,m}^2 dt = 0,$$

with $h(x) = x \log x + (1-x) \log(1-x)$ and, given $u : [0, 1] \rightarrow \mathbb{R}$,

$$\|u\|_{-1,m}^2 := \sup_J \left\{ 2 \int_0^1 J(\theta)u(\theta)d\theta - \int_0^1 m(1-m)(J'(\theta))^2 d\theta \right\},$$

where the supremum is over all smooth test functions J .

We now state the main result of that section. Hereafter, when we assert that a sequence of curves of probability measures (in $\mathcal{P}(\mathcal{X}_n) \hookrightarrow \mathcal{P}(\mathcal{X})$) converges to a deterministic curve $\{m_t\}$ (in \mathcal{X}), we mean that it converges to a curve of Dirac measures $\{\delta_{m_t}\}$. Definition 4.1 below gives a more precise meaning in the case of particle systems. It is important to stress that we study convergence of probability measures, which are deterministic objects.

Theorem 3.1 *Let (a_n) be an increasing diverging sequence of positive numbers. We first assume that the topology on $\mathcal{P}(\mathcal{X})$ has the following property:*

(P) *For any sequence $(\rho_t^n v_n)$ of smooth curves of positive probability measures that converges to some deterministic curve $\{m_t\}$, the following inequalities hold:*

$$\liminf_{n \rightarrow \infty} \frac{1}{a_n} \text{Ent}_{v_n}(\rho_T^n) \geq \mathcal{H}(m_T) \tag{8}$$

$$\liminf_{n \rightarrow \infty} \frac{1}{a_n} \int_0^T \mathcal{I}_n(\rho_t^n) dt \geq \int_0^T g(m_t) dt \tag{9}$$

$$\liminf_{n \rightarrow \infty} \frac{1}{a_n} \int_0^T \mathcal{A}_n(\rho_t^n, \psi_t^n) dt \geq \int_0^T |\dot{m}_t|^2 dt, \tag{10}$$

where ψ_t^n is such that (ρ_t^n, ψ_t^n) solves (5).

Now consider a sequence $(\rho_t^n v_n)$ of gradient flows (so that there is equality in (6)), assume that the initial sequence $(\rho_0^n v_n)$ does converge in distribution to some m_0 , and that moreover

$$\lim_{n \rightarrow \infty} \frac{1}{a_n} \text{Ent}_{v_n}(\rho_0^n) = \mathcal{H}(m_0).$$

Then, any possible weak limit $\{m_t\}$ of $(\rho_t^n v_n)$ is almost surely a gradient flow of the energy \mathcal{H} , starting from m_0 . In particular, if gradient flows starting from a given initial data are unique, $(\rho_t^n v_n)$ weakly converges to a Dirac measure concentrated on the unique gradient flow of \mathcal{H} starting from m_0 .

Moreover, for any $t \in [0, T]$, we have

$$\frac{1}{a_n} \text{Ent}_{v_n}(\rho_t^n) \xrightarrow{n \rightarrow \infty} \mathcal{H}(m_t).$$

Above (a_n) is a sequence of weights that corresponds to the correct scaling of the system. For particle systems on the discrete torus of length n in dimension d under diffusive scaling, we would take $a_n = n^d$. This result is a slight variation of the abstract method developed in [19], to which we refer for more details. The main difference (apart from the setting which is restricted to gradient flows in spaces of probability measures arising from reversible Markov chains) is that we consider curves of probability measures that converge to a deterministic curve, rather than any possible limit.

One of the interesting features of this technique is that it does not require an assumption of uniform semi-convexity on the sequence of relative entropies, which can be hard to establish for interacting particle systems (see [7] for the general theory of geodesic convexity of the entropy for Markov chains, and [9] for the study of this property for interacting particle systems on the complete graph). Such an assumption of semi-convexity is known as a lower bound on Ricci curvature for the Markov chain, by analogy with the situation for Brownian motion on a Riemannian manifold. For the simple exclusion on the discrete torus, it seems reasonable to conjecture that curvature is non-negative, but this is still an unsolved problem.

Proof First of all, for any weak limit \mathcal{Q} of the laws of the trajectories, we also have

$$\liminf_{n \rightarrow \infty} \frac{1}{a_n} \int_0^T \mathcal{E}(\rho_t^n, \log \rho_t^n) dt \geq \mathcal{Q} \left[\int_0^T g(m_t) dt \right]$$

and

$$\liminf_{n \rightarrow \infty} \frac{1}{a_n} \int_0^T \mathcal{A}(\rho_t^n, \psi_t^n) dt \geq \mathcal{Q} \left[\int_0^T |\dot{m}_t|^2 dt \right],$$

where we denote by $\{m_t\}$ a random trajectory with law \mathcal{Q} . This is a direct consequence of the following lemma (whose proof is given below):

Lemma 3.2 *Let (f_n) be a sequence of real-valued, non-negative functions on a space (Ω, \mathbb{P}) , and assume that there exists a function f such that for any sequence of random variables (X_n) that converges in law to a deterministic limit x , we have*

$$\liminf_{n \rightarrow \infty} \mathbb{E}[f_n(X_n)] \geq f(x).$$

Then, for any sequence (X_n) of random variables that converges in law to a random variable X_∞ , we have

$$\liminf_{n \rightarrow \infty} \mathbb{E}[f_n(X_n)] \geq \mathbb{E}[f(X_\infty)].$$

We now use Proposition 2.2 with the gradient flows $\{\rho_t^n \nu_n\}$, and pass to the limit in

$$\frac{1}{a_n} \left(\text{Ent}_{\nu_n}(\rho_T^n) - \text{Ent}_{\nu_n}(\rho_0^n) + \frac{1}{2} \int_0^T \mathcal{E}(\rho_t^n, \log \rho_t^n) dt + \frac{1}{2} \int_0^T \mathcal{A}(\rho_t^n, \psi_t^n) dt \right) = 0,$$

and therefore

$$\mathcal{Q}[\mathcal{H}(m_T)] - \mathcal{H}(m_0) + \frac{1}{2} \mathcal{Q} \left[\int_0^T g(m_t) dt \right] + \frac{1}{2} \mathcal{Q} \left[\int_0^T |\dot{m}_t|^2 dt \right] \leq 0. \quad (11)$$

Since the above quantity is an expectation of a non-negative functional, we see that

$$\mathcal{H}(m_T) - \mathcal{H}(m_0) + \frac{1}{2} \int_0^T g(m_t) dt + \frac{1}{2} \int_0^T |\dot{m}_t|^2 dt = 0, \quad \mathcal{Q}\text{-almost surely.}$$

This means that \mathcal{Q} -almost surely, $\{m_t\}$ is a gradient flow of \mathcal{H} . If uniqueness of gradient flows with initial condition m_0 holds, convergence immediately follows.

Convergence of the relative entropy at time T necessarily holds, since otherwise it would contradict (11). Finally, it also holds at any other time $t \in [0, T]$, since one can rewrite the same result on the time-interval $[0, t]$.

We still have to prove Lemma 3.2. This proof is taken from [6].

Proof (of Lemma 3.2) Consider a sequence (X_n) that converges in law to a random variable X_∞ . Using the almost-sure representation theorem, there exists a sequence (Y_n) such that for any n , Y_n has the same law as X_n , and (Y_n) almost surely converges to Y_∞ . If we condition the whole sequence on the event $\{Y_\infty = y\}$, then (Y_n) almost surely converges to y . Then we have, using Fatou's lemma

$$\begin{aligned} \liminf_{n \rightarrow \infty} \mathbb{E}[f_n(X_n)] &= \liminf_{n \rightarrow \infty} \mathbb{E}[f_n(Y_n)] = \liminf_{n \rightarrow \infty} \mathbb{E}_{Y_\infty} [\mathbb{E}[f_n(Y_n) | Y_\infty]] \\ &\geq \mathbb{E}_{Y_\infty} \left[\liminf_{n \rightarrow \infty} \mathbb{E}[f_n(Y_n) | Y_\infty] \right] \geq \mathbb{E}_{Y_\infty} [f(Y_\infty)] = \mathbb{E}[f(X_\infty)]. \end{aligned}$$

4 Symmetric Simple Exclusion Process (SSEP)

4.1 Model: Definitions and Notations

To make notations easier we consider the interacting particle systems on the one-dimensional torus $\mathbb{T}_n = \{0, \dots, n-1\}$, but the result is valid in any dimension

$d \geq 1$. Let us define $\mathcal{X}_n := \{0, 1\}^{\mathbb{T}_n}$, $\mathcal{X} := \{0, 1\}^{\mathbb{Z}}$, and $\mathbb{T} = [0, 1)$ the continuous torus. We create a Markov process $\{\eta_t^n ; t \geq 0\}$ on the state space \mathcal{X}_n , which satisfies for any $\eta \in \mathcal{X}_n$:

- $\eta(i) = 1$ if there is a particle at site $i \in \mathbb{T}_n$,
- $\eta(i) = 0$ if the site i is empty,
- any particle waits independently an exponential time and then jumps to one of its neighbouring sites with probability $1/2$, provided that the chosen site is empty.

We are looking at the evolution of the Markov process in the diffusive time scale, meaning that time is accelerated by n^2 . The generator is given for $f : \mathcal{X}_n \rightarrow \mathbb{R}$ by

$$\mathcal{L}_n(f)(\eta) := n^2 \sum_{i \in \mathbb{T}_n} \eta(i)(1 - \eta(i + 1))(f(\eta^{i,i+1}) - f(\eta)),$$

where $\eta^{i,j}$ is the configuration obtained from η exchanging the occupation variables $\eta(i)$ and $\eta(j)$. The hydrodynamics behavior of the SSEP is well-known, and we refer the reader to [11] for a survey. Let ν_α^n be the Bernoulli product measure of parameter $\alpha \in (0, 1)$, the invariant measures for the dynamics. Under ν_α^n , the variables $\{\eta(i)\}_{i \in \mathbb{T}_n}$ are independent with marginals given by

$$\nu_\alpha^n\{\eta(i) = 1\} = \alpha = 1 - \nu_\alpha^n\{\eta(i) = 0\}.$$

Let us fix once and for all $\alpha \in (0, 1)$ and denote by ρ_t^n the probability density of the law of η_t^n (whose time evolution is generated by $n^2 \mathcal{L}_n$) with respect to ν_α^n .

To prove convergence, we need to embed our particle configurations in a single metric space. For each configuration $\eta \in \mathcal{X}_n$, we construct a measure on \mathcal{X} associated to η , denoted by $\pi^n(\eta)$. We do it here through the *empirical measures*:

$$\pi_t^n(d\theta) := \pi^n(\eta_t^n)(d\theta) = n^{-1} \sum_{i \in \mathbb{T}_n} \eta_t^n(i) \delta_{i/n}(d\theta), \tag{12}$$

where δ_θ stands for the Dirac measure concentrated on $\theta \in \mathbb{T}$. Let us denote by $\mathcal{M}_+ = \mathcal{M}_+(\mathcal{X})$ the space of finite positive measures on \mathcal{X} endowed with the weak topology. Assume moreover that, for each n , $\pi^n : \mathcal{X}_n \rightarrow \mathcal{M}_+$ is a continuous function. Our goal is to prove the convergence of the flow of measures $(\pi^n(\eta_t^n))$. In particular, π^n inherits the Markov property from η^n .

We start by defining properly two notions of convergence. For any function $G : \mathcal{X} \rightarrow \mathbb{R}$ and any measure π on \mathcal{X} , we denote by $\langle \pi, G \rangle$ the integral of G with respect to the measure π . In the following $T > 0$ is fixed.

Definition 4.1 Let (π_t^n) be a sequence of flows of measures, each element belonging to the Skorokhod space $\mathcal{D}([0, T], \mathcal{M}_+)$. For each n , let \mathcal{Q}_n be the probability measure on $\mathcal{D}([0, T], \mathcal{M}_+)$ corresponding to $\{\pi_t^n ; t \in [0, T]\}$.

1. We say that the sequence (π_t^n) converges to the deterministic flow $\{\pi_t\}$ if the probability measure \mathcal{Q}_n converges to the Dirac probability measure concentrated on the deterministic flow $\{\pi_t\}$.
2. Fix $t \in [0, T]$. We say that (π_t^n) converges in probability to the deterministic measure $\pi_t \in \mathcal{M}_+$ if, for all smooth test functions $G : \mathcal{X} \rightarrow \mathbb{R}$, and all $\delta > 0$,

$$\mathcal{Q}_n \left[\left| \langle \pi_t^n, G \rangle - \langle \pi_t, G \rangle \right| > \delta \right] \xrightarrow[n \rightarrow \infty]{} 0. \tag{13}$$

The next proposition gives the equivalence between the two notions above.

Proposition 4.1 *Let (π_t^n) be a sequence of flows of measures which converges to a deterministic flow $\{\pi_t\}$. Assume that $t \in [0, T] \mapsto \pi_t \in \mathcal{M}_+$ is continuous (with respect to the weak topology). Then, for any $t \in [0, T]$ fixed, (π_t^n) converges in probability to $\pi_t \in \mathcal{M}_+$.*

Proof By assumption, the limiting probability measure on $\mathcal{D}([0, T], \mathcal{M}_+)$ is concentrated on weakly continuous trajectories. Therefore, the limiting flow is almost surely continuous, and the map $\{\pi_t ; t \in [0, T]\} \mapsto \pi_t$ is continuous from $\mathcal{D}([0, T], \mathcal{M}_+)$ to \mathcal{M}_+ . Then, for $t \in [0, T]$, (π_t^n) converges in distribution to π_t . Since the latter is deterministic, this induces convergence in probability.

We recall here the main result, that we are going to prove in a different way. Recall that π_t^n is the empirical measure defined in (12) and \mathcal{Q}_n is the probability measure on $\mathcal{D}([0, T], \mathcal{M}_+)$ corresponding to the flow $\{\pi_t^n\}$.

Theorem 4.2 (Hydrodynamic limits for the SSEP) *Fix a density profile $m_0 : \mathbb{T} \rightarrow [0, 1]$ and let (μ^n) be a sequence of probability measures such that, under μ^n , the sequence $(\pi_0^n(d\theta))$ converges in probability to $m_0(\theta)d\theta$. In other words,*

$$\limsup_{n \rightarrow \infty} \mu^n \left[\left| n^{-1} \sum_{i \in \mathbb{T}_n} G(i/n) \eta(i) - \int_{\mathbb{T}} G(\theta) m_0(\theta) d\theta \right| > \delta \right] = 0,$$

for any $\delta > 0$ and any smooth function $G : \mathbb{T} \rightarrow \mathbb{R}$. Assume moreover that this initial data is well-prepared, in the sense that:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \text{Ent}_{\nu_\alpha^n}(\mu^n) = \int_{\mathbb{T}} h(m_0(\theta)) d\theta - h\left(\int_{\mathbb{T}} m_0(\theta) d\theta \right), \tag{14}$$

where h has been defined in Example 3.1. Then, for any $t > 0$, the sequence $\{\pi_t^n\}_{n \in \mathbb{N}}$ converges in probability to the deterministic measure $\pi_t(d\theta) = m(t, \theta)d\theta$ where m is solution to the heat equation (7) on $\mathbb{R}_+ \times \mathbb{T}$. The entropy also converges:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \text{Ent}_{\nu_\alpha^n}(\mu_t^n) = \int_{\mathbb{T}} h(m(t, \theta)) d\theta - h\left(\int_{\mathbb{T}} m(t, \theta) d\theta \right).$$

Note that $\int m(t, \theta)d\theta$ is actually a constant, given by the fixed density of particles. The convergence of the entropy is equivalent to the local Gibbs behavior (see [12]). Hence, the assumptions and conclusions are those obtained with the relative entropy method of [21]. However, the techniques and restrictions are the same as for the entropy method of [10]: we do not use smoothness of solutions to the hydrodynamic PDE, but we use the replacement lemma (see Sect. 4.3), which relies on the two-block estimate, rather than the one-block estimate alone as in [21].

4.2 The Gradient Flow Approach to Theorem 4.2

We are going to apply Theorem 3.1 to obtain Theorem 4.2. The main steps are as follows:

1. We first need to prove that the sequence (\mathcal{Q}_n) is relatively compact, so that there exists a converging subsequence. Such an argument was already part of the entropy method of [10]. We refer to [11][Chap.4, Sect. 2] for the proof in the context of the simple exclusion process.
2. In order to prove (8), we have to investigate the convergence of the relative entropy with respect to the invariant measure ν_α^n towards the free energy associated to the limiting PDE (7), which in our case reads as

$$\mathcal{H}(m) = \int_{\mathbb{T}} h(m(\theta))d\theta - h\left(\int_{\mathbb{T}} m(\theta)d\theta\right).$$

This result is actually equivalent to the large deviation principle for ν_α^n (see for example [15]), and is standard (see [11]). Moreover, if our initial data is close (in relative entropy) to a slowly varying Bernoulli product measure¹ associated to m , which satisfies $\nu_{\rho(\cdot)}^n\{\eta(i) = 1\} = m(i/n)$, then its relative entropy with respect to ν_α^n converges to the limiting free energy, so that we can easily have (14).

3. We prove the lower bound for the entropy production along curves (9) and the lower bound for the slopes (10) in Sect. 4.3.
4. When passing to the limit, we obtain that for any weak limit \mathcal{Q} of (\mathcal{Q}_n) ,

$$\mathcal{Q}\left[\int_{\mathbb{T}} h(m(T, \theta))d\theta - \int_{\mathbb{T}} h(m(0, \theta))d\theta + \frac{1}{2} \int_0^T \int_{\mathbb{T}} m(1-m)\left(\frac{\partial(h'(m))}{\partial\theta}\right)^2 d\theta dt + \frac{1}{2} \int_0^T \|\dot{m}_t\|_{-1,m}^2 dt\right] \leq 0.$$

¹This is also the assumption used to make Yau’s relative entropy method work, see [21].

Since the expression inside the expectation is the characterization of solutions to the heat equation as minimizing-movement curves, it is non-negative, and almost surely m is a solution to the heat equation. Uniqueness of solutions starting from m_0 allows us to conclude.

4.3 Bounds and Convergence

Here we prove that (9) and (10) are satisfied for the density ρ_t^n of the SSEP accelerated in time, assuming that the empirical measure $(\pi_t^n(d\theta))$ converges to a deterministic curve $m_t(\theta)d\theta$. Let us start with (10). The argument is based on a duality argument (Proposition 4.3) and on the replacement lemma (Lemma 4.4) which is commonly used in the literature (see for example [11]).

Proposition 4.3 *Consider a couple (ρ_t, ψ_t) satisfying the continuity equation (5) for almost every $t \geq 0$. For any smooth (in time) function $J : [0, T] \times \mathcal{X} \rightarrow \mathbb{R}$,*

$$\begin{aligned} \int_0^T \mathcal{A}(\rho_t, \psi_t) dt &\geq 2 \sum_{x \in \mathcal{X}} J(T, x) \rho_T(x) v(x) - 2 \sum_{x \in \mathcal{X}} J(0, x) \rho_0(x) v(x) \\ &- 2 \int_0^T \sum_{x \in \mathcal{X}} \partial_t J(t, x) \rho_t(x) v(x) dt - \int_0^T \sum_{x, y} (J(t, x) - J(t, y))^2 \hat{\rho}_t(x, y) K(x, y) v(x) dt. \end{aligned}$$

Proof From the continuity equation (5), we have

$$\begin{aligned} &\sum_{x \in \mathcal{X}} J(T, x) \rho_T(x) v(x) - \sum_{x \in \mathcal{X}} J(0, x) \rho_0(x) v(x) \\ &= \int_0^T \sum_{x \in \mathcal{X}} \partial_t J(t, x) \rho_t(x) v(x) + J(t, x) \dot{\rho}_t(x) v(x) dt = \int_0^T \sum_{x \in \mathcal{X}} \partial_t J(t, x) \rho_t(x) v(x) dt \\ &\quad - \int_0^T \sum_{x, y \in \mathcal{X}} J(t, x) (\psi_t(y) - \psi_t(x)) \hat{\rho}_t(x, y) K(x, y) v(x) dt. \end{aligned}$$

We symmetrize in x and y the last term, and get

$$\begin{aligned} 2 \sum_{x \in \mathcal{X}} J(T, x) \rho_T(x) v(x) - 2 \sum_{x \in \mathcal{X}} J(0, x) \rho_0(x) v(x) &= 2 \int_0^T \sum_{x \in \mathcal{X}} \partial_t J(t, x) \rho_t(x) v(x) dt \\ &- \int_0^T \sum_{x, y \in \mathcal{X}} (J(t, x) - J(t, y)) (\psi_t(y) - \psi_t(x)) \hat{\rho}_t(x, y) K(x, y) v(x) dt, \end{aligned}$$

and therefore

$$\begin{aligned} \frac{1}{2n} \int_0^T \mathcal{A}(\rho_t, \psi_t) dt &\geq \frac{1}{n} \sum_{x \in \mathcal{X}} J(T, x) \rho_T(x) v(x) - \frac{1}{n} \sum_{x \in \mathcal{X}} J(0, x) \rho_0(x) v(x) \\ &- \frac{1}{n} \int_0^T \sum_{x \in \mathcal{X}} \partial_t J(t, x) \rho_t(x) v(x) dt - \frac{1}{2n} \int_0^T \sum_{x, y} (J(t, x) - J(t, y))^2 \hat{\rho}_t(x, y) K(x, y) v(x) dt. \end{aligned}$$

To apply Proposition 4.3 to the SSEP, we consider observables of the form

$$J(t, \eta) = \sum_{i \in \mathbb{T}_n} G\left(t, \frac{i}{n}\right) \eta(i) \tag{15}$$

for smooth functions $G : [0, T] \times \mathbb{T} \rightarrow \mathbb{R}$. For any $\ell \in \mathbb{N}$ and $i \in \mathbb{T}_n$, we denote by $\eta^\ell(i)$ the empirical density of particles in a box of size $2\ell + 1$ centered at i :

$$\eta^\ell(i) := \frac{1}{2\ell + 1} \sum_{|j-i| \leq \ell} \eta(j).$$

Hereafter we also denote by τ_x the translated operator that acts on local functions $g : \{0, 1\}^{\mathbb{Z}} \rightarrow \mathbb{R}$ as $(\tau_x g)(\eta) := g(\tau_x \eta)$, and $\tau_x \eta$ is the configuration obtained from η by shifting: $(\tau_x \eta)_y = \eta_{x+y}$. The main tool that we are going to use is the well-known replacement lemma, which is a consequence of the averaging properties of the SSEP. We recall the main statement and refer the reader to [10, 11] for a proof:

Lemma 4.4 (Replacement Lemma) *Denote by \mathbb{P}_{μ^n} the probability measure on the Skorokhod space $\mathcal{D}([0, T], \mathcal{X}_n)$ induced by the Markov process $\{\eta_t^n\}_{t \geq 0}$ starting from μ^n . Then, for every $\delta > 0$ and every local function g ,*

$$\limsup_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}_{\mu^n} \left[\int_0^T n^{-1} \sum_{x \in \mathbb{T}_n} \tau_x V_{\varepsilon n}(\eta_s) ds \geq \delta \right] = 0,$$

where

$$V_\ell(\eta) = \left| \frac{1}{2\ell + 1} \sum_{|y| \leq \ell} \tau_y g(\eta) - \tilde{g}(\eta^\ell(0)) \right|$$

and $\tilde{g} : (0, 1) \rightarrow \mathbb{R}$ corresponds to the expected value: $\tilde{g}(\alpha) := \int g(\eta) d\nu_\alpha(\eta)$.

We are now able to conclude the proof. We treat separately the terms in the right-hand side of Proposition 4.3, taking J as in (15). Since, for any fixed t , $(\pi_t^n(d\theta))$ converges in probability to $\pi_t(d\theta) = m_t(\theta)d\theta$ we have

$$\frac{1}{n} \sum_{\eta \in \mathcal{X}_n} J(T, \eta) \rho_T^n(\eta) v_\alpha^n(\eta) = \mathcal{Q}_n[\langle \pi_T^n, G \rangle] \xrightarrow{n \rightarrow \infty} \int_{\mathbb{T}} G(T, \theta) m_T(\theta) d\theta.$$

And the same happens at initial time for ρ_0^n . Similarly,

$$\frac{1}{n} \int_0^T \sum_{\eta \in \mathcal{X}_n} \frac{\partial J}{\partial t}(t, \eta) \rho_t^n(\eta) v_\alpha^n(\eta) dt \longrightarrow \int_0^T \int_{\mathbb{T}} \frac{\partial G}{\partial t}(t, \theta) m_t(\theta) d\theta dt.$$

Then, we write

$$\begin{aligned} & \int_0^T \frac{1}{n} \sum_{\eta, \eta' \in \mathcal{X}_n} (J(t, \eta) - J(t, \eta'))^2 \widehat{\rho}_t^n(\eta, \eta') K_n(\eta, \eta') v_\alpha^n(\eta) dt \\ &= \frac{n^2}{n} \int_0^T \sum_{\eta, i} \left[G\left(t, \frac{i}{n}\right) - G\left(t, \frac{i+1}{n}\right) \right]^2 \eta(i)(1 - \eta(i+1)) \widehat{\rho}_t^n(\eta, \eta^{i, i+1}) v_\alpha^n(\eta) dt. \end{aligned}$$

We now use the logarithmic inequality (3) and write that the latter is smaller than

$$n \int_0^T \sum_{\eta, i} \left[G\left(t, \frac{i}{n}\right) - G\left(t, \frac{i+1}{n}\right) \right]^2 \eta(i)(1 - \eta(i+1)) \frac{\rho_t^n(\eta) + \rho_t^n(\eta^{i, i+1})}{2} v_\alpha^n(\eta) dt.$$

From the invariance property of v_α^n with respect to the change of variables $\eta \rightarrow \eta^{i, i+1}$, and from the smoothness of G we get that the above quantity is equal to

$$\frac{1}{n} \int_0^T \sum_{\eta \in \mathcal{X}_n} \sum_{i \in \mathbb{T}_n} \left[G'\left(t, \frac{i}{n}\right) \right]^2 \eta(i)(1 - \eta(i+1)) \rho_t^n(\eta) v_\alpha^n(\eta) dt + o\left(\frac{1}{n}\right), \quad (16)$$

where G' denotes the space derivative of G .

Above we want to replace $\eta(i)(1 - \eta(i+1))$ by $m(i/n)(1 - m(i/n))$. For $\varepsilon > 0$ we define the approximation of the identity $i_\varepsilon(u) = (2\varepsilon)^{-1} \mathbf{1}\{|u| \leq \varepsilon\}$. With that notation, $\eta_t^{\varepsilon n}(0)$ is very close to $\langle \pi_t^n, i_\varepsilon \rangle$. Let us denote $h(\eta) := \eta(0)(1 - \eta(1))$. Since G is a smooth function, (16) equals

$$\frac{1}{n} \int_0^T \sum_{\eta \in \mathcal{X}_n} \sum_{i \in \mathbb{T}_n} \frac{1}{2\varepsilon n + 1} \sum_{|j-i| \leq \varepsilon n} \left[G'\left(t, \frac{j}{n}\right) \right]^2 \tau_i h(\eta) \rho_t^n(\eta) v_\alpha^n(\eta) dt + O(\varepsilon^2).$$

A summation by parts shows that the previous term can be written as

$$\frac{1}{n} \int_0^T \sum_{\eta \in \mathcal{X}_n} \sum_{i \in \mathbb{T}_n} \left[G'\left(t, \frac{i}{n}\right) \right]^2 \frac{1}{2\varepsilon n + 1} \sum_{|j-i| \leq \varepsilon n} \tau_j h(\eta) \rho_t^n(\eta) v_\alpha^n(\eta) dt + O(\varepsilon^2).$$

By Lemma 4.4, this expression is then equal to

$$\frac{1}{n} \int_0^T \sum_{\eta \in \mathcal{X}_n} \sum_{i \in \mathbb{T}_n} \left[G'\left(t, \frac{i}{n}\right) \right]^2 \tau_i \tilde{h}(\langle \pi_t^n(\eta), i_\varepsilon \rangle) v_\alpha^n(\eta) dt + R_{n, \varepsilon, T},$$

where $R_{n,\varepsilon,T}$ vanishes in probability as n goes to infinity and then ε goes to 0. From the convergence in probability of (π_t^n) , the last expression converges to

$$\int_0^T \int_{\mathbb{T}} m_t(\theta)(1 - m_t(\theta))(G'(t, \theta))^2 d\theta dt.$$

As a result, since the convergences above are valid for any smooth function G ,

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} \int_0^T \mathcal{A}(\rho_t^n, \psi_t^n) dt &\geq \int_0^T \sup_G \left\{ 2 \int_{\mathbb{T}} G \dot{m}_t d\theta - \int_{\mathbb{T}} m_t(1 - m_t)(G')^2 d\theta \right\} dt \\ &= \int_0^T \|\dot{m}_t\|_{-1,m}^2 dt. \end{aligned}$$

In the same way, we need to prove

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \int_0^T \mathcal{E}(\rho_t^n, \log \rho_t^n) dt \geq \int_0^T \int_{\mathbb{T}} m(1 - m) \left(\frac{\partial(h'(m))}{\partial \theta} \right)^2 d\theta dt.$$

Since the arguments are essentially the same as for the slopes, we shall be more brief in the exposition. We denote

$$\nabla_n G\left(t, \frac{i}{n}\right) = n \left[G\left(t, \frac{i+1}{n}\right) - G\left(t, \frac{i}{n}\right) \right].$$

By duality, we have

$$\begin{aligned} &\frac{1}{2n} \int_0^T \mathcal{E}(\rho_t^n, \log \rho_t^n) dt \\ &\geq \int_0^T \sum_{\eta,i} (\log \rho_t^n(\eta^{i,i+1}) - \log \rho_t^n(\eta)) \nabla_n G\left(t, \frac{i}{n}\right) \eta_i (1 - \eta_{i+1}) \widehat{\rho}_t^n(\eta, \eta^{i,i+1}) v_\alpha^n(\eta) dt \\ &\quad - \frac{1}{2n} \int_0^T \sum_{\eta,i} \left[\nabla_n G\left(t, \frac{i}{n}\right) \right]^2 \eta_i (1 - \eta_{i+1}) \widehat{\rho}_t^n(\eta, \eta^{i,i+1}) v_\alpha^n(\eta) dt \\ &\geq \int_0^T \sum_{\eta,i} (\rho_t^n(\eta^{i,i+1}) - \rho_t^n(\eta)) \nabla_n G\left(t, \frac{i}{n}\right) \eta_i (1 - \eta_{i+1}) v_\alpha^n(\eta) dt \\ &\quad - \frac{1}{2n} \int_0^T \sum_{\eta,i} \left[\nabla_n G\left(t, \frac{i}{n}\right) \right]^2 \eta_i (1 - \eta_{i+1}) \rho_t^n(\eta) v_\alpha^n(\eta) dt \\ &= - \int_0^T \sum_{\eta,i} \rho_t^n(\eta) \left[\nabla_n G\left(t, \frac{i}{n}\right) - \nabla_n G\left(t, \frac{i-1}{n}\right) \right] \eta_i v_\alpha^n(\eta) dt \\ &\quad - \frac{1}{2n} \int_0^T \sum_{\eta,i} \left[\nabla_n G\left(t, \frac{i}{n}\right) \right]^2 \eta_i (1 - \eta_{i+1}) \rho_t^n(\eta) v_\alpha^n(\eta) dt \end{aligned}$$

Using the replacement lemma, passing to the supremum in G , and to the limit,

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{n} \int_0^T \mathcal{E}(\rho_t^n, \log \rho_t^n) dt \\ & \geq \sup_G \left\{ -2 \int_0^T \int_{\mathbb{T}} G'' m_t \, d\theta dt - \int_0^T \int_{\mathbb{T}} m_t (1 - m_t) (G')^2 d\theta dt \right\} \\ & = \sup_G \left\{ -2 \int_0^T \int_{\mathbb{T}} G'' m_t (1 - m_t) h''(m_t) d\theta dt - \int_0^T \int_{\mathbb{T}} m_t (1 - m_t) (G')^2 d\theta dt \right\} \\ & = \int_0^T \int_{\mathbb{T}} m_t (1 - m_t) \left(\frac{\partial h'(m_t)}{\partial \theta} \right)^2 d\theta dt, \end{aligned}$$

and this is exactly what we were seeking to prove.

Acknowledgments M.F. : I would like to thank Hong Duong, Matthias Erbar, Vaios Laschos and André Schlichting for discussions on convergence of gradient flows. Part of this work was done while I was staying at the Hausdorff Institute for Mathematics in Bonn, whose support is gratefully acknowledged. I also benefited from funding from GDR MOMAS and from NSF FRG grant DMS-1361185. M.S.: This work has been supported by the French Ministry of Education through the grant ANR (EDNHS), and also by CAPES (Brazil) and IMPA (Instituto de Matematica Pura e Aplicada, Rio de Janeiro) through a post-doctoral fellowship.

References

1. Adams, S., Dirr, N., Peletier, M.A., Zimmer, J.: From a large-deviations principle to the Wasserstein gradient flow: a new micro-macro passage. *Commun. Math. Phys.* **307**, 791–815 (2011)
2. Adams, S., Dirr, N., Peletier, M. A., Zimmer, J.: Large deviations and gradient flows. *Phil. Trans. R. Soc. A* **371**(2005), 0341 (2013)
3. Ambrosio, L., Gigli, N., Savaré, G.: *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lectures in Mathematics ETH Zurich, 2nd edn. Birkhauser Verlag, Basel (2008)
4. Ambrosio, L., Savaré, G., Zambotti, L.: Existence and stability for Fokker-Planck equations with log-concave reference measure. *Probab. Theory Relat. Fields* **145**, 517–564 (2009)
5. De Giorgi, E., Marino, A., Tosques, M.: Problems of evolution in metric spaces and maximal decreasing curve. *Att. Acc. Naz. Linc. R. Cl. Sci. Fis. Mat. Nat.* **8(68)**(3), 180–187 (1980)
6. Erbar, M., Fathi, M., Laschos, V., Schlichting, A.: Gradient flow structure for McKean-Vlasov equations on discrete spaces. [arXiv:1601.08098](https://arxiv.org/abs/1601.08098)
7. Erbar, M., Maas, J.: Ricci curvature of finite Markov chains via convexity of the entropy. *Arch. Ration. Mech. Anal.* **206**(3), 997–1038 (2012)
8. Fathi, M.: A gradient flow approach to large deviations for diffusion processes. *Journal de Mathématiques Pures et Appliquées* (2014). [arXiv:1405.3910](https://arxiv.org/abs/1405.3910)
9. Fathi, M., Maas, J.: Entropic Ricci curvature bounds for discrete interacting systems. *Ann. Appl. Probab.* (2015)
10. Guo, M.Z., Papanicolaou, G.C., Varadhan, S.R.S.: Nonlinear diffusion limit for a system with nearest neighbor interactions. *Commun. Math. Phys.* **118**, 31–59 (1988)
11. Kipnis, C., Landim, C.: *Scaling Limits of Interacting Particle Systems*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 320. Springer, Berlin (1999)

12. Kosygina, E.: The behavior of the specific entropy in the hydrodynamic scaling limit. *Ann. Probab.* **29**(3), 1086–1110 (2001)
13. Maas, J.: Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.* **261**(8), 2250–2292 (2011)
14. Maas, J., Mielke, A.: Gradient structures for chemical reactions with detailed balance: I. Modeling and large-volume limit (2015)
15. Mariani, M.: A Gamma-convergence approach to large deviations. Preprint [arXiv:1204.0640v1](https://arxiv.org/abs/1204.0640v1) (2012)
16. Mielke, A.: Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Partial. Differ. Equ.* **48**(1–2), 1–31 (2013)
17. Otto, F.: The geometry of dissipative evolution equations: the porous medium equation. *Commun. Partial. Differ. Equ.* **26**(1–2), 101–174 (2001)
18. Sandier, E., Serfaty, S.: Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Commun. Pure Appl. Math.* **57**(12), 1627–1672 (2004)
19. Serfaty, S.: Gamma-convergence of gradient flows on Hilbert and metric spaces and applications. *Discret. Contin. Dyn. Syst.* **A 31**(4), 1427–1451 (2011)
20. Vazquez, J.L.: *The Porous Medium Equation: Mathematical Theory*. Oxford Mathematical Monographs. A Clarendon Press Publication, Oxford (2006)
21. Yau, H.T.: Relative entropy and hydrodynamics of Ginzburg-Landau models. *Lett. Math. Phys.* **22**, 63–80 (1991)