**University of Bath**

**Articles you may be interested in**

# Entropic and gradient flow formulations for nonlinear diffusion

Nicolas Dirr,[1,a)] Marios Stamatakis,[2,b)] and Johannes Zimmer[2,c)]

[1]*School of Mathematics, Cardiff University, Senghennydd Road, Cardiff CF24 4AG, United Kingdom*
[2]*Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, United Kingdom*

Nonlinear diffusion $\partial_t \rho = \Delta(\Phi(\rho))$ is considered for a class of nonlinearities $\Phi$. It is shown that for suitable choices of $\Phi$, an associated Lyapunov functional can be interpreted as thermodynamic entropy. This information is used to derive an associated metric, here called thermodynamic metric. The analysis is confined to nonlinear diffusion obtainable as hydrodynamic limit of a zero range process. The thermodynamic setting is linked to a large deviation principle for the underlying zero range process and the corresponding equation of fluctuating hydrodynamics. For the latter connections, the thermodynamic metric plays a central role. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4960748]

## I. INTRODUCTION

There is a wealth of results on nonlinear diffusion equations of the kind

$$\partial_t \rho = \Delta(\Phi(\rho)), \tag{1}$$

as the monumental monograph by Vázquez[24] for the porous medium equation, $\Phi(\rho) = \rho^m$, demonstrates. Recently, Bodineau, Lebowitz, Mouhot and Villani[4] studied drift-driven nonlinear diffusion equations with inhomogeneous Dirichlet data and established the existence of Lyapunov functionals for such equations.

Starting with the work of Jordan, Kinderlehrer, and Otto,[20] diffusion equations have been written as gradient flow of the entropy. For the example of linear diffusion, $\partial_t \rho = \Delta\rho$, with the Boltzmann entropy density $S_{\text{lin}} := \rho \log(\rho)$, it is immediate that $\partial_t \rho = \text{div}(\rho \nabla S'_{\text{lin}})$. The seminal insight of Ref. 20 is that the operator involved, $K(\rho)\xi := -\text{div}(\rho\nabla\xi)$, is associated to a metric and hence defines a geometry; the evolution is then determined by the steepest descent of the entropy $\mathcal{S}$ in this associated geometry. Here $\mathcal{S}$ is the classic entropy given by $\mathcal{S}(\rho) = \int S_{\text{lin}}(\rho(x))dx$ if $\rho$ is absolutely continuous with respect to the Lebesgue measure and $\mathcal{S}(\rho) = +\infty$ if $\rho$ is not absolutely continuous. More generally, gradient flows are formally equations of the form $\partial_t \rho = -K(\rho)d\mathcal{S}(\rho)$, where $\mathcal{S}$ is a functional defined on the state space, $d\mathcal{S}$ the Fréchet differential, and $K(\rho)$ an operator from the cotangent space to the tangent, the so-called Onsager operator. This formulation states that more than that $\mathcal{S}$ is a Lyapunov functional; it states that the evolution is given by the steepest descent evolution of $\mathcal{S}$ in the geometry given by $K$.

In this short note, we ask a simple question: For Equation (1), with a suitable choice of $\Phi$, can one show that an associate Lyapunov functional is in fact a thermodynamic entropy, that is, the entropy associated with an underlying particle system? This requires that (1) is the hydrodynamic limit of the underlying particle system. Furthermore, we will then write (1) as formal gradient flow (steepest descent) of the entropy, in a metric we derive as the metric associated with the thermodynamic entropy. We thus call this metric (or geometry) the thermodynamic metric (or geometry). In some cases, the formal setting can be shown to define rigorously a gradient flow.

---

a)DirrNP@cardiff.ac.uk
b)M.G.Stamatakis@bath.ac.uk
c)zimmer@maths.bath.ac.uk

**57**, 081505-1

We employ basic tools, techniques, and results from statistical mechanics and optimal transport. Yet, the result seems interesting in a twofold way. First, from the statistical mechanics perspective, the thermodynamic entropy is well known (see, e.g., Grosskinsky[15]), but has not been placed in a gradient flow context, i.e., linked to the associated thermodynamic metric. Second, the nonlinear diffusion $\partial_t \rho = \Delta \Phi(\rho)$ has been studied in a classic Wasserstein setting, with the 2-Wasserstein metric,[2] and has been shown to be the gradient flow, with respect to the 2-Wasserstein metric, of the free energy $\mathcal{F}$ given by

$$\mathcal{F}(\rho) = \int F(\rho)\,\mathrm{d}x + F'(\infty)\rho^{\perp}(\mathbb{T}^d).$$

Here $F$ is characterised by the ordinary differential equation $\Phi(\rho) = \rho F'(\rho) + F(\rho)$ and assumed to satisfy $F'(\infty) := \sup_{t>0} \frac{F(t)}{t} = +\infty$ (see Ref. 2, Theorem 11.2.5, for less restrictive assumptions). This gradient flow setting is different from the setting presented in this article.

It is well known that one equation can have different incarnations as gradient flow, that is, as steepest descent of a functional in a metric space; linear diffusion is a classic example with infinitely many such formulations. Then the functional is a Lyapunov functional. Here we show that in the setting of this article, the Lyapunov function has a thermodynamic interpretation, being the thermodynamic entropy of the so-called zero range process. We point out that the existence theory can be established without this thermodynamic interpretation. That is, the knowledge of Lyapunov functionals, which may or may not be entropies, can be key to a successful existence theory, as Ref. 13 testifies. At the same time, large deviation theory singles out one particular metric in the cases under consideration. We show in Section IV that this gives, at least formally, immediately rise to a variational scheme of Jordan-Kinderlehrer-Otto (JKO) type.[20]

We remark that different particle processes can lead to the same equation as hydrodynamic limit, but that equation will then typically have different representations as pairs of metric and entropy. For example, the symmetric exclusion process and a system of independent random walks both lead to the linear diffusion equation, but with two different entropies and correspondingly two different metrics, see Ref. 1.

It turns out that in the setting of this article, the same pair of a weighted Wasserstein metric and the thermodynamic entropy arises both from large deviation theory and a weighted version of the classical Benamou-Brenier formula. The weighted Wasserstein metric is (formally) conformally equivalent to the classical one, with the weight being related to the diffusion of a tagged particle. Namely, it is often convenient to write (1) as

$$\partial_t \rho = \Delta(\sigma(\rho)\rho) = \mathrm{div}(\sigma(\rho)\nabla\rho) + \mathrm{div}(\rho\nabla\sigma(\rho)), \tag{2}$$

since $\sigma(\rho) = \Phi(\rho)/\rho$ will be related to the diffusivity of a tagged particle in the underlying zero range process, as discussed in Subsection II B. Roughly speaking, the equation which is the hydrodynamic limit of the zero range process looks like the Kolmogorov equation of a single tagged particle. This diffusivity (sometimes called *self-diffusion coefficient*) will turn out to be closely related to the *thermodynamic metric* of the process discussed below.

*Plan of the paper*: We focus on the so-called zero-range process as underlying microscopic model; this process and its properties are described in Section II. In particular, the entropy (density) $S$ is identified there; in Section III the associated metric is determined. We show that this metric can also be obtained differently, by a large deviation argument; this is the content of Section IV, which also sketches connections to associated models of fluctuating hydrodynamics.

## II. PARTICLE MODELS

### A. The zero range process

In this subsection, we briefly summarise the zero range process on a flat torus; see Refs. 15 and 16 for more information on zero range process and the condensation phenomenon they exhibit for particular choices of jump rates.

A zero range process is an interacting particle system. Let $\Lambda$ be lattice, that is, an ordered collection of boxes indexed by an (at most countable) index set. Each box can be empty or contain a finite number of particles. The particles are indistinguishable. They evolve according to an irreducible Markov kernel on $\Lambda$ in the following way. A particle at a given site $x \in \Lambda$ interacts only with the particles located at the same site (hence zero range). Specifically, the rate at which one particle leaves a site $x$ depends only on the total number of particles at site $x$, described by the (local) *jump rate function* $g \colon \mathbb{N}_0 := \mathbb{N} \cup \{0\} \longrightarrow \mathbb{R}_0^+$, which satisfies $g(k) = 0$ if and only if $k = 0$. Here $g$ is assumed to be Lipschitz continuous, i.e., $\sup_{k \in \mathbb{N}_0} |g(k + 1) - g(k)| < +\infty$.

In particular, for the hydrodynamic limit in the flat torus $\mathbb{T}^d$ we consider the diffusively rescaled zero range processes with finite lattice, the discrete $d$-dimensional torus $\mathbb{T}_N^d \cong (\mathbb{Z}/N\mathbb{Z})^d \cong \{0, 1, \ldots, N - 1\}^d$ as $N \to +\infty$. The state space of the system is the set $\mathbb{M}_N := \mathbb{N}_0^{\mathbb{T}_N^d}$ of all configurations of particles $\eta \colon \mathbb{T}_N^d \longrightarrow \mathbb{N}_0$; the system is in state $\eta$ if $x$ contains $\eta(x)$ particles for all $x \in \mathbb{T}_N^d$. Let $p$ denote the symmetric nearest neighbour transition probability (renormalised so that it has total mass equal to $2d$)

$$p(x, y) := \sum_{j=1}^d \mathbb{1}_{\{-e_j, e_j\}}(y - x).$$

The symmetric (nearest neighbour) zero range process with jump rate $g$ is the Markov jump process on the state space $\mathbb{M}_N$ with formal generator,

$$L_N f(\eta) = N^2 \sum_{x, y \in \mathbb{T}_N^d} \{f(\eta^{x,y}) - f(\eta)\} g(\eta(x)) p(y - x),$$

where $f \colon \mathbb{M}_N \longrightarrow \mathbb{R}$ is a function on the space of configurations, $\eta^{x,y} := \eta - \mathbb{1}_{\{x\}} + \mathbb{1}_{\{y\}}$ is the configuration resulting from $\eta$ by moving a particle from $x$ to $y$, and the term $N^2$ comes from the diffusive rescaling of the process.

The microscopic stochastic dynamics defined by this generator can be described as follows. Starting from an initial configuration $\eta_0$, the system waits an exponential time $\tau_1$ with rate $N^2 \lambda(\eta_0) := N^2 \sum_{x \in \mathbb{T}_N^d} g(\eta_0(x))$, at which time one particle is moved from $x$ to $y$ with probability $\frac{g(\eta_0(x))}{\lambda(\eta_0)} p(y - x)$. Then the processes starts again with $\eta_{\tau_1} := \eta_0^{x,y}$ in place of $\eta_0$.

The equilibrium states of zero range processes are well known. They are characterised by the equation $\nu L_N = 0$ [Ref. 21, Prop. A.4.1]. If $\nu$ is a translation invariant measure on the configuration space $\mathbb{M}_N$, all its one-site marginals $\nu_x(\cdot) := \nu\{\eta(x) = \cdot\}$, $x \in \mathbb{T}_N^d$, must be equal. In particular, $\varphi := \mathbb{E}_\nu g(\eta(0)) = \mathbb{E}_\nu g(\eta(x))$. So if $\nu$ is a translation invariant and product equilibrium state with *fugacity* $\varphi := \mathbb{E}_\nu g(\eta(0))$, then $\nu L_N = 0$ can be solved to yield

$$\nu\{\eta \,|\, \eta(0) = k\} = \nu\{\eta \,|\, \eta(0) = 0\} \cdot \frac{\varphi^k}{g!(k)}, \quad (x, k) \in \mathbb{T}_N^d \times \mathbb{N}_0,$$

where $g!(k) := g(1) \cdots g(k)$ and $g!(0) = 1$. Since $\nu$ is a probability measure, the *partition function*

$$Z(\varphi) := \sum_{k=0}^{\infty} \frac{\varphi^k}{g!(k)}$$

converges at $\varphi$ and $\nu\{\eta \,|\, \eta(0) = 0\} = \frac{1}{Z(\varphi)}$. Conversely, for $\varphi$ in the proper domain $D_Z := \{\varphi \in \mathbb{R}_0^+ \,|\, Z(\varphi) < +\infty\}$, the translation invariant product measure $\bar{\nu}_\varphi^N$ in $\mathbb{M}_N$ with one-site marginal,

$$\bar{\nu}_\varphi^N\{\eta(0) = k\} := \frac{1}{Z(\varphi)} \frac{\varphi^k}{g!(k)},$$

is an equilibrium state of the zero range process. Consequently the family of the translation invariant and product equilibrium states of the zero range process is $\{\bar{\nu}_\varphi^N\}_{\varphi \in D_Z}$. From a statistical mechanics point of view, this family is the grand canonical ensemble. Of course, in order for non-trivial equilibrium states to exist, the critical fugacity

$$\varphi_c := \liminf_{k \to +\infty} \sqrt[k]{g!(k)} \tag{3}$$

must be strictly positive to ensure that the partition function $Z$ has non-trivial radius of convergence.

For the description of the evolution of the empirical density of the zero range process, it is useful to have a reparametrisation of the grand canonical ensemble by the density. This can be achieved by introducing the *density function* $R\colon D_Z \longrightarrow [0, +\infty]$, which is given by

$$R(\varphi) := \mathbb{E}_{\bar{\nu}_\varphi^N} \eta(0) = \frac{\varphi}{Z(\varphi)} \sum_{k=0}^\infty k \frac{\varphi^{k-1}}{g!(k)} = \frac{\varphi Z'(\varphi)}{Z(\varphi)}. \tag{4}$$

This function is analytic (on $[0, \varphi_c)$) and strictly increasing [Ref. 21, Section 2.3] with $D_R = D_Z \cap D_{Z'}$. The critical density is then $\rho_c := \sup R(D_R) \in (0, +\infty]$. In this article, we restrict our attention to zero range processes with superlinear jump rate $g$, i.e., for some constant $a_0 > 0$

$$g(k) \geq a_0 k \text{ for all } k \geq 1, \tag{5}$$

which enforces $\varphi_c = \rho_c = +\infty$ (we refer the reader to Ref. 16 for zero range processes that exhibit phase transitions). Then the inverse $\Phi := R^{-1}$ is well defined on all of $\mathbb{R}_0^+$; by setting $\nu_\rho^N := \bar{\nu}_{\Phi(\rho)}^N$, the grand canonical ensemble is reparametrised by the density. We note that $\Phi(\rho)$ is stochastically represented as the mean local jump rate with respect to the translation invariant and product equilibrium state $\nu_\rho^N$ of density $\rho$, since

$$\mathbb{E}_{\nu_\rho^N} g(\eta(0)) = \frac{1}{Z(\Phi(\rho))} \sum_{k=1}^\infty g(k) \frac{\Phi(\rho)^k}{g!(k)} = \Phi(\rho). \tag{6}$$

The *thermodynamic entropy density $S$* of the zero range process is defined as the Legendre transform of the pressure $P(\lambda) = \log(Z(\exp(\lambda)))$; here it is (see Ref. 16, Eq. (3.2); we use the opposite sign convention)

$$S(\rho) = \rho \log \Phi(\rho) - \log(Z(\Phi(\rho))) \tag{7}$$

and thus

$$S'(\rho) = \log(\Phi(\rho)) + \rho \frac{\Phi'(\rho)}{\Phi(\rho)} - \frac{Z'(\Phi(\rho))}{Z(\Phi(\rho))} \Phi'(\rho). \tag{8}$$

Since $\Phi = R^{-1}$, inserting $\varphi = \Phi(\rho)$ in (4) we get

$$\rho = R(\Phi(\rho)) = \frac{Z'(\Phi(\rho))}{Z(\Phi(\rho))} \Phi(\rho),$$

and so

$$\frac{Z'(\Phi(\rho))}{Z(\Phi(\rho))} \Phi'(\rho) = \rho \frac{\Phi'(\rho)}{\Phi(\rho)},$$

and hence (8) simplifies to

$$S'(\rho) = \log(\Phi(\rho)) = \log(\sigma(\rho)\rho). \tag{9}$$

The large deviations functional with speed $\frac{1}{N^d}$ (see also Subsection IV A) of the occupation variables $\{\eta(x)\}_{x \in \mathbb{Z}^d}$ with respect to the equilibrium state $\nu_{\rho_*}^\infty$ on the infinite lattice $\mathbb{Z}^d$ is given by [Ref. 21, Lemma 6.1.7]

$$I_{\rho_*}(\rho) = \rho \log \frac{\Phi(\rho)}{\Phi(\rho_*)} - \log \frac{Z(\Phi(\rho))}{Z(\Phi(\rho_*))}.$$

Thus it differs from the thermodynamic entropy by an affine function that depends only on the choice of the equilibrium state $\nu_{\rho_*}^N$ through its density $\rho_*$.

The integral functional $\mathcal{S}$ corresponding to the thermodynamic entropy $S$ is given by

$$\mathcal{S}(\rho) = \int_{\mathbb{T}^d} S(\rho^{ac}) \, dx + S'(\infty) \rho^\perp(\mathbb{T}^d), \tag{10}$$

where $S'(\infty) = \lim_{\rho \to +\infty} S(\rho)/\rho = \log \varphi_c$ is the rate of linear growth at infinity of the thermodynamic entropy, and $\rho = \rho^{ac} + \rho^\perp$ is the Radon-Nikodym decomposition of $\rho$, $\rho^{ac} \ll \mathcal{L}_{\mathbb{T}^d}$, $\rho^\perp \perp$

$\mathcal{L}_{\mathbb{T}^d}$ (with respect to Lebesgue measure $\mathcal{L}_{\mathbb{T}^d}$ on the torus). In particular, in the case that $\varphi_c = +\infty$, the thermodynamic entropy has superlinear growth $S'(\infty) = +\infty$, and thus $\mathcal{S}(\rho) = +\infty$ if $\rho \not\ll \mathcal{L}_{\mathbb{T}^d}$. The integral functional $\mathcal{I}_{\rho_*}$ associated to the rate functional $I_{\rho_*}$ as in (10),

$$\mathcal{I}_{\rho_*}(\rho) = \int_{\mathbb{T}^d} I_{\rho_*}(\rho^{ac}) \, dx + I'_{\rho_*}(\infty)\rho^{\perp}(\mathbb{T}^d),$$

is (see the proof of Lemma 5.1.6 in Ref. 21) the rate functional for the large deviations principle (with speed $\frac{1}{N^d}$) satisfied by the sequence

$$\Big(\frac{1}{N^d} \sum_{x \in \mathbb{T}_N^d} \eta(x)\delta_{x/N}\Big)_{\sharp} \nu_{\rho_*}^N \in \mathcal{P}(\mathcal{M}_+(\mathbb{T}^d)), \quad N \in \mathbb{N},$$

of the empirical embeddings of the zero range process with respect to an equilibrium state of density $\rho_*$. In the last display $\mathcal{M}_+(\mathbb{T}^d)$ denotes the set of all non-negative Borel measures on the torus, and $\mathcal{P}$ the set of probability measures; the push-forward $f_{\sharp}\mu$ of a measure $\mu$ by a (measurable) map $f$ is defined as the measure satisfying $f_{\sharp}\mu(A) = \mu(f^{-1}(A))$ for all (measurable) sets $A$. Since $S$ and $I_{\rho_*}$ differ by an affine function, the integral functionals $\mathcal{S}$ and $\mathcal{I}_{\rho_*}$ differ by a constant and consequently $\mathcal{S}$ and $\mathcal{I}_{\rho_*}, \rho_* \in \mathbb{R}_0^+$ define the same gradient flow (with respect to any metric).

It is well known (see Ref. 21, Theorem 5.1.1, for a precise formulation) that the hydrodynamic limit of the zero range process is given by (1), with $\Phi$ being the mean jump rate function. This means that starting the process from a sequence $\{\mu_0^N\}$ of initial distributions associated to a profile $\rho_0 \in L^1(\mathbb{T}^d)$, i.e., from a sequence of initial laws $\mu_0^N$, $N \geq 1$, such that for all $\delta > 0$ and $G \in C(\mathbb{T}^d)$,

$$\lim_{N \to +\infty} \mu_0^N \left\{ \left| \frac{1}{N^d} \sum_{x \in \mathbb{T}_N^d} G\left(\frac{x}{N}\right)\eta(x) - \int_{\mathbb{T}^d} G(x)\rho_0(x) \, dx \right| > \delta \right\} = 0,$$

then the law $\mu_t^N$ of the process at each time $t > 0$ is associated to the profile $\rho_t$, where $(\rho_t)_{t \in \mathbb{R}_0^+}$ is the unique weak solution of (1) starting from $\rho_0$. Furthermore, if the initial profile is of class $C^{2+\theta}$ for some $\theta > 0$, then the relative entropy method shows that asymptotically as $N \to +\infty$ the law $\mu_t^N$ looks like the product measure $\nu_{\rho_t(\cdot)}^N$ with slowly varying profile $\rho_t$, which is characterised by the marginals

$$\nu_{\rho_t(\cdot)}^N\{\eta(x) = k\} = \frac{1}{Z(\Phi(\rho_t(x/N)))} \frac{\Phi(\rho_t(x/N))^k}{g!(k)},$$

see Ref. 21, [Theorem 6.1.1] for a precise formulation.

## B. Asymptotic behaviour of tagged particles

Given a symmetric zero-range process, we can "tag" a particle that is originally at the origin, say and follow its movements, denoting its position at time $t$ by $X(t)$. It is natural to ask whether there is a central limit theorem for this tagged particle, that is, whether the law of $N^{-1}X(N^2 t)$ converges to a diffusion $dX = \sigma(X)dW$, where $W$ is a standard Brownian motion, and if so, what the diffusivity $\sigma$ is. This central limit theorem has been shown to hold in many situations, both in equilibrium (i.e., with the initial distribution of the particles given by an invariant measure for the process) and in non-equilibrium, see, e.g., Ref. 21 and the references therein. It turns out[19] that for $\rho > 0$,

$$\sigma(\rho) = \frac{\Phi(\rho)}{\rho},$$

where $\Phi$ and $\rho(t, x)$ as in (1). In Sec. III, this intrinsic diffusivity will be related to what we will call below the thermodynamic metric of the process.

## III. THERMODYNAMIC FORMULATION

We now phrase (2) with nonlinear $\sigma$ in a Wasserstein setting. We extend the classic Benamou-Brenier formulation[3] of optimal transport by a weighted version.

As a starting point, note that with the notation from Sec. II (see, e.g., (9)),

$$\partial_t \rho = \Delta(\Phi(\rho)) = \mathrm{div}(\Phi(\rho)\nabla S'(\rho)) = \mathrm{div}(\rho\sigma\nabla S'(\rho)), \tag{11}$$

with $\sigma(\rho) = \Phi(\rho)/\rho$. In order to simplify notation, we will often omit the dependence of $\sigma$ on $\rho$.

In this section we will show that this structure allows us to interpret (2) as a gradient flow of the entropy (10) with respect to a metric which is associated to the operator $\mathrm{div}(\Phi(\rho)\nabla\cdot)$ appearing in the equation above. This metric will be shown to be a weighted Wasserstein metric with weight $\sigma^{-1}$.

Here we will assume that our spatial domain is the $d$-dimensional flat torus $\mathbb{T}^d$. For general bounded domains and homogeneous Dirichlet boundary data the reasoning is analogous. The case of inhomogeneous boundary data is covered in Ref. 4.

We introduce a metric tensor, following the work of Otto.[23] See also the presentation in Ref. 25, Section 8. Here we work with the set $\mathcal{P}_{ar}\mathbb{T}^d$ of probability measures on the torus with continuous Lebesgue densities. What follows can be seen as the Benamou-Brenier formulation of optimal transport with a nonconstant metric. In this context, a point $\rho \in \mathcal{P}_{ar}\mathbb{T}^d$ in this set has formally as tangent space the space of all infinitesimal variations at $t = 0$ of smooth curves $(-\epsilon, \epsilon) \ni t \mapsto \rho_t$ with $\rho_0 = \rho$. By the conservation of mass, these variations satisfy $\int \partial_t|_{t=0}\rho_t \, dx = 0$ and thus the tangent space as a set can be identified with the set of all continuous functions $\zeta \colon \mathbb{T}^d \longrightarrow \mathbb{R}$ such that $\int \zeta \, dx = 0$. The tangent spaces are thus all equal as sets, while the metric we are going to define depends on the base point $\rho$. As in the standard case (e.g., Ref. 23), a metric is defined on each tangent space by associating to each tangent vector $\zeta = \partial_t|_{t=0}\rho_t$ (i.e., a derivative of a path in the space of densities/measures) a vector field representing this variation, via the equation

$$\zeta + \mathrm{div}(\rho v) = 0, \tag{12}$$

with unknown the vector field $v \colon \mathbb{T}^d \to \mathbb{R}^d$. To ensure uniqueness of the vector field $v$ representing the tangent vector $\zeta$, we minimise the *weighted* kinetic energy under constraint (12),

$$\inf_v \left\{ \int_{\mathbb{T}^d} \rho\sigma^{-1}|v|^2 \, dx \,\middle|\, \zeta + \mathrm{div}(\rho v) = 0 \right\}. \tag{13}$$

In the classic case (no division by $\sigma$), this term has an interpretation as twice the kinetic energy, and we thus call (13) (twice) the *weighted kinetic energy*. We take outer variations of the form $v + \epsilon\rho^{-1}w$ with divergence-free $w$; note that the latter property ensures that (12) is preserved. Then the extremality condition for (13) reads

$$\int_{\mathbb{T}^d} \sigma^{-1}\langle v, w\rangle \, dx = 0$$

for all divergence-free $w$, where $\langle\cdot,\cdot\rangle$ is the usual Euclidean scalar product on $\mathbb{R}^d$. By the Helmholtz-Hodge decomposition of vector fields into divergence-free and gradient fields, this implies

$$\sigma^{-1}v = \nabla\xi \tag{14}$$

for some real-valued function $\xi$, or equivalently $v = \sigma\nabla\xi$. Note that thus $\zeta = -\mathrm{div}(\rho v) = -\mathrm{div}(\sigma\rho\nabla\xi)$. As (for given suitable $\zeta$) the equation $-\mathrm{div}(\rho\sigma\nabla\xi) = \zeta$ has a unique solution, there is (up to additive constants) only one such $\xi = \xi_\zeta$, where the latter notation indicates the dependence on $\zeta$. In this way, we can associate to each tangent vector $\zeta$ the vector field $v_\zeta := \sigma(\rho)\nabla\xi_\zeta$. Then on the tangent space at $\rho$, we define the metric

$$g_\rho(\zeta_1, \zeta_2) := \int_{\mathbb{T}^d} \frac{\rho}{\sigma(\rho)}\langle v_{\zeta_1}, v_{\zeta_2}\rangle \, dx = \int_{\mathbb{T}^d} \rho\sigma(\rho)\langle\nabla\xi_{\zeta_1}, \nabla\xi_{\zeta_2}\rangle \, dx. \tag{15}$$

For each $\rho$, one can think of $\sigma^{-1}(\rho)\langle v, v\rangle$ as a Riemannian metric on a manifold which is topologically $\mathbb{R}^d$. This penalises transport in regions with slow diffusion, where $\sigma(\rho)$ is small. In the classic setting,[23] the division by $\sigma$ is absent. As in classic formal Otto calculus, the cotangent space is a

space of functions modulo constants and thus the function $\xi_\zeta$ corresponding to a tangent vector $\zeta$ via (14) is the cotangent vector corresponding to $\zeta$ via the metric $g$.

Let now $\mathcal{S}$ be any functional of the form (10) for some integrand $S \colon \mathbb{R}_0^+ \longrightarrow \mathbb{R}$, which, by the regularity assumptions made earlier on the probability measures under consideration, is in fact given by

$$\mathcal{S}(\rho) := \int_\Omega S(\rho)\,\mathrm{d}x. \tag{16}$$

We recall that in the usual $L^2$ pairing, the differential $d\mathcal{S}$ is represented by the variational derivative $\frac{\delta\mathcal{S}(\rho)}{\delta\rho} = S'(\rho)$. Analogously, if we denote by $\langle\cdot|\cdot\rangle_\rho$ the dual pairing between tangent and cotangent space at $\rho$, and if $\zeta$ is represented by $v_\zeta = \sigma\nabla\xi_\zeta$, an integration by parts yields

$$\begin{aligned}
\langle d\mathcal{S}(\rho)|\zeta\rangle_\rho &= \int \zeta\frac{\delta\mathcal{S}(\rho)}{\delta\rho}\,\mathrm{d}x = -\int \mathrm{div}(\rho v_\zeta)S'(\rho)\,\mathrm{d}x \\
&= \int \rho\langle v_\zeta, \nabla S'(\rho)\rangle\,\mathrm{d}x = \int \frac{\rho}{\sigma(\rho)}\langle v_\zeta, \sigma(\rho)\nabla S'(\rho)\rangle\,\mathrm{d}x \\
&= g_\rho(\zeta, -\mathrm{div}(\rho\sigma\nabla S'(\rho))).
\end{aligned} \tag{17}$$

The last equality follows by definition of the metric $g$, since if $\bar\zeta := -\mathrm{div}(\rho\sigma\nabla S'(\rho))$ then $v_{\bar\zeta} = \sigma(\rho)\nabla S'(\rho)$. In words, formula (17) says that the cotangent vector $d\mathcal{S}(\rho)$ is represented via the metric tensor $g$ by the tangent vector $-\mathrm{div}(\sigma\rho\nabla S'(\rho))$. Thus in our case the Onsager operator $K = K_\sigma$ corresponding to the weighted metric $g$ maps $d\mathcal{S}$ to the tangent vector $\zeta = -\mathrm{div}(\sigma\rho\nabla[S'(\rho)])$, that is,

$$K(\rho)d\mathcal{S}(\rho) = -\mathrm{div}(\sigma\rho\nabla S'(\rho)).$$

Note that the identification $\zeta \mapsto \xi_\zeta$ given via the minimisation of (13) and (14) is in fact the inverse of the Onsager operator $K$.

A gradient flow with respect to a metric is distinguished by the fact that the velocity $\partial_t\rho_t$ of a path $t \mapsto \rho_t$ is related via the metric to the rate of change of a functional, i.e., here $\zeta = -\mathrm{div}(\sigma\rho\nabla S'(\rho))$, and so the gradient flow equation for a functional of the form (10) becomes

$$\partial_t\rho_t = -K(\rho)d\mathcal{S}(\rho) = \mathrm{div}(\sigma\rho\nabla S'(\rho)).$$

So now if we choose the thermodynamic entropy (7) as the integrand in the functional $\mathcal{S}$, with its derivative given by (8), then we recover (2). In summary if we choose the thermodynamic entropy, then precisely the choice (15) for the metric recovers the nonlinear diffusion equation (1); we thus call (15) the *thermodynamic metric*.

Weighted kinetic energies as (13) have been introduced and studied before, in particular by Dolbeault, Nazaret and Savaré,[10] and Carrillo, Lisini and Slepčev.[5] We sketch how their results can be used to obtain existence and uniqueness for solutions to the gradient flow corresponding to (1) in the thermodynamic setting. The global metric corresponding to the kinetic energy (13) is given by

$$W_{\sigma;2}^2(\mu, \nu) := \inf \int_0^1 \int_{\mathbb{T}^d} \rho\sigma^{-1}|v|^2\,\mathrm{d}x\,\mathrm{d}t,$$

where the infimum is taken over all pairs $(\rho_t, v_t)_{0 \le t \le 1}$ with $\rho_0 = \mu$, $\rho_1 = \nu$ that satisfy the usual continuity equation $\partial_t\rho + \mathrm{div}(\rho v) = 0$. Equivalently, one can consider

$$W_{\sigma;2}^2(\mu, \nu) := \inf \int_0^1 \int_{\mathbb{T}^d} \Phi(\rho)|v|^2\,\mathrm{d}x\,\mathrm{d}t$$

with the constraint $\partial_t\rho + \mathrm{div}(\Phi(\rho)v) = 0$; the latter variational formulation is Ref. 10, Eq. (1.6) in our notation. Then, as proved in Refs. 10 and 5, concavity of $\Phi$ and a type of generalised McCann conditions are sufficient conditions to establish existence and uniqueness of the gradient flow of the thermodynamic entropy with respect to the weighted Wasserstein metric $W_\sigma$. These generalised $d$-dimensional McCann conditions [Ref. 5, Definition 4.5] take in our context the form

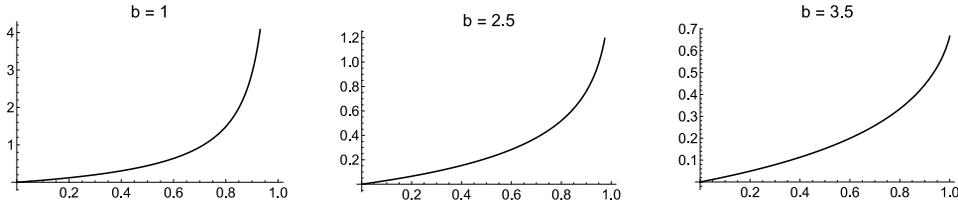$$\Phi(\rho)\Phi'(\rho) \ge (1 - 1/d)\int_0^\rho \Phi'(r)^2\,\mathrm{d}r.$$

FIG. 1.  Graph of the density function $R_b = \Phi_b^{-1}$ in the Evans model for $b = 1$, $b = 2.5$ and $b = 3.5$.

We note that in dimension $d = 1$ the McCann condition is equivalent to $\Phi'(\rho) \geq 0$ for all $\rho \geq 0$, which is true for all mean jump rate functions $\Phi$ of zero range processes.

Let us also note that not all mean jump rate functions of zero range processes are concave. Two simple examples where it is concave are $\Phi(\rho) = \rho$ and $\Phi(\rho) = \rho/(\rho + 1)$, which correspond to $g(k) = k$ and $g(k) \equiv 1$, respectively. Although the latter jump rate does not satisfy the superlinearity assumption (5), one can consider the jump rate $g_\epsilon(k) = g(k) + \epsilon k$ for some $\epsilon > 0$, which according to (6) gives the mean jump rate $\Phi_\epsilon(\rho) = \Phi(\rho) + \epsilon \rho$. More generally, as one can see by plots the mean jump rate function corresponding to the *Evans model*,

$$g_b(k) = \left(1 + \frac{b}{k}\right) \mathbb{1}_{\{k \geq 1\}}$$

is concave for all parameters $b \geq 0$. The density function $R_b = \Phi_b^{-1}$ in the Evans model is given by Ref. 16 [Eq. (4.5)]

$$R_b(\varphi) = \frac{\varphi_2 F_1(2, 2, 2 + b, \varphi)}{(1 + b)_2 F_1(1, 1, 1 + b, \varphi)}, \quad 0 \leq \varphi,$$

where $_2F_1(a, b, c; z) := \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{z^k}{k!}$ is the hypergeometric function and

$$(a)_k := \begin{cases} a(a + 1) \cdot \cdots \cdot (a + k - 1) & \text{if } k \geq 1 \\ 1 & \text{if } k = 0 \end{cases}$$

is the rising Pochhammer symbol. Fig. 1 gives plots for $R_b$ for some values of $b$.

An example of a zero-range process with non-convex density function is given by the *Landim jump rate* function given by

$$g_b(k) = \begin{cases} k & \text{if } k = 0, 1 \\ \left(\dfrac{k}{k - 1}\right)^b & \text{if } k \geq 2 \end{cases}.$$

In this case, $g_b!(k) = k^b$ and the partition function is $Z_b(\varphi) = 1 + \sum_{k=1}^{\infty} \frac{\varphi^k}{k^b} =: 1 + \mathrm{Li}_b(\varphi)$, where $\mathrm{Li}_b$ is the polylogarithmic function. This results in the density function

$$R_b(\varphi) = \frac{\mathrm{Li}_{b-1}(\varphi)}{1 + \mathrm{Li}_b(\varphi)}.$$

Fig. 2 indicates that $R_b$ is convex for small $b > 0$, concave for large $b > 0$, and neither of the two for intermediate values. We close this section with the rather obvious remark that equations of porous medium type, $\Phi(\rho) = \rho^m$, are not covered by the setting described here, as for the zero-range process $\Phi'(0)$ is strictly positive.

## IV. CONNECTION TO LARGE DEVIATION THEORY AND FLUCTUATIONS

### A. Large deviation theory

In Section III, we used the thermodynamic entropy $\mathcal{S}$ from Section II and used it to identify the associated (weighted Wasserstein) metric (15); indeed, the calculation in Section III shows that
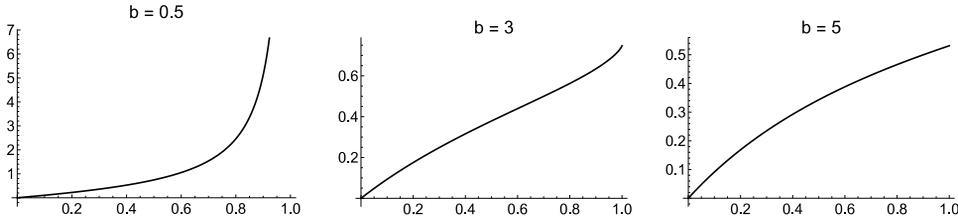
FIG. 2.  Graph of the density function $R_b = \Phi_b^{-1}$ in the Landim model for $b = 0.5$, $b = 3$, and $b = 5$.

this metric is associated to the operator $\mathrm{div}(\Phi(\rho)\nabla\cdot)$ in (11) which gives the gradient flow formulation of the thermodynamic entropy. We now describe an alternative link between the zero-range process and this metric. The tool is Large Deviation Theory, which we sketch first for a toy model. This theory is a step beyond a central limit hydrodynamic limit theorem; on the most superficial level, large deviation theory describes the most likely event of improbable events in terms of an exponential (un-)likelihood. As explained with the toy model, the probability of stochastic events of $x^\epsilon$ is determined to be of the form $\mathbb{P}(\{x^\epsilon(t) \sim x(t)\}) \sim \exp\left(-\frac{1}{\epsilon}I(x(t))\right)$; here $\sim$ means asymptotic equivalence (precise formulations are given in any textbook on Large deviation Theory, for example Ref. 9). The functional $I$ is called the *rate functional* and the central finding of this section is that the rate functional for the zero-range process has as minimiser the nonlinear diffusion equation (1), which appears in the norm defined by the weighted Wasserstein metric (15).

We first explain this link using a toy model of a stochastically perturbed one-dimensional gradient flow,

$$dX^\epsilon = -V'(X^\epsilon)dt + \sqrt{\epsilon}\,dW,$$

where $dW$ denotes a Wiener process. Heuristically, one expects that for small noise, $\epsilon$ small, trajectories will tend to be close to the deterministic trajectory solving $\dot{x} = -V'(x)$. Large deviation theory describes the probability of observing a realisation, here a trajectory, by means of a *rate functional* $I$; the probability of finding a given trajectory is essentially $\exp(-1/\epsilon\, I)$ as sketched now. For the toy model, the large deviations rate functional depends on the final time $T$ and is denoted $I_T$; it is

$$I_T(\dot{x}, x) = \int_0^T \frac{1}{2}(\dot{x} + V'(x))^2 \, ds \tag{18}$$

$$= \int_0^t \frac{1}{2}|\dot{x}|^2 \, ds + \frac{1}{2}\int_0^T (V'(x))^2 \, ds + V(x(T)) - V(x(0)) \tag{19}$$

with $\dot{x}$ denoting the time derivative of $x$. In essence, this means that

$$\mathbb{P}(\{x^\epsilon(0) \sim x, \, x^\epsilon(T) \sim y\}) \sim \exp\left(-\frac{1}{\epsilon}\inf I_T(\dot{x}, x)\right),$$

where the infimum is taken over paths $x = x(s)$ with $x(0) = x$ and $x(T) = y$. In order to find the most likely position at $t = T$, we have to minimise first over all paths such that $x(0) = x$ and $x(T) = y$ and then over all terminal values $x(T)$.

Note that the first integral in (19) is bounded from below by the distance $d(x(0), x(T))$ between the initial and the final points. Choosing $T = h \ll 1$ and re-scaling, we see that we have to minimise

$$\frac{1}{2h}d^2(x(0), x(h)) + V(x(h)) + \frac{h}{2}\int_0^1 (V'(x(sh)))^2 \, ds.$$

If $d(x(0), x(h)) \gg h$, then the last term is clearly of lower order, while if $d(x(0), x(h)) = O(h)$ and $V$ is smooth enough, then the last integral equals

$$\frac{h}{2}\int_0^1 (V'(x(sh)))^2 \, ds = \frac{h}{2}V'(x(0)) + O(h^2),$$

so again the term relevant for the minimisation is of lower order. Formally this means that we can find the most likely position $x(h)$ by minimising the functional[20]

$$\frac{1}{2h}d^2(x(0), x(h)) + V(x(h)). \tag{20}$$

The analogous computation starting with the large deviation rate functional for the empirical measure of $n$ Brownian particles (with $n$ playing the role of the large parameter $1/\epsilon$ in the toy model; here the rate functional is defined on function spaces),[7] then one obtains an equation with the same structure as (20), with $d$ being the Wasserstein norm, and the entropy $S$ in place of $V$ (see Ref. 11 for the proof). Thus, one obtains the celebrated Jordan-Kinderlehrer-Otto functional[20] as expansion of the associated large deviation rate functional.

We now apply this reasoning in the context of nonlinear diffusion and more specifically the zero range process. The large deviation rate function for the zero range process is of the form

$$\int_0^T \|\partial_t \rho(s) - \Delta\Phi(\rho(s))\|^2_{H^{-1}_{\Phi(\rho(s))}} \, \mathrm{d}s \tag{21}$$

(see Ref. 22 and references therein); here the weighted negative Sobolev norm is defined as follows. Let for some non-negative bounded function $w$

$$\langle u, v \rangle_{H^1_w} := \int_{\mathbb{T}^d} w \nabla u \cdot \nabla v \, \mathrm{d}x,$$

and denote by $\Delta^{-1}_w u$ the solution of

$$\mathrm{div}(w\nabla(\Delta^{-1}_w u)) = u. \tag{22}$$

Then

$$\langle u, v \rangle_{H^{-1}_w} = \langle \Delta^{-1}_w u, \Delta^{-1}_w v \rangle_{H^1_w} = -\int_{\mathbb{T}^d} u \Delta^{-1}_w v \, \mathrm{d}x,$$

where the last identity used integration by parts and (22).

As seen from the reasoning above, by expanding the square in (21), we obtain a metric term

$$\int_{\mathbb{T}^d} \Phi \left| \nabla(\Delta^{-1}_{\Phi(\rho_t)}\partial_t\rho) \right|^2 \mathrm{d}x = g_\sigma(\partial_t\rho_t, \partial_t\rho_t),$$

where $g$ is as in (15). This follows since if $v_\zeta = \sigma\nabla\xi_\zeta$ is the vector field representing a tangent vector $\zeta$ at $\rho$ via (14), then $\zeta = -\mathrm{div}(\rho v_\zeta) = -\mathrm{div}(\rho\sigma\nabla\xi_\zeta) = -\mathrm{div}(\Phi(\rho)\nabla\xi_\zeta)$ and thus $\Delta^{-1}_{\Phi(\rho)}(\zeta) = \xi_\zeta$.

The null Lagrangian (the mixed term) equals

$$\int_0^T \int_{\mathbb{T}^d} \partial_t\rho \Delta^{-1}_{\Phi(\rho)}(\Delta\Phi(\rho)) \, \mathrm{d}x \, \mathrm{d}t,$$

so it remains to compute $\Delta^{-1}_{\Phi(\rho)}(\Delta\Phi(\rho))$. As $\log\Phi(\rho)$ solves

$$\mathrm{div}(\Phi(\rho)\nabla\log\Phi(\rho)) = \Delta\Phi(\rho),$$

we have $\Delta^{-1}_{\Phi(\rho)}(\Delta\Phi(\rho)) = \log(\Phi(\rho))$, so if $S'(\rho) = \log(\Phi(\rho))$ as in (8), we obtain

$$\int_0^T \int_{\mathbb{T}^d} \partial_t\rho \Delta^{-1}_\Phi(\Delta\Phi(\rho)) \, \mathrm{d}x \, \mathrm{d}t = \int_{\mathbb{T}^d} S(\rho(T)) \, \mathrm{d}x - \int_{\mathbb{T}^d} S(\rho(0)) \, \mathrm{d}x \tag{23}$$

and recover the thermodynamic entropy $\mathcal{S}$ of (16).

The remaining quadratic term is for small time intervals formally of lower order. Thus in summary, an expansion of the large deviation rate functional (21) for nonlinear diffusion associated with the zero range process yields a variational formulation of Jordan-Otto-Kinderlehrer type, with the metric being the weighted Wasserstein metric considered in this article, and the functional being the thermodynamic entropy.

## B. Fluctuating hydrodynamics

Subsection IV A showed that the thermodynamic metric is related not only to the limit equation, but also to the probability of deviations from it. At least at a formal level, this connection is even stronger, as we now explain. Formally, the metric is given by a symmetric linear operator, which has an inverse that can be interpreted as covariance matrix. If we add a suitable space-time noise to the limit partial differential equation, then it formally has not only the same large deviations rate functionals, but finite moments of the fluctuations (that is, the difference between stochastic process and deterministic limit) converge to the same process as for the original particle model. This is now explained in more detail. Consider the equations

$$d\rho^\epsilon = \Delta(\Phi(\rho^\epsilon))dt + \mathrm{div}(\sqrt{\epsilon\sigma(\rho^\epsilon)\rho^\epsilon}\dot{W}), \tag{24}$$

$$\partial_t\rho = \Delta(\Phi(\rho)), \tag{25}$$

where $\dot{W}$ is space-time white noise. We point out that the stochastic PDE (SPDE) (24) is not expected to have strong (i.e., pathwise) solutions. To our knowledge, the suitable notion of generalised solutions is an open problem. Consider, however, the fluctuations, i.e., the differences between solutions of (24) and the deterministic PDE (25), re-scaled with the scaling of the central limit theorem. As

$$Z^\epsilon := \rho - \rho^\epsilon$$

is small, we can linearise around $\rho$ and obtain formally that a limit (in law) $Z^0$ of $\epsilon^{-1/2}Z^\epsilon$ should solve the *linear* SPDE

$$dZ^0 = \Delta(\Phi'(\rho)Z^0)dt + \mathrm{div}(\sqrt{\sigma(\rho^\epsilon)\rho^\epsilon}\dot{W}), \tag{26}$$

i.e., it is an Ornstein-Uhlenbeck process. Here $\Phi'$ is the *self-diffusion* coefficient of the zero range process, while $\sigma$ is the *bulk-diffusion* coefficient.

This convergence statement for the fluctuations can be stated in the following weaker form. Consider for smooth $\gamma : \mathbb{R}^d \to \mathbb{R}$ the process

$$Y^\epsilon := \epsilon^{-1/2}\langle\gamma, \rho - \rho^\epsilon\rangle,$$

then by (26), an application of Ito's formula (see, e.g., Ref. 6 [Chap. 4.5]), and the (formal) limit $\epsilon \to 0$, we get for any smooth $F : \mathbb{R} \to \mathbb{R}$ that

$$F(Y^0(t)) - F(Y^0(0)) - \int_0^t F'(Y^0(s))\langle Y^0(s), \Phi'(\rho(s))\Delta\gamma\rangle ds$$

$$- \frac{1}{2}\int_0^t F''(Y^0(s))\langle\sigma(\rho(s))\rho(s)\nabla\gamma, \nabla\gamma\rangle ds \quad (27)$$

is a *martingale*.

Note that in the last term, the expression $\langle\sigma(\rho(s))\rho(s)\nabla\gamma, \nabla\gamma\rangle$ can be written as $\|\gamma\|^2_{H^1_\Phi(\rho)}$, which implies that the large deviations should be related to the metric with the inverse operator, i.e., $\|\gamma\|^2_{H^{-1}_\Phi(\rho)}$, as found in Subsection IV A.

We now relate the martingale (27) to the zero range process. We consider the empirical measure of the zero range process, i.e., for smooth $\gamma$,

$$\langle\rho^N, \gamma\rangle := \frac{1}{N^d}\sum_{x\in\mathbb{T}^d_N}\gamma\left(\frac{x}{N}\right)\eta(x),$$

and the fluctuations,

$$Y^N := N^{-d/2}\langle\gamma, \rho - \rho^N\rangle.$$

It can be shown[12,14,18] for a large class of zero range processes that for any smooth test function, $F(Y^N)$ converge to a solution of the martingale problem (27). In other words, the limit of the fluctuations of the zero range process is in law equal to those of the SPDE (24). This can be seen as justification of the complicated noise term in (24); precisely this noise term is related to

fluctuations of the zero range process via the martingale expression (27). We remark that (24) is of fluctuating hydrodynamics type. In the linear case $\Phi(\rho) = \rho$, the multiplicative conservative noise in (24) has been derived by Dean,[8] and using quite different lines of arguments by Sturm and von Renesse;[26] the link between the noise and the geometry has been explained by a formal large deviation argument.[17]

While the discussion in this paper is limited to a sub-class of zero range processes, it therefore does not cover all nonlinearities $\Phi(\rho)$ but only those that occur as hydrodynamic limit of such a process. Yet, we remark that a solution of (24) for a certain range of $\epsilon > 0$ would formally yield a stochastic process whose hydrodynamic limit is a gradient flow of the entropy defined by (7) with respect to the weighted Wasserstein metric, and whose large deviation rate functional is as in Subsection IV A. The existence of solutions to (24), however, is to our knowledge in general an open problem.

## ACKNOWLEDGMENTS

[1] Adams, S., Dirr, N., Peletier, M., and Zimmer, J., "Large deviations and gradient flows," Philos. Trans. R. Soc., A **371**(2005), 20120341 (2013).
[2] Ambrosio, L., Gigli, N., and Savaré, G., *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Lectures in Mathematics ETH Zürich (Birkhäuser Verlag, Basel, 2005).
[3] Benamou, J.-D. and Brenier, Y., "A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem," Numer. Math. **84**(3), 375–393 (2000).
[4] Bodineau, T., Lebowitz, J., Mouhot, C., and Villani, C., "Lyapunov functionals for boundary-driven nonlinear drift-diffusion equations," Nonlinearity **27**(9), 2111–2132 (2014).
[5] Carrillo, J. A., Lisini, S., Savaré, G., and Slepčev, D., "Nonlinear mobility continuity equations and generalized displacement convexity," J. Funct. Anal. **258**(4), 1273–1309 (2010).
[6] Da Prato, G. and Zabczyk, J., *Stochastic Equations in Infinite Dimensions*, 2nd ed. Encyclopedia of Mathematics and its Applications Vol. 152 (Cambridge University Press, Cambridge, 2014).
[7] Dawson, D. A. and Gärtner, J., "Large deviations from the McKean-Vlasov limit for weakly interacting diffusions," Stochastics **20**(4), 247–308 (1987).
[8] Dean, D. S., "Langevin equation for the density of a system of interacting Langevin processes," J. Phys. A: Math. Gen. **29**(24), L613 (1996).
[9] den Hollander, F., *Large Deviations*, Fields Institute Monographs (American Mathematical Society, Providence, RI, 2000), Vol. 14.
[10] Dolbeault, J., Nazaret, B., and Savaré, G., "A new class of transport distances between measures," Calculus Var. Partial Differ. Equations **34**(2), 193–231 (2009).
[11] Duong, M. H., Laschos, V., and Renger, M., "Wasserstein gradient flows from large deviations of many-particle limits," ESAIM: Control Optim. Calculus Var. **19**(4), 1166–1188 (2013).
[12] Ferrari, P. A., Presutti, E., and Vares, M. E., "Nonequilibrium fluctuations for a zero range process," Ann. Inst. H. Poincaré Probab. Statist. **24**(2), 237–268 (1988).
[13] Fornaro, S., Lisini, S., Savaré, G., and Toscani, G., "Measure valued solutions of sub-linear diffusion equations with a drift term," Discrete Contin. Dyn. Syst. **32**(5), 1675–1707 (2012).
[14] Gielis, G., Koukkous, A., and Landim, C., "Equilibrium fluctuations for zero range processes in random environment," Stochastic Processes Appl. **77**(2), 187–205 (1998).
[15] Grosskinsky, S., "Phase transitions in nonequilibrium stochastic particle systems with local conservation laws," Ph.D. thesis, Technische Universität München, 2004.
[16] Grosskinsky, S., Schütz, G. M., and Spohn, H., "Condensation in the zero range process: Stationary and dynamical properties," J. Statist. Phys. **113**(3-4), 389–410 (2003).
[17] Jack, R. and Zimmer, J., "Geometrical interpretation of fluctuating hydrodynamics in diffusive systems," J. Phys. A: Math. Theor. **47**, 485001 (2014).
[18] Jankowski, H., "Density fluctuations for the zero range process with colour," e-print arXiv:math/0607505v1 (2006).
[19] Jara, M., Landim, C., and Sethuraman, S., "Nonequilibrium fluctuations for a tagged particle in one-dimensional sublinear zero-range processes," Ann. Inst. Henri Poincaré Probab. Stat. **49**(3), 611–637 (2013).
[20] Jordan, R., Kinderlehrer, D., and Otto, F., "The variational formulation of the Fokker-Planck equation," SIAM J. Math. Anal. **29**(1), 1–17 (1998).
[21] Kipnis, C. and Landim, C., *Scaling Limits of Interacting Particle Systems*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] Vol. 320 (Springer-Verlag, Berlin, 1999).

[22] Koukkous, A. and Guiol, H., "Large deviations for a zero mean asymmetric zero range process in random media," Technical report, Relatorio de Pesquisa RP 16/00 IMECC-UNICAMP, 2000; preprint arXiv:math/0009110v1.

[23] Otto, F., "The geometry of dissipative evolution equations: The porous medium equation," Commun. Partial Differ. Equations **26**(1-2), 101–174 (2001).

[24] Vázquez, J. L., *The Porous Medium Equation: Mathematical Theory*, Oxford Mathematical Monographs (The Clarendon Press; Oxford University Press, Oxford, 2007).

[25] Villani, C., *Topics in Optimal Transportation*, Graduate Studies in Mathematics Vol. 58 (American Mathematical Society, Providence, RI, 2003).

[26] von Renesse, M.-K. and Sturm, K.-T., "Entropic measure and Wasserstein diffusion," Ann. Probab. **37**(3), 1114–1191 (2009).