Communications in Contemporary Mathematics
Vol. 15, No. 5 (2013) 1350017 (43 pages)

DOI: 10.1142/S021919971350017X

VARIATIONAL FORMULATION OF THE FOKKER–PLANCK EQUATION WITH DECAY: A PARTICLE APPROACH

MARK A. PELETIER* and D. R. MICHIEL RENGER†

Department of Mathematics and Computer Sciences
Institute for Complex Molecular Systems
Technische Universiteit Eindhoven
5600 MB Eindhoven, The Netherlands
**m.a.peletier@tue.nl*
†d.r.m.renger@tue.nl

MARCO VENERONI

Department of Mathematics F. Casorati
University of Pavia, Pavia, Italy
marco.veneroni@unipv.it

Received 31 October 2012
Accepted 4 March 2013
Published 30 May 2013

We introduce a stochastic particle system that corresponds to the Fokker–Planck equation with decay in the many-particle limit, and study its large deviations. We show that the large-deviation rate functional corresponds to an energy-dissipation functional in a Mosco-convergence sense. Moreover, we prove that the resulting functional, which involves entropic terms and the Wasserstein metric, is again a variational formulation for the Fokker–Planck equation with decay.

Keywords: Large deviations; Fokker–Planck equation with decay; Wasserstein metric; gradient flow; Mosco convergence.

Mathematics Subject Classification 2010: 35Q84, 60F10, 49S05

1. Introduction

1.1. *On the origin of Wasserstein gradient flows*

Since the introduction of the Wasserstein gradient flows in 1997–98 [27, 28, 41, 43] it has become clear that a very large number of well-known parabolic partial differential equations and other evolutionary systems can be written as gradient flows. Examples of these are nonlinear drift-diffusion equations [2], diffusion-drift equations with non-local interactions [9], higher-order parabolic equations [42, 23, 26, 33, 24], moving-boundary problems [42, 45], and chemical reactions [36]. The parallel development of rate-independent systems introduced similar variational structures for friction [19], delamination [30], plasticity [34], phase transformations [39], hysteresis [38], and various other phenomena. Further generalizations are suggested by taking limits of gradient flows, as in the case of Kramers' equation for chemical reactions [5].

This multitude of gradient-flow structures does raise questions. Before 1997, for instance, it was widely believed that convection–diffusion equations could not be gradient flows. This belief was contradicted by [27, 28]; apparently the question "which systems can be gradient flows" is a non-trivial one. As another example, common building blocks of these gradient-flow structures, such as the Wasserstein metric, appear to be mathematical, non-physical constructs — can one give these an interpretation in terms of physics, chemistry, or other modeling contexts?

In [1] the authors give a suggestion for an organizing principle behind the observed variety in systems and gradient flows. For the example of the entropy-Wasserstein gradient flow (see below) they show how the gradient-flow structure itself is closely related to the probabilistic structure of a system of stochastic particles. This connection explains many aspects of the gradient flow, such as the origin of both the entropy and the Wasserstein metric and the interpretation of the discrete-time approximation.

The result of [1] also suggests that this connection between gradient-flow structures and stochastic particle systems may be much more general. In this paper we explore this idea for the following diffusion equation with convection and decay:

$$\partial_t u = \Delta u + \operatorname{div}(u\nabla\Psi) - \lambda u, \quad \text{in } \mathbb{R}^d \times (0,\infty), \tag{1}$$

with $\Psi \in C_b^2(\mathbb{R}^d)$ and $\lambda \geq 0$. We contribute two main results to the theory of this type of equations: first, we derive a new gradient-flow formulation for Eq. (1), and secondly, since this formulation is constructed along the lines of [1], we automatically connect this gradient flow to microscopic systems of diffusing particles, and show that the gradient-flow structure arises from the probabilistic structure of these particle systems.

The paper is organized as follows. In the remainder of this introductory section we develop the required concepts and formulate the main aim of this paper in a little more detail. Next, we recall the central notions of this paper in Sec. 2. We proceed with our microscopic models and the corresponding results in Secs. 3 and 4, and we wrap up with a general discussion in Sec. 5. In the Appendix we give a description and the proof of an existing large-deviation result in a language that is more suited to this paper.

1.2. *Variational formulations*

In this paper we study iterative variational schemes on some space $\mathcal{X}$ of the form

$$\text{Given } \rho^{k-1}, \text{ choose } \rho^k \in \operatorname*{arg\,min}_{\rho\in\mathcal{X}} \mathcal{K}^h(\rho\,|\,\rho^{k-1}), \tag{2}$$

which will approximate the solution of an evolution equation as $h \to 0$. The following examples illustrate the main ideas.

Example 1 (Hilbert-space gradient flows). If $\mathcal{X}$ is a Hilbert space and the functional $\mathcal{K}^h$ is of the form

$$\mathcal{K}^h(\rho\,|\,\overline{\rho}) = \mathcal{F}(\rho) + \frac{1}{2h}\|\rho - \overline{\rho}\|^2 \tag{3}$$

for some smooth functional $\mathcal{F}$, then the minimization problem (2) gives the stationarity condition

$$\frac{\rho^k - \rho^{k-1}}{h} = -\mathrm{grad}\,\mathcal{F}(\rho^k).$$

In this one recognizes the backward Euler approximation of the continuous-time gradient flow:

$$\partial_t \rho = -\mathrm{grad}\,\mathcal{F}(\rho). \tag{4}$$

The time-discrete variational form (3) illustrates how in gradient flows the evolution is driven by a trade-off between two competing effects. An *energy functional* $\mathcal{F} : \mathcal{X} \to \mathbb{R} \cup \{\infty\}$ drives the system towards lower values of the energy; at the same time a *dissipation mechanism* (here quantified by the norm $\|\cdot\|$) acts as a selection principle among all directions that decrease $\mathcal{F}$.

If one chooses $\mathcal{X} = L^2(\mathbb{R}^d)$ and $\mathcal{F}(\rho) = \frac{1}{2}\int |\nabla\rho|^2$, then (4) simply becomes the diffusion equation. However, it is not possible to describe convection in this way. The next example shows that convection–diffusion equations are nevertheless gradient flows, in a more general context.

Example 2 (Wasserstein gradient flows). Instead of a Hilbert space, we now consider the metric space $\mathcal{X} = \mathcal{P}_2(\mathbb{R}^d)$ of probability measures with finite second moment, equipped with the Wasserstein metric d (see Sec. 2.1). Similarly to (3), let (where the subscript *FP* stands for "Fokker–Planck"):

$$\mathcal{K}^h_{FP}(\rho\,|\,\overline{\rho}) := \frac{1}{2}\mathcal{F}(\rho) - \frac{1}{2}\mathcal{F}(\overline{\rho}) + \frac{1}{4h}d^2(\overline{\rho}, \rho), \tag{5}$$

where $\mathcal{F}(\rho) = \mathcal{S}(\rho) + \mathcal{E}(\rho)$ is the Helmholtz free energy, and

$$\begin{aligned} \mathcal{S}(\rho) &:= \begin{cases} \displaystyle\int \log f(y)\rho(dy) & \text{if } \rho(dy) = f(y)dy, \\ \infty & \text{otherwise}, \end{cases} \\ \mathcal{E}(\rho) &:= \int \Psi(y)\rho(dy), \end{aligned} \tag{6}$$

are the (negative) Gibbs–Boltzmann entropy and the energy arising from a potential Ψ. Note that in comparison to (3) we have subtracted the free energy of the previous state, and multiplied the expression by $1/2$. Both are done in view of the connection to large-deviation rate functionals that we establish below; of course neither change affects the minimization properties of $\mathcal{K}^h_{FP}(\cdot\,|\,\overline{\rho})$.

It was first observed by Jordan, Kinderlehrer and Otto [27, 28] that the time-discrete process defined by (2) and (5) converges to the solution of the Fokker–Planck equation:

$$\partial_t u = \Delta u + \operatorname{div}(u\nabla\Psi) \quad \text{in } \mathbb{R}^d \times (0,\infty). \tag{7}$$

We see that, in the same sense as the previous example, the Fokker–Planck equation is a gradient flow of free energy with respect to the Wasserstein metric. For future reference, we duplicate their main theorem as follows (where the superscript a denotes absolutely continuous).

Theorem 1 ([28]). *Let $\rho^0 \in \mathcal{P}_2^a(\mathbb{R}^d)$, and define the sequence $\{\rho^{h,k}\}_{k\geq 0}$ by:*

$$\begin{aligned} &\rho^{h,0} = \rho^0, \\ &\rho^{h,k} \in \operatorname*{arg\,min}_{\rho\in\mathcal{P}_2(\mathbb{R}^d)} \mathcal{K}^h_{FP}(\rho\,|\,\rho^{h,k-1}), \quad k \geq 1. \end{aligned}$$

These minimizers exist uniquely, and as $h \to 0$, the function $\rho^{h,\lfloor t/h\rfloor}$ converges weakly in $L^1(\mathbb{R}^d \times (0,T))$ to the solution of (7) with initial condition ρ^0.

Actually, [28] provides an argument to extend this result to weak convergence in $L^1(\mathbb{R}^d)$ for almost every $t \in (0,T)$ and strong convergence in $L^1(\mathbb{R}^d, (0,T))$.

While various generalizations of Hilbert-space gradient flows were known for some time [3, 12, 32], this result meant a breakthrough by extending the concept to a large and important class of evolution equations. In addition to inspiring a great amount of research into gradient flows in Wasserstein spaces and in general metric spaces, in a variety of functional-analytic settings [43, 35, 4, 46], it also gave rise to many fruitful connections between partial differential equations, optimal transport theory, geometry, functional inequalities, and probability; see [48, 49] for an overview.

Example 3 (Exponential decay). As in some other cases [3, 32], it will be useful to consider more general time-discrete constructions, namely of the form

$$\mathcal{K}^h(a\,|\,\overline{a}) = \mathcal{F}(a;\overline{a}) + f^h(\overline{a}, a), \tag{8}$$

for some function f^h. In this example, fix some $0 < r^h < 1$ and let the state space be $\mathcal{X} = \mathbb{R}^+$. Take for $\mathcal{F}$ a mixing entropy with parameter $\overline{a}$,

$$\mathcal{F}(a;\overline{a}) := a\log a + (\overline{a} - a)\log(\overline{a} - a), \quad \text{for } 0 < a < \overline{a}, \tag{9}$$

and for f^h the expression[a]

$$f^h(\overline{a}, a) := -a\log r^h - (\overline{a} - a)\log(1 - r^h). \tag{10}$$

[a]As suggested by one of the referees, this particular form (8) + (9) arises as the quenched large-deviation rate of a system of independent exponentially distributed decay processes, with $a = \frac{1}{n}\#\{\text{non-decayed } X_i(h)\}$ and $\bar{a} = \frac{1}{n}\#\{\text{non-decayed } X_i(0)\}$.

Then, the unique minimizer of (8) is $a = r^h\overline{a}$. While this construction may appear to be a convoluted way of arriving at this result, in fact it appears *naturally* in the context of a specific stochastic system of particles, as we show below. In the limit $h \to 0$ it will describe the term $-\lambda u$ in (1) which is associated with decay, as is illustrated by the following simple result.

Theorem 2. *Let $\mathcal{K}^h$ be given as in (8)–(10) with $r^h := e^{-\lambda h}$. Let $a^0 \in \mathbb{R}^+$ be fixed and define the sequence $\{a^{h,k}\}_{k\geq 0}$ by*

$$a^{h,0} = a^0,$$

$$a^{h,k} \in \operatorname*{arg\,min}_{a\in\mathbb{R}^+} \mathcal{K}^h(a \,|\, a^{h,k-1}), \quad k \geq 1.$$

Then as $h \to 0$ the function $t \mapsto a^{h,\lfloor t/h \rfloor}$ converges in time to the solution $t \mapsto a^0 e^{-\lambda t}$ of $\partial_t u = -\lambda u$.

The proof follows from remarking that $a^{h,k} = a^0 e^{-\lambda kh}$.

Below we will consider this construction in integrated form:

$$\mathcal{K}^h_{Dc}(\rho \,|\, \overline{\rho}) := -\mathcal{S}(\overline{\rho}) + \mathcal{S}(\rho) + \mathcal{S}(\overline{\rho} - \rho) - |\rho| \log r^h - |\overline{\rho} - \rho| \log(1 - r^h)$$

(the subscript Dc stands for "Decay equation") on the space of non-negative Borel measures $\mathcal{M}^+(\mathbb{R}^d)$ with the total variation norm $|\rho| := \rho(\mathbb{R}^d)$. Observe that compared to (8)–(10), we have an additional term $-\mathcal{S}(\overline{\rho})$. This term does not influence the minimizer, but we have added it here to ensure that the minimum is 0, which will be needed below.

Synthesis of Examples 2 and 3. In the results that we prove in this paper, the last two examples are merged in a single variational scheme. In the simplest case, for instance, where $\Psi \equiv 0$, the discrete algorithm approximating (1) becomes

$$\begin{aligned}\rho^k \in \operatorname*{arg\,min}_{\rho\in\mathcal{M}^+(\mathbb{R}^d)} \inf_{\rho_{ND}:|\rho+\rho_{ND}|=|\rho^{k-1}|} & -\frac{1}{2}\mathcal{S}(\rho + \rho_{ND}) - \frac{1}{2}\mathcal{S}(\rho^{k-1}) \\ & + \frac{1}{4h} d^2(\rho + \rho_{ND}, \rho^{k-1}) + \mathcal{S}(\rho) + \mathcal{S}(\rho_{ND}) - |\rho| \log r^h \\ & - |\rho_{ND}| \log(1 - r^h). \end{aligned} \tag{11}$$

To interpret the formula above, one should realize that the infimum over the measure ρ_{ND} in the formula above represents a choice: in each time step, the system designates a portion $\rho_{ND} \geq 0$ for decay (the index *ND* stands for "Normal to Decayed"), while the other part $\rho \geq 0$ remains "normal".

The terms inside the infimum can be written as $\mathcal{K}^h_{FP}(\rho+\rho_{ND} \,|\, \rho^{k-1}) + \mathcal{K}^h_{Dc}(\rho \,|\, \rho+\rho_{ND})$, and one can understand the structure of (11) through this splitting. The functional $\mathcal{K}^h_{FP}(\rho + \rho_{ND} \,|\, \rho^{k-1})$ characterizes a single time-step of diffusion of ρ^{k-1}, according to Theorem 1. Decay is left out of this step, since the joint mass $\rho + \rho_{ND}$ is independent of the distribution over normal (ρ) and decayed matter (ρ_{ND}). In a second step, given a choice for $\rho + \rho_{ND}$, the second functional $\mathcal{K}^h_{Dc}(\rho \,|\, \rho + \rho_{ND})$

describes how the total mass $\rho + \rho_{ND}$ is divided over ρ and ρ_{ND}, according to Theorem 2. As such, we can interpret $\rho + \rho_{ND}$ as an intermediate state between ρ^{k-1} and ρ.

1.3. *From microscopic model to large deviations*

We claimed above that the approximation scheme arises naturally in the context of stochastic particle systems. We now describe this context. It is well known (going back at least to Einstein [20]) that the diffusion equation

$$\partial_t u = \Delta u, \quad \text{in } \mathbb{R}^d \times (0,\infty), \tag{12}$$

is the macroscopic (hydrodynamic, continuum) limit of a wide range of stochastic particle systems [13]. Here we focus on one such system, composed of independent Brownian particles.

More specifically, let all particles $1,\ldots,n$ be initially distributed according to some fixed $\overline{\rho} \in \mathcal{P}(\mathbb{R}^d)$, and, for a fixed time interval $h > 0$, let each particle $i = 1,\ldots,n$ move to a new position Y_i^h, where the probability of moving from x to y is given by the density (which is identical for all particles)

$$\theta^h(y-x) := \frac{1}{(4\pi h)^{d/2}} \exp\left(-\frac{|x-y|^2}{4h}\right). \tag{13}$$

The empirical measure $L_n^h := n^{-1}\sum_{i=1}^n \delta_{Y_i^h}$ then is a random probability measure that describes the distribution of all n particles in space at time h. This measure converges (as $n \to \infty$) to $\overline{\rho} * \theta^h$, the solution of (12) at time h with initial condition $\overline{\rho}$.

The speed of this convergence is characterized by a *large-deviation principle*, which we discuss in Sec. 2.2. It states that the probability of finding L_n^h close to some $\rho \in \mathcal{P}(\mathbb{R}^d)$ converges exponentially to zero with rate $n\mathcal{J}_{Df}^h(\rho\,|\,\overline{\rho})$ (the subscript stands for "Diffusion equation"):

$$\mathrm{Prob}(L_n^h \approx \rho\,|\,L_n^0 \approx \overline{\rho}) \sim \exp(-n\mathcal{J}_{Df}^h(\rho\,|\,\overline{\rho})) \quad \text{as } n \to \infty.$$

The *rate functional* $\mathcal{J}_{Df}^h(\cdot\,|\,\overline{\rho})$ is non-negative and minimized by the solution of (12) at time h.

1.4. *From large deviations to Wasserstein gradient flow*

When restricting ourselves to the diffusion equation (12), the gradient-flow functional (5) reduces to

$$\mathcal{K}_{Df}^h(\rho\,|\,\overline{\rho}) := \frac{1}{2}\mathcal{S}(\rho) - \frac{1}{2}\mathcal{S}(\overline{\rho}) + \frac{1}{4h}d^2(\overline{\rho},\rho).$$

Recent results [1, 18] have shown that, under suitable assumptions, not only the minimizers of $\mathcal{J}_{Df}^h$ and $\mathcal{K}_{Df}^h$ have the same limit, but the two are in fact strongly related. Since we expect this statement to be generally true, we pose it here as a conjecture. It will be convenient to introduce the set:

$$\mathcal{P}_2^{\mathcal{S}}(\mathbb{R}^d) := \left\{\rho \in \mathcal{P}(\mathbb{R}^d) : \int |x|^2 d\rho < \infty, \mathcal{S}(\rho) < \infty\right\}.$$

Conjecture 3. *For any fixed* $\overline{\rho} \in \mathcal{P}_2^{\mathcal{S}}(\mathbb{R}^d)$ *there holds*

$$\mathcal{J}_{Df}^h(\cdot\,|\,\overline{\rho}) - \frac{1}{4h}d^2(\overline{\rho},\cdot) \xrightarrow[h\to 0]{M} \frac{1}{2}\mathcal{S}(\cdot) - \frac{1}{2}\mathcal{S}(\overline{\rho}) = \mathcal{K}_{Df}^h(\cdot\,|\,\overline{\rho}) - \frac{1}{4h}d^2(\overline{\rho},\cdot) \tag{14}$$

in the sense of Mosco convergence, where the lower bound holds in $\mathcal{P}_2(\mathbb{R}^d)$ *with the narrow topology, and the recovery sequence holds in the topology defined by convergence in Wasserstein distance plus convergence in entropy* $\mathcal{S}$ (*see Sec.* 2.3).

This conjecture was first proven in [1] under the restriction that both ρ and $\overline{\rho}$ in $\mathcal{J}_{Df}^h(\rho\,|\,\overline{\rho})$ are sufficiently close to uniform distributions on a bounded interval in $\mathbb{R}$. In [18], the result was generalized to $\mathbb{R}$ for any $\overline{\rho}$ with bounded Fisher information.

Note that the term $-(4h)^{-1}d^2(\overline{\rho},\,\cdot)$ appears on both sides of (14). The role of this term is to compensate the singular behavior of both $\mathcal{J}_{Df}^h$ and $\mathcal{K}_{Df}^h$ in the limit $h \to 0$. Morally, the conjecture states that

$$\text{as } h \to 0, \quad \mathcal{J}_{Df}^h(\cdot\,|\,\overline{\rho}) \approx \mathcal{K}_{Df}^h(\cdot\,|\,\overline{\rho}).$$

This connection shows how the functional $\mathcal{K}_{Df}^h$, which defines the time-discretized gradient flow, can be interpreted physically: as the large-deviation rate functional of the microscopic model.

1.5. *Overview of this work*

In this paper we extend the results of [1, 18] to Eq. (1). Although the results in the latter already includes the Fokker–Planck equation (7), this paper uses very different techniques and yields results under different assumptions on the potential Ψ. The main results of this paper are of the same form as Theorem 1 and Conjecture 3.

We divide the arguments, and the paper, into two parts. In the first part we discuss diffusion with drift but without decay ($\Psi \not\equiv 0$, $\lambda = 0$ in (1)). First we construct a system of Brownian particles with drift that models the Fokker–Planck equation (7), and then derive a corresponding large-deviation principle. In our first main result, Theorem 9, we show that for small times the large-deviation rate functional of the micro model relates to $\mathcal{K}_{FP}^h$ in the same sense as in Conjecture 3 for the diffusion equation. Note that the expression for the gradient-flow functional $\mathcal{K}_{FP}^h$ is already known from [28]; the novelty of the current result lies in the connection to the microscopic particle system.

The second part of the paper concerns the diffusion equation with decay ($\lambda > 0$, and for ease of notation we first take $\Psi \equiv 0$):

$$\partial_t u = \Delta u - \lambda u, \quad \text{in } \mathbb{R}^d \times (0, \infty). \tag{15}$$

Again, we devise a particle system that models this equation microscopically, and derive a corresponding large-deviation principle. In the second main result of this paper, Theorem 11, we show that the large-deviation rate functional relates to an energy-dissipation functional (74) in the same way as in Conjecture 3. Finally, in Theorem 12 we show that the minimizers of this new functional indeed approximate the solution of (15) in the sense of Theorem 1. In this case, the novelty lies in both the expression of the energy-dissipation functional, and in its connection to the microscopic system.

2. Background

2.1. *Wasserstein distance*

In the Kantorovich formulation of the optimal transport problem, a transport plan between two measures $\overline{\rho}, \rho \in \mathcal{P}(\mathbb{R}^d)$ is a measure in the set

$$\Gamma(\overline{\rho}, \rho) := \{q \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d) : \pi^1 q = \overline{\rho} \text{ and } \pi^2 q = \rho\},$$

where we denote the marginals of q by

$$\pi^1 q(B) := q(B \times \mathbb{R}^d) \quad \text{and} \quad \pi^2 q(B) := q(\mathbb{R}^d \times B) \quad \text{for all Borel sets } B \subset \mathbb{R}^d.$$

In the particular case of the 2-Wasserstein distance (henceforth simply called the Wasserstein distance), the unit cost of transporting an infinitesimal mass from position x to y is taken to be $|x-y|^2$. One can then ask for the optimal transport plan that transports all mass from a measure $\overline{\rho}$ to another measure ρ. The minimum cost defines a metric on the space $\mathcal{P}_2(\mathbb{R}^d) := \{\rho \in \mathcal{P}(\mathbb{R}^d) : \int |x|^2 d\rho < \infty\}$.

Definition 4 (Wasserstein distance).

$$d^2(\overline{\rho}, \rho) := \inf_{q \in \Gamma(\overline{\rho}, \rho)} \iint |x-y|^2 q(dx\, dy).$$

An important property of the Wasserstein distance is that a sequence $\{\rho_h\}_h \in \mathcal{P}_2(\mathbb{R}^d)$ converges to ρ in the Wasserstein distance as $h \to 0$ if and only if [48, Theorem 7.12]

(1) $\rho_h \rightharpoonup \rho$ (see Sec. 2.3),
(2) $\int x^2 \rho_h(dx) \to \int x^2 \rho(dx)$.

Observe that the Wasserstein distance is still meaningful for measures $\overline{\rho}, \rho \in \mathcal{M}^+(\mathbb{R}^d)$ that are not necessarily probability measures, as long as $|\overline{\rho}| = |\rho|$. With this generalization we have that

$$d^2(\rho_1+\rho_2,\rho_3+\rho_4) \le d^2(\rho_1,\rho_3)+d^2(\rho_2,\rho_4) \quad \text{for all } \rho_{1,2,3,4} \text{ with } |\rho_1|=|\rho_3| \text{ and } |\rho_2|=|\rho_4|. \tag{16}$$

This property will be used later in the paper.

2.2. *Large deviations*

Recall from the law of large numbers that with probability 1, in the large-n limit the expectation $\mathbb{E}L_1^h$ is the only event that occurs (see, for example, [16, Theorem 11.4.1]). In this limit, any other event is considered a large deviation from this expected behavior. A large-deviation principle characterizes the unlikeliness of such event by the speed of convergence of its probability to 0. To illustrate this, we briefly switch to a more abstract notation.

Definition 5. A sequence X_n of random variables with variables in a topological space $\mathcal{X}$ satisfies the *large-deviation principle* with speed n and *rate functional* $\mathcal{J}:\mathcal{X}\to[0,\infty]$ whenever:

(1) $\mathcal{J}$ is not identically ∞, and $\mathcal{J}^{-1}[0,c]$ is compact for all $c<\infty$;
(2) $\liminf_{n\to\infty}\frac{1}{n}\log\mathrm{Prob}(X_n\in U)\ge-\inf_{x\in U}\mathcal{J}(x)$ for all open sets $U\subset\mathcal{X}$;
(3) $\limsup_{n\to\infty}\frac{1}{n}\log\mathrm{Prob}(X_n\in C)\le-\inf_{x\in C}\mathcal{J}(x)$ for all closed sets $C\subset\mathcal{X}$.

The rate functional $\mathcal{J}$ is non-negative and achieves its minimum of zero at the most probable behavior of X_n. The right-hand infimum reflects the general principle that "any large deviation is done in the least unlikely of all the unlikely ways" [15, p. 10]. A related mathematical result is the *contraction principle* [14, Theorem 4.2.1], which states the following. Let $p:\mathcal{X}\to\mathcal{Y}$ be a continuous map, and $Y_n:=p(X_n)$ the corresponding random variables. Then Y_n satisfies a large-deviation principle similar to the one above, with rate functional $\inf_{x\in\mathcal{X}:p(x)=y}\mathcal{J}(x)$. This contraction principle will be used throughout this paper. For instance, it explains the role of the minimization in (11).

2.3. *Mosco convergence*

A useful tool in the study of sequences of minimization problems is Γ-convergence [10]. In particular, it is often used in the study of large deviations [1, Lemma 2] and gradient flows (cf. [12, 47]). Moreover, in [31], Γ-convergence is used to connect large deviations to optimal transport. In many cases, it is convenient to require that the recovery sequence of the Γ-convergence exists in a stronger topology (cf. [4, Remark 2.0.5] or [37]): the resulting notion of convergence is known as Mosco convergence [40]. In results that are related to this paper, a further analysis reveals that Mosco convergence is indeed satisfied (cf. [1, Theorem 3; 18, Theorem 1.1]). In this sense it provides a natural notion for the purpose of this study.

Definition 6. Let $\mathcal{X}$ be a space with two first-countable (e.g., metrizable) topologies $\tau_w \subset \tau_s$. A sequence of functionals $\{\mathcal{F}^h\}_h$ on $\mathcal{X}$ Mosco-converges[b] to $\mathcal{F} : \mathcal{X} \to \mathbb{R} \cup \{\infty\}$ as $h \to 0$, written as $\mathcal{F}^h \xrightarrow{M} \mathcal{F}$, whenever

(1) (*Lower bound*) for any sequence $\rho^h \xrightarrow[h\to 0]{\tau_w} \rho$ in $\mathcal{X}$ there holds

$$\liminf_{h\to 0} \mathcal{F}^h(\rho^h) \geq \mathcal{F}(\rho);$$

(2) (*Recovery sequence*) for all $\rho \in \mathcal{X}$ there is a sequence $\rho^h \xrightarrow[h\to 0]{\tau_s} \rho$ in $\mathcal{X}$ such that

$$\limsup_{h\to 0} \mathcal{F}^h(\rho^h) \leq \mathcal{F}(\rho).$$

In this paper we take $\mathcal{X} = \mathcal{P}_2^{\mathcal{S}}(\mathbb{R}^d)$ (defined in Sec. 1.4), and for τ_w we take the narrow topology, characterized by narrow convergence:

$$\rho^h \rightharpoonup \rho \text{ if and only if } \int \phi(x)\rho^h(dx) \to \int \phi(x)\rho(dx) \quad \text{for all } \phi \in C_b(\mathbb{R}^d).$$

For the strong topology τ_s, we take the weakest topology such that all functionals $\rho \mapsto \int x^2\rho(dx)$, $\rho \mapsto \mathcal{S}(\rho)$ and $\rho \mapsto \int \phi(x)\rho(dx)$ for all $\phi \in C_b(\mathbb{R}^d)$ are continuous. Since this topology is first-countable, convergence in $(\mathcal{P}_2^{\mathcal{S}}, \tau_s)$ is characterized by convergence in the Wasserstein topology plus convergence of the entropy functional $\mathcal{S}$. In fact, we prove below that convergence in this topology implies strong L^1-convergence of a subsequence and its entropies. These important facts will be used to prove the Mosco convergence Theorems 9 and 11. Let $\mathcal{L}^d$ be the d-dimensional Lebesgue measure.

Lemma 7. *Let $\rho^h \to \rho$ in $\mathcal{P}_2^{\mathcal{S}}(\mathbb{R}^d)$ in the strong topology, i.e.:*

$$d(\rho^h, \rho) \to 0, \text{ in the Wasserstein metric,} \tag{17}$$

$$\mathcal{S}(\rho^h) \to \mathcal{S}(\rho). \tag{18}$$

Then ρ^h and ρ are $\mathcal{L}^d$-absolutely continuous and can be identified with their densities, i.e. $\rho^h, \rho \in L^1(\mathbb{R}^d)$, and there is a subsequence such that

$$\rho^h \to \rho, \tag{19}$$

$$\rho^h \log \rho^h \to \rho \log \rho, \tag{20}$$

strongly in $L^1(\mathbb{R}^d)$.

[b] We slightly generalize the usual concept of Mosco convergence, where $\mathcal{X}$ should be a Banach space where the weak topology is defined by duality with $\mathcal{X}^*$.

Proof. Step I — Decomposition of the entropy. To deal with the fact that $\mathcal{S}$ is not bounded from below, we rewrite $\mathcal{S}$ in the following way. Define, for any $\alpha \in \mathbb{R}$ with $\alpha > d$

$$c^{-1} := \int_{\mathbb{R}^d} \frac{1}{(1+|x|)^\alpha} dx,$$

$$\nu(dx) = \nu(x)dx = \frac{c}{(1+|x|)^\alpha} dx,$$

and let $\mathcal{H}$ be the relative entropy on two probability measures $\gamma, \nu \in \mathcal{P}(\mathbb{R}^d)$:

$$\mathcal{H}(\gamma \,|\, \nu) := \begin{cases} \displaystyle\int \frac{d\gamma}{d\nu}(x) \log \frac{d\gamma}{d\nu}(x)\nu(dx) & \text{if } \gamma \ll \nu, \\ +\infty & \text{otherwise.} \end{cases} \tag{21}$$

(Note that $\mathcal{S}(\rho) = \mathcal{H}(\rho \,|\, \mathcal{L}^d)$.) Then for any $\rho \in \mathcal{P}_2^{\mathcal{S}}$, we can write

$$\begin{aligned} \mathcal{S}(\rho) = \int_{\mathbb{R}^d} \rho \log \rho dx &= \int_{\mathbb{R}^d} \frac{\rho}{\nu} \log\left(\frac{\rho}{\nu}\right) \nu dx + \int_{\mathbb{R}^d} \rho \log(\nu) dx \\ &= \mathcal{H}(\rho \,|\, \nu) + \log c - \alpha \int_{\mathbb{R}^d} \rho \log(1+|x|) dx. \end{aligned} \tag{22}$$

By (17) and [4, Lemma 5.1.7]

$$\int_{\mathbb{R}^d} \rho^h(x)\phi(x)dx \to \int_{\mathbb{R}^d} \rho(x)\phi(x)dx \tag{23}$$

for all continuous functions $\phi : \mathbb{R}^d \to \mathbb{R}$ such that $|\phi(x)| \le A + B|x|^2$ for all $x \in \mathbb{R}^d$, for some $A, B \ge 0$. This implies that the last term on the right-hand side of (22) converges:

$$\alpha \int_{\mathbb{R}^d} \rho^h(x) \log(1+|x|)dx \to \alpha \int_{\mathbb{R}^d} \rho(x) \log(1+|x|)dx, \tag{24}$$

so that the study of $\mathcal{S}(\rho^h)$ can be reduced to the study of $\mathcal{H}(\rho^h \,|\, \nu)$.

Step II — Convergence of the plans. Define the measures $\gamma^h \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R})$ by

$$\int_{\mathbb{R}^d \times \mathbb{R}} \psi(x,y)\gamma^h(dx\,dy) = \int_{\mathbb{R}^d} \psi\left(x, \frac{\rho^h(x)}{\nu(x)}\right) \nu(x)dx \quad \text{for all } \psi \in C_b(\mathbb{R}^d \times \mathbb{R}).$$

The marginals $\pi^1\gamma^h$ and $\pi^2\gamma^h$ then satisfy

$$\begin{aligned} \int_{\mathbb{R}^d} \phi(x)\pi^1\gamma^h(dx) &= \int_{\mathbb{R}^d} \phi(x)\nu(x)dx, \\ \int_{\mathbb{R}} \varphi(y)\pi^2\gamma^h(dy) &= \int_{\mathbb{R}^d} \varphi\left(\frac{\rho^h(x)}{\nu(x)}\right) \nu(x)dx, \end{aligned} \tag{25}$$

for all $\phi \in C_b(\mathbb{R}^d)$ and for all $\varphi \in C_b(\mathbb{R})$. We claim that

- there exists a $\gamma \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R})$ such that, up to subsequences, $\gamma^h \rightharpoonup \gamma$ (narrowly);
- the barycentric projection (27) of the limit γ, with respect to ν, is ρ/ν.

In order to prove the first part of the claim, we note that by [4, Lemma 5.2.2], if the marginals of γ^h are tight, then γ^h is also tight, and thus (by [4, Theorem 5.1.3]) relatively compact, with respect to the narrow topology of $\mathcal{P}(\mathbb{R}^d \times \mathbb{R})$. By (25) the first marginal does not depend on h. For the second marginal we use the following integral condition for tightness (see [4, Remark 5.1.5]): "*if there exists a function* $G : \mathbb{R} \to [0, +\infty]$, *whose sublevels are compact in* $\mathbb{R}$, *such that*

$$\sup_{h \in \mathbb{N}} \int_{\mathbb{R}} G(y) \pi^2 \gamma^h(dy) < +\infty,$$

then $\{\pi^2 \gamma^h\}$ *is tight*". We can choose, as in [4, Eq. (9.4.2)], the non-negative, lower semicontinuous, strictly convex function

$$G(s) := \begin{cases} s(\log s - 1) + 1 & \text{if } s > 0, \\ 1 & \text{if } s = 0, \\ +\infty & \text{if } s < 0, \end{cases}$$

defined on $\mathbb{R}$, and observe that

$$\int_{\mathbb{R}} G(y) \pi^2 \gamma^h(dy) = \int_{\mathbb{R}^d} G\left(\frac{\rho^h(x)}{\nu(x)}\right) \nu(x) dx = \mathcal{H}(\rho^h \,|\, \nu).$$

The last term is bounded, owing to (23), (22), and (24). We conclude that γ^h is relatively compact and therefore, up to subsequences, γ^h converges to a measure $\gamma \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R})$.

In order to prove the second part of the claim, note that by disintegration of measures [4, Theorem 5.3.1], there exists a family $\{\mu_x\}_{x \in \mathbb{R}^d} \subset \mathcal{P}(\mathbb{R})$ such that

$$\int_{\mathbb{R}^d \times \mathbb{R}} \psi(x, y) \gamma(dx dy) = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}} \psi(x, y) \mu_x(dy) \right) \nu(dx) \tag{26}$$

for every Borel map $\psi : \mathbb{R}^d \times \mathbb{R} \to [0, +\infty]$. We want to identify the barycentric projection of γ with respect to ν, that is, the function

$$x \mapsto \int_{\mathbb{R}} y \mu_x(dy), \tag{27}$$

with ρ/ν. This can be done if we can choose as test function ψ a function of the form $(x, y) \to \phi(x) y$, with $\phi \in C_b(\mathbb{R}^d)$. Since such a function is not bounded, we first need to check that it is uniformly integrable. Since $\mathcal{H}(\rho^h \,|\, \nu)$ is bounded, there is a constant $C_1 > 0$ such that, for all $R > 1$,

$$\begin{aligned}
C_1 &> \sup_h \int_{\mathbb{R}^d} G\left(\frac{\rho^h(x)}{\nu(x)}\right) \nu(x)dx \\
&\geq \sup_h \int_{\{\rho^h > R\}} G\left(\frac{\rho^h(x)}{\nu(x)}\right) \nu(x)dx \\
&= \sup_h \int_{\{\rho^h > R\}} \rho^h(x) \log\left(\rho^h(x)\frac{(1+|x|)^\alpha}{c}\right) dx \\
&\geq \sup_h \int_{\{\rho^h > R\}} \rho^h(x) \log R - \rho^h \log c + \alpha\rho^h(x) \log(1+|x|)dx \\
&\geq \log(R) \sup_h \int_{\{\rho^h > R\}} \rho^h(x)dx - \log c - \sup_h \alpha \int_{\mathbb{R}^d} \rho^h(x) \log(1+|x|)dx \\
&\overset{(24)}{\geq} \log(R) \sup_h \int_{\{\rho^h > R\}} \rho^h(x)dx - C_2.
\end{aligned}$$

Therefore,

$$\lim_{R\to\infty} \sup_h \int_{\{\rho^h > R\}} \rho^h dx \leq \lim_{R\to\infty} \frac{C_1 + C_2}{\log(R)} = 0, \tag{28}$$

i.e. ρ^h is uniformly integrable. Since for every $\phi \in C_b(\mathbb{R}^d)$

$$\begin{aligned}
\lim_{R\to\infty} \sup_h \int_{\{\phi(x)y \geq R\}} \phi(x)y\, \gamma^h(dxdy) &\leq \lim_{R\to\infty} \sup_h \|\phi\|_\infty \int_{\{|y| \geq R/\|\phi\|_\infty\}} y\, \gamma^h(dxdy) \\
&= \lim_{R\to\infty} \sup_h \|\phi\|_\infty \int_{\{\rho^h \geq R/\|\phi\|_\infty\}} \rho^h(x)dx \overset{(28)}{=} 0,
\end{aligned}$$

we conclude that the function $\mathbb{R}^d \times \mathbb{R} \ni (x,y) \mapsto \phi(x)y \in \mathbb{R}$ is uniformly integrable with respect to the measures $\{\gamma^h\}$. Uniform integrability, owing to [4, Lemma 5.1.7], yields

$$\begin{aligned}
\lim_{h\to\infty} \int_{\mathbb{R}^d \times \mathbb{R}} \phi(x)y\, \gamma^h(dxdy) &= \int_{\mathbb{R}^d \times \mathbb{R}} \phi(x)y\, \gamma(dxdy) \\
&\overset{(26)}{=} \int_{\mathbb{R}^d} \phi(x)\left(\int_{\mathbb{R}} y\mu_x(dy)\right) \nu(dx).
\end{aligned}$$

On the other hand, by (18) we know that

$$\lim_{h\to\infty} \int_{\mathbb{R}^d \times \mathbb{R}} \phi(x)y\, \gamma^h(dxdy) = \lim_{h\to\infty} \int_{\mathbb{R}^d} \phi(x)\frac{\rho^h(x)}{\nu(x)}\nu(dx) = \int_{\mathbb{R}^d} \phi(x)\frac{\rho(x)}{\nu(x)}\nu(dx).$$

We conclude that the weak limit of the densities is equal to the barycentric projection of the limit plans:

$$\frac{\rho(x)}{\nu(x)} = \int_{\mathbb{R}} y\mu_x(dy) \quad \text{for a.e. } x \in \mathbb{R}^d. \tag{29}$$

Step III — Pointwise convergence. We compute

$$\begin{aligned}\liminf_{h\to\infty}\mathcal{H}(\rho^h\,|\,\nu) &= \liminf_{h\to\infty}\int_{\mathbb{R}^d} G\left(\frac{\rho^h(x)}{\nu(x)}\right)\nu(dx)\\ &= \liminf_{h\to\infty}\int_{\mathbb{R}^d\times\mathbb{R}} G(y)\gamma^h(dxdy)\\ &\geq \int_{\mathbb{R}^d\times\mathbb{R}} G(y)\gamma(dxdy)\\ &= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}} G(y)\mu_x(dy)\right)\nu(dx)\\ &\geq \int_{\mathbb{R}^d} G\left(\int_{\mathbb{R}} y\mu_x(dy)\right)\nu(dx) \qquad (30)\\ &= \int_{\mathbb{R}^d} G\left(\frac{\rho(x)}{\nu(x)}\right)\nu(dx) = \mathcal{H}(\rho\,|\,\nu), \qquad (31)\end{aligned}$$

where, in the last three steps, we used (26), Jensen's inequality, and (29). Collecting all the computations we have

$$\begin{aligned}\mathcal{H}(\rho\,|\,\nu) &\overset{(22)}{=} \mathcal{S}(\rho) - \log c + \alpha\int_{\mathbb{R}^d}\rho(x)\log(1+|x|)dx\\ &\overset{(18),(24)}{=} \lim_{h\to\infty}\left\{\mathcal{S}(\rho^h) - \log c + \alpha\int_{\mathbb{R}^d}\rho^h(x)\log(1+|x|)dx\right\}\\ &\overset{(22)}{=} \liminf_{h\to\infty}\mathcal{H}(\rho^h\,|\,\nu)\\ &\overset{(31)}{\geq} \mathcal{H}(\rho\,|\,\nu).\end{aligned}$$

Therefore, the inequality in (30) must be an equality, which, by strict convexity of G, implies that μ_x is a Dirac delta concentrated in $\frac{\rho(x)}{\nu(x)}$, for a.e. $x\in\mathbb{R}^d$. As a consequence

$$\frac{\rho^h(x)}{\nu(x)} \to \frac{\rho(x)}{\nu(x)} \quad \text{for a.e. } x\in\mathbb{R}^d,$$

and therefore

$$\rho^h(x)\to\rho(x) \quad \text{for a.e. } x\in\mathbb{R}^d. \qquad (32)$$

Step IV — Strong convergence. To prove the strong convergence results (19) and (20), recall the following theorem from [8, Theorem 1] for any measure κ on $\mathbb{R}^d$ and non-negative $\rho^h,\rho\in L^1(\kappa)$:

$$\text{If} \int\rho^h d\kappa \to \int\rho d\kappa \quad \text{and} \quad \rho^h(x)\to\rho(x)\ \kappa\text{-a.e.},\ \text{then } \rho^h\to\rho \text{ strongly in } L^1(\kappa). \qquad (33)$$

Clearly, (19) follows from (32) and (33) by taking $\kappa=\mathcal{L}^d$.

In order to prove (20), let $G^h := G(\rho^h/\nu)$, $G^0 := G(\rho/\nu)$. Since G is continuous and $\rho^h \to \rho$ almost everywhere,

$$G^h(x) \to G^0(x) \quad \text{for a.e. } x \in \mathbb{R}^d. \tag{34}$$

Moreover, from the proof of (19), we know that

$$\int_{\mathbb{R}^d} G^h(x)\nu(dx) = \mathcal{H}(\rho^h \,|\, \nu) \to \mathcal{H}(\rho \,|\, \nu) = \int_{\mathbb{R}^d} G^0(x)\nu(dx). \tag{35}$$

Again by (33), now with $\kappa = \nu$, it follows from (34) and (35) that $G^h \to G^0$ strongly in $L^1(\nu)$. Therefore, because the density of ν is uniformly bounded

$$G^h \to G^0 \text{ strongly in } L^1(\mathbb{R}^d). \tag{36}$$

It now follows from (19) and (36) together with

$$\begin{aligned} \rho^h \log \rho^h &= G^h \nu + \rho^h \log(\nu) + \rho^h - \nu \\ &= G^h f + \rho^h(\log(c) + 1) - \alpha\rho^h \log(1 + |\cdot|) - \nu \end{aligned}$$

that, in order to prove (20) we only need to check that

$$\rho^h \log(1 + |\cdot|) \to \rho \log(1 + |\cdot|) \text{ strongly in } L^1(\mathbb{R}^d).$$

This follows from the uniform integrability of the first moments of ρ^h and from the strong L^1-convergence of ρ^h. Precisely, since $d(\rho^h, \rho) \to 0$, then ρ^h has uniformly integrable p-moments for all $p \in (0, 2)$. In particular, for every $\varepsilon > 0$ there exists $R_\varepsilon > 0$ such that

$$\sup_h \int_{|x| \geq R_\varepsilon} |x|\rho^h(x)dx \leq \varepsilon.$$

For all $\varepsilon > 0$ we estimate

$$\begin{aligned} &\int_{\mathbb{R}^d} |\rho^h(x) \log(1 + |x|) - \rho(x) \log(1 + |x|)|dx \\ &\quad \leq \int_{|x| < R_\varepsilon} |\rho^h(x) - \rho(x)| \log(1 + |x|)dx + \int_{|x| \geq R_\varepsilon} |x|\rho^h(x)dx + \int_{|x| \geq R_\varepsilon} |x|\rho(x)dx \\ &\quad \leq \|\rho^h - \rho\|_{L^1} \log(1 + R_\varepsilon) + 2\varepsilon \end{aligned}$$

and therefore, for all $\varepsilon > 0$

$$\lim_{h \to \infty} \int_{\mathbb{R}^d} |\rho^h \log(1 + |x|)dx - \rho \log(1 + |x|)|dx \leq 2\varepsilon.$$

By the arbitrariness of ε, we conclude strong L^1-convergence. □

3. Diffusion with Drift

In this section we discuss the case of diffusion with drift but without decay ($\Psi \not\equiv 0$, $\lambda = 0$), i.e. Eq. (7). First we describe the particle system that we use as a microscopic model for this equation, and derive the corresponding large-deviation principle. Next, we show that the large-deviation rate functional relates to the energy-dissipation functional (5) in a Mosco-convergence sense.

3.1. *Microscopic model*

Consider a system of n independent (i.e. non-interacting) point particles in $\mathbb{R}^d$. We wish $\overline{\rho} \in \mathcal{P}(\mathbb{R}^d)$ to represent the distribution of initial positions, and implement this as in [31]. For each n choose $x_i \in \mathbb{R}^d, 1 \le i \le n$ such that

$$\frac{1}{n}\sum_{i=1}^{n} \delta_{x_i} \rightharpoonup \overline{\rho} \quad \text{as } n \to \infty.$$

We then set the (deterministic) initial position[c] of particle $i \in \{1, \dots, n\}$ to be x_i.

The dynamics of the system is determined by the probability for particle i to move from x_i to a (random) position Y_i^h in some fixed time $h > 0$. We take this transition probability to be the fundamental solution $\eta^t(y; x)$ of the drift-diffusion equation (7), in the following sense.

Definition 8. We say that a mapping $\eta : \mathbb{R}^d \times [0, \infty) \to \mathcal{P}(\mathbb{R}^d)$ is a *fundamental solution* of the Fokker–Planck equation (7) whenever

(1) $\eta^{x,t}(B)$ is measurable in $x \in \mathbb{R}^d$ and $t \in [0, \infty)$ for all fixed Borel sets $B \subset \mathbb{R}^d$,
(2) for all $\phi \in C_b^{2,1}(\mathbb{R}^d \times [0, \infty))$ and $(x, T) \in \mathbb{R}^d \times [0, \infty)$ there holds:

$$\int_0^T \int (\partial_t \phi + \Delta\phi - \nabla\Psi \cdot \nabla\phi)\, \eta^{x,t}(dy)dt = \int \phi(y, T)\eta^{x,T}(dy) - \phi(x, 0).$$

If we assume that $\Psi \in C_b^2(\mathbb{R}^d)$, that is $\Psi \in C^2(\mathbb{R}^d)$ and $|\Psi|, |\nabla\Psi|$, and $|\Delta\Psi|$ are all bounded, then there exists an absolutely continuous fundamental solution with a density in $C^{2,1}(\mathbb{R}^d \times (0, \infty))$ [22, Theorem 1.10]. We can thus identify this fundamental solution $\eta^{x,t}$ with its density $\eta^t(\cdot\,; x)$.

Using this fundamental solution as the transition probability, the empirical measure $L_n^h = n^{-1}\sum_{i=1}^n \delta_{Y_i^h}$ will converge almost surely to $\overline{\rho} * \eta^h$, which is the solution to (7) at time h with initial condition $\overline{\rho}$ [16, Theorem 11.4.1]. In this sense the proposed system is indeed a microscopic precursor of this equation.

3.2. *From large deviations to Wasserstein gradient flow*

The sequence L_n^h satisfies a large-deviation principle with rate n and rate functional (see Corollary A.2 in the Appendix):

$$\mathcal{J}_{FP}^h(\rho \,|\, \overline{\rho}) := \inf_{q \in \Gamma(\overline{\rho}, \rho)} \mathcal{H}\big(q \,|\, \overline{\rho}\eta^h\big), \tag{37}$$

where $\mathcal{H}$ is the relative entropy (21) on $\mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$, and, by abuse of notation we write $(\overline{\rho}\eta^h)(dx\ dy) = \overline{\rho}(x)\eta^h(y; x)dx\, dy$.

We now prove the following relationship between this rate functional $\mathcal{J}_{FP}^h$ and the gradient-flow functional $\mathcal{K}_{FP}^h$ (given by (5)).

[c] This way of enforcing the initial distribution $\overline{\rho}$ is different from the approach of [1]. It provides a more direct result, and is easier to interpret; see Remark A.3 for a discussion.

Theorem 9. *Assume that Conjecture* 3 *holds, and that* $\Psi \in C_b^2(\mathbb{R}^d)$. *Then for any* $\overline{\rho} \in \mathcal{P}_2^{\mathcal{S}}(\mathbb{R}^d)$

$$\mathcal{J}_{FP}^h(\cdot\,|\,\overline{\rho}) - \frac{1}{4h}d^2(\overline{\rho},\cdot) \xrightarrow[h\to 0]{M} \frac{1}{2}\mathcal{S}(\cdot) - \frac{1}{2}\mathcal{S}(\overline{\rho}) + \frac{1}{2}\mathcal{E}(\cdot) - \frac{1}{2}\mathcal{E}(\overline{\rho}),$$

$$= \mathcal{K}_{FP}^h(\cdot\,|\,\overline{\rho}) - \frac{1}{4h}d^2(\overline{\rho},\cdot). \tag{38}$$

The proof relies heavily on an estimate of the fundamental solution η^h. To explain this estimate morally, observe that if Ψ is affine, i.e. $\Psi(x) = c\cdot x$, then the force field $\nabla\Psi$ is homogeneous, leading to constant drift c. In this simple case, the fundamental solution can be written explicitly:

$$\eta^t(y;x) = \frac{1}{(4\pi t)^{d/2}}e^{-|y-(x-ct)|^2/4t} = \theta^t(y-x)e^{-\frac{1}{2}c\cdot y+\frac{1}{2}c\cdot x-\frac{1}{4}|c|^2 t}, \tag{39}$$

where θ^t is again the diffusion kernel (13). Although for an arbitrary Ψ an analytic expression for the fundamental solution is generally difficult to find, the expression (39) above suggests that it can be estimated by something similar for small times. Below we see that this is indeed the case. We expect that this estimate is not a new result, but since we have not been able to find it in the literature we include the proof here for completeness.[d]

Lemma 10. *Assume* $\Psi \in C_b^2(\mathbb{R}^d)$, *and let* η *be the fundamental solution from Definition* 8. *Then there are* $\beta_0,\beta_1 \in \mathbb{R}$ *such that for every* $t > 0$:

$$\theta^t(y-x)e^{-\frac{1}{2}\Psi(y)+\frac{1}{2}\Psi(x)+\beta_0 t} \le \eta^t(y;x) \le \theta^t(y-x)e^{-\frac{1}{2}\Psi(y)+\frac{1}{2}\Psi(x)+\beta_1 t} \tag{40}$$

for almost every $x,y\in\mathbb{R}^d$.

Proof. For brevity we assume that $x = 0$ and $\Psi(0) \equiv 0$, and we omit the dependence on x. For $\beta\in\mathbb{R}$ define:

$$\zeta_\beta(y,t) := \eta^t(y) - \theta^t(y)e^{-\frac{1}{2}\Psi(y)+\beta t}.$$

By partial integration we obtain for all $0<\epsilon<T$ and $\phi\in C_b^{2,1}(\mathbb{R}^d\times[\epsilon,T])$:

$$\int_\epsilon^T\int(\partial_t\phi(y,t)+\Delta\phi(y,t)-\nabla\Psi(y)\cdot\nabla\phi(y,t))\,\zeta_\beta(y,t)dydt$$

$$= \int_\epsilon^T\int\phi(y,t)f_\beta(y,t)dydt + \int\phi(y,T)\zeta_\beta(y,T)dy - \int\phi(y,\epsilon)\zeta_\beta(y,\epsilon)dy \tag{41}$$

with:

$$f_\beta(y,t) := \left(-\frac{1}{2}\Delta\Psi(y)+\frac{1}{4}|\nabla\Psi(y)|^2+\beta\right)\theta^t(y)e^{-\frac{1}{2}\Psi(y)+\beta t}.$$

Because $\nabla\Psi$ and $\Delta\Psi$ are bounded, there are $\beta_0,\beta_1\in\mathbb{R}$ such that:

$$f_{\beta_0}(y,t) \le 0 \le f_{\beta_1}(y,t). \tag{42}$$

[d]See, for example, [6] for a similar, but not strong enough result.

First we exploit this inequality for β_1. Let ϕ be the solution of the adjoint problem:

$$-\partial_t\phi = \Delta\phi - \nabla\Psi\cdot\nabla\phi \tag{43}$$

with end condition:

$$\phi^T(y) := H(\zeta_{\beta_1}(y,T)),$$

where H is the Heaviside function. Again by [22, Theorem 1.10] there exists a positive fundamental solution η^* and hence a positive bounded solution $\phi \in C^{2,1}(\mathbb{R}^d \times [0,T))$ to (43). However, (41) requires the test functions to be in $C_b^{2,1}(\mathbb{R}^d \times (0,T])$. To this aim we approximate ϕ in the following way. First, let ϕ_n^T be a sequence in $C_0^\infty(\mathbb{R}^d)$ such that

$$\phi_n^T \to \phi^T \text{ weakly-}* \text{ in } L^\infty(\mathbb{R}^d).$$

Next, let $\phi_n \in C_b^{2,1}(\mathbb{R}^d \times [0,T])$ be the solution of (43) with approximated end condition ϕ_n^T. For this sequence (41) becomes:

$$\begin{aligned}
0 \quad &= \quad \int_\epsilon^T\int \phi_n(y,t) f_{\beta_1}(y,t)dydt + \int \phi_n^T(y)\zeta_{\beta_1}(y,T)dy - \int \phi_n(y,\epsilon)\zeta_{\beta_1}(y,\epsilon)dy \\
&\xrightarrow[\epsilon\to 0]{\text{(i)}} \int_0^T\int \phi_n(y,t) f_{\beta_1}(y,t)dydt + \int \phi_n^T(y)\zeta_{\beta_1}(y,T)dy \\
&\xrightarrow[n\to\infty]{\text{(ii)}} \int_0^T\int \phi(y,t) f_{\beta_1}(y,t)dydt + \int H(\zeta_{\beta_1}(y,T))\zeta_{\beta_1}(y,T)dy,
\end{aligned} \tag{44}$$

using properties (i) and (ii) that we will prove below. From this we infer for the positive part of ζ_{β_1}:

$$0 \le \int \zeta_{\beta_1}^+(y,T)dy \overset{(44)}{=} -\int_0^T\int \underbrace{\phi(y,t)}_{\ge 0}\underbrace{f_{\beta_1}(y,t)}_{\ge 0}dydt \le 0.$$

Analogously we use the other inequality from (42) and conclude that for all $T>0$:

$$\zeta_{\beta_1}(y,T) \le 0 \le \zeta_{\beta_0}(y,T) \text{ for almost every } y\in\mathbb{R}^d,$$

which proves the statement.

Finally, we prove the two limits in (44).

(i) The argument follows from $\zeta_{\beta_1}(x,\epsilon) \to 0$ weakly in $L^1(\mathbb{R}^d)$ as $\epsilon\to 0$. Then for any fixed n:

$$\begin{aligned}
\left|\int (\phi_n(y,\epsilon) - \phi_n(y,0))\,\zeta_{\beta_1}(y,\epsilon)dy\right| &= \left|\int\int_0^\epsilon \partial_t\phi_n(y,t)dt\zeta_{\beta_1}(y,\epsilon)dy\right| \\
&\le \underbrace{\epsilon}_{\to 0}\underbrace{\|\partial_t\phi_n\|_{L^\infty(\mathbb{R}^d\times[0,T])}}_{\text{bounded}}\underbrace{\left|\int \zeta_{\beta_1}(y,\epsilon)dy\right|}_{\to 0} \xrightarrow[\epsilon\to 0]{} 0.
\end{aligned}$$

Hence:

$$\int \phi_n(y,\epsilon)\zeta_{\beta_1}(y,\epsilon)dy = \int (\phi_n(y,\epsilon)-\phi_n(y,0))\,\zeta_{\beta_1}(y,\epsilon)dy + \int \phi_n(y,0)\zeta_{\beta_1}(y,\epsilon) \xrightarrow[\epsilon\to 0]{} 0.$$

(ii) For the second convergence in (44), we can assume that the approximation of the end condition satisfies:

$$0 \le \phi_n^T(y) \le \phi^T(y) \quad \text{for all } y \in \mathbb{R}^d.$$

Therefore:

$$|\phi_n(y,t)f_{\beta_1}(y,t)| \le |\phi(y,t)f_{\beta_1}(y,t)| \le \underbrace{\|\phi^T\|_{L^\infty(\mathbb{R}^d)}|f_{\beta_1}(y,t)|}_{\in L^1(\mathbb{R}^d\times(0,T))}.$$

Since for the fundamental solution η^* of the adjoint problem (43) there holds $z \mapsto \eta^{*t}(y,z) \in L^1(\mathbb{R}^d)$, we have:

$$\phi_n(y,t) = \int \eta^{*t}(y,z)\phi_n^T(z)dz \xrightarrow[n\to\infty]{} \int \eta^{*t}(y,z)\phi^T(z)dz = \phi(y,t)$$

pointwise. The Dominated Convergence theorem then gives

$$\phi_n f_{\beta_1} \xrightarrow[n\to\infty]{L^1} \phi f_{\beta_1}.$$

□

Observe that the factors 1/2 in the exponent of (40) correspond to the factors 1/2 of the energy in expression (5). We are now ready to prove the Mosco-convergence result.

Proof of Theorem 9. To prove the lower bound, take any sequence $\rho^h \rightharpoonup \rho$ in $\mathcal{P}_2^{\mathcal{S}}(\mathbb{R}^d)$ and calculate

$$\begin{aligned}
&\liminf_{h\to 0} \mathcal{J}_{FP}^h(\rho^h\,|\,\overline{\rho}) - \frac{1}{4h}d^2(\overline{\rho},\rho^h)\\
&\overset{(37)}{=} \liminf_{h\to 0} \inf_{q\in\Gamma(\overline{\rho},\rho^h)} \mathcal{H}(q\,|\,\overline{\rho}\eta^h) - \frac{1}{4h}d^2(\overline{\rho},\rho^h)\\
&\overset{(40)}{\ge} \liminf_{h\to 0} \inf_{q\in\Gamma(\overline{\rho},\rho^h)} \mathcal{H}(q\,|\,\overline{\rho}\theta^h)\\
&\qquad - \iint\left(-\frac{1}{2}\Psi(y)+\frac{1}{2}\Psi(x)+\beta_1 h\right)q(dxdy) - \frac{1}{4h}d^2(\overline{\rho},\rho^h)\\
&= \liminf_{h\to 0} \inf_{q\in\Gamma(\overline{\rho},\rho^h)} \mathcal{H}(q\,|\,\overline{\rho}\theta^h) - \frac{1}{4h}d^2(\overline{\rho},\rho^h) + \frac{1}{2}\mathcal{E}(\rho^h) - \frac{1}{2}\mathcal{E}(\overline{\rho}) - \beta_1 h\\
&\ge \frac{1}{2}\mathcal{S}(\rho) - \frac{1}{2}\mathcal{S}(\overline{\rho}) + \frac{1}{2}\mathcal{E}(\rho) - \frac{1}{2}\mathcal{E}(\overline{\rho}),
\end{aligned}$$

where the last inequality follows from Conjecture 3 and the (narrow) continuity of $\rho \mapsto \mathcal{E}(\rho)$.

To construct a recovery sequence, fix a $\rho \in \mathcal{P}_2^{\mathcal{S}}(\mathbb{R}^d)$ and take a recovery sequence $\rho^h \to \rho$ from Conjecture 3, in the strong topology of $\mathcal{P}_2^{\mathcal{S}}(\mathbb{R}^d)$. Then similarly:

$$\begin{aligned}
&\limsup_{h\to 0} \mathcal{J}^h_{FP}(\rho^h \,|\, \overline{\rho}) - \frac{1}{4h} d^2(\overline{\rho}, \rho^h) \\
&\quad \overset{(37)}{=} \limsup_{h\to 0} \inf_{q\in\Gamma(\overline{\rho},\rho^h)} \mathcal{H}(q \,|\, \overline{\rho}\eta^h) - \frac{1}{4h} d^2(\overline{\rho}, \rho^h) \\
&\quad \overset{(40)}{\leq} \limsup_{h\to 0} \inf_{q\in\Gamma(\overline{\rho},\rho^h)} \mathcal{H}(q \,|\, \overline{\rho}\theta^h) - \frac{1}{4h} d^2(\overline{\rho}, \rho^h) + \frac{1}{2}\mathcal{E}(\rho^h) - \frac{1}{2}\mathcal{E}(\overline{\rho}) - \beta_0 h \\
&\quad \leq \frac{1}{2}\mathcal{S}(\rho) - \frac{1}{2}\mathcal{S}(\overline{\rho}) + \frac{1}{2}\mathcal{E}(\rho) - \frac{1}{2}\mathcal{E}(\overline{\rho}). \qquad \square
\end{aligned}$$

4. Diffusion with Drift and Decay

In this section we discuss the case of diffusion with decay. For brevity, we first consider the case without drift ($\Psi \equiv 0$, $\lambda > 0$). First we describe the particle system that we use as a microscopic model for this equation, and calculate the corresponding large-deviation principle. We proceed with the main results for this equation: Mosco convergence to an energy-dissipation functional, and convergence of the approximation scheme to the solution of the diffusion-decay equation. Finally, we discuss how the system can be generalized to include drift, and how the decay can be generalized to diffusion–reaction equations.

4.1. *Microscopic model*

In contrast to the case without decay, the diffusion-decay equation (15) is not mass-conserving, implying that the Wasserstein distance between two time instances of a solution is not defined. To overcome this difficulty, we assume that all decayed matter continues to exist after its decay, but in a different form. We thus distinguish between *normal*, non-decayed matter, denoted by N, and *decayed* or *dark matter*, denoted by D.

The microscopic model now consists of a finite number n of independent non-interacting point particles moving in $\mathbb{R}^d \times \{N, D\}$. Similarly to the non-decaying model, we fix an initial distribution $\overline{\rho} \in \mathcal{P}(\mathbb{R}^d \times \{N, D\})$ and initial positions $x_i \in \mathbb{R}^d$ and states $\mu_i \in \{N, D\}$ such that:

$$\frac{1}{n} \sum_{\substack{i=1\\ \mu_i = N}}^{n} \delta_{x_i} \rightharpoonup \overline{\rho}_N \quad \text{and} \quad \frac{1}{n} \sum_{\substack{i=1\\ \mu_i = D}}^{n} \delta_{x_i} \rightharpoonup \overline{\rho}_D \text{ as } n \to \infty.$$

For the dynamics of the system we assume that the motion of all particles in $\mathbb{R}^d$ is independent of their motion in $\{N, D\}$ (this construction will yield separate terms in the rate functional for both processes). We take the motion in $\mathbb{R}^d$ during some fixed time step $h > 0$ to be Brownian, i.e. governed by the transition probability θ^h from (13). For the motion in $\{N, D\}$, we assume that the time after which a particle changes from N to D is exponentially distributed with rate λ. Since decay is a one-way street, the probability for a particle to change back from D to N is zero. This results in a probability for a particle to change from state μ to ν during the time step h of

$$r^h_{\mu\nu} := \begin{cases} e^{-\lambda h}, & \mu = N, \nu = N, \\ 1 - e^{-\lambda h}, & \mu = N, \nu = D, \\ 0, & \mu = D, \nu = N, \\ 1, & \mu = D, \nu = D. \end{cases}$$

Denote $L^h_n := n^{-1} \sum_{i=1}^n \delta_{(Y^h_i, \nu^h_i)}$, where $Y^h_i \in \mathbb{R}^d$ and $\nu^h_i \in \{N, D\}$ are the random position and state of the ith particle at time h. Indeed, L^h_n converges almost surely to the solution at time h of the system [16, Theorem 11.4.1]

$$\begin{cases} \partial_t u_N = \Delta u_N - \lambda u_N, & \mathbb{R}^d \times (0, \infty), \\ \partial_t u_D = \Delta u_D + \lambda u_N, & \mathbb{R}^d \times (0, \infty), \end{cases} \tag{45}$$

with initial condition $(\overline{\rho}_N, \overline{\rho}_D)$. In this sense, the thus defined particle system is a microscopic interpretation of the diffusion-decay equation (15) (if we ignore the dark matter).

4.2. *Large deviations to gradient flow to PDE*

While the inspiration for this paper was Eq. (1), the construction above suggests to consider not only (1) but also the augmented system of Eq. (45) (and its extensions to non-zero Ψ). For this reason we derive a large-deviation principle and a corresponding energy-dissipation functional for this system, and afterwards simplify by contraction, leading to results for (1).

Let $M^h_n := n^{-1} \sum_{i=1}^n \delta_{(x_i, \mu_i, Y^h_i, \nu^h_i)}$ be the empirical measure of the initial and final configurations corresponding to the particle system defined above. Then (see Theorem A.1) the sequence M^h_n satisfies a large-deviation principle in $\mathcal{P}(\mathbb{R}^d \times \{N, D\} \times \mathbb{R}^d \times \{N, D\})$ with rate n and rate functional

$$\begin{cases} \displaystyle\sum_{\substack{\mu = N,D \\ \nu = N,D}} \mathcal{H}(q_{\mu\nu} \,|\, \overline{\rho}_\mu r^h_{\mu\nu} \theta^h) & \text{if } q(\cdot \times \{N\} \times \mathbb{R}^d \times \{N, D\}) = \overline{\rho}_N(\cdot) \\ & \text{and} \quad q(\cdot \times \{D\} \times \mathbb{R}^d \times \{N, D\}) = \overline{\rho}_D(\cdot), \\ \infty & \text{otherwise,} \end{cases}$$

writing $q_{\mu\nu}(dx\, dy) = q(dx \times \{\mu\} \times dy \times \{\nu\})$. We note that definitions (6) and (21) indeed allow for non-negative Borel measures that are not necessarily probability measures.

In contrast to the previous case without decay, the special structure of the decay forces us to keep track of more information: not only of the total amount of dark matter, but of both the pre-existing dark matter and the normal matter that is converted to dark matter in the present time step, separately. We thus obtain a large-deviation principle for the triple empirical measures $\frac{1}{n} \sum_{i=1}^n \delta_{(\mu_i, Y^h_i, \nu^h_i)}$ with rate n

and rate functional (the subscript stands for "Diffusion equation with Decay")

$$\mathcal{J}^h_{DfDc}(\rho_{NN},\rho_{ND},\rho_{DD}\,|\,\overline{\rho}_N,\overline{\rho}_D)$$
$$:= \inf\Bigg\{\sum_{\mu\nu=\mathrm{NN,ND,DD}}\inf_{q_{\mu\nu}\in\Gamma(\overline{\rho}_{\mu\nu},\rho_{\mu\nu})}\mathcal{H}(q_{\mu\nu}\,|\,\overline{\rho}_\mu r^h_{\mu\nu}\theta^h) : \overline{\rho}_{NN},\overline{\rho}_{ND}\in\mathcal{M}^+(\mathbb{R}^d)$$
$$\text{such that } \overline{\rho}_{NN}+\overline{\rho}_{ND}=\overline{\rho}_N\Bigg\}. \tag{46}$$

Here $\rho_{\mu\nu}$ is the final-time matter of type ν that was initially of type μ, and similarly $\overline{\rho}_{\mu\nu}$ is that part of the initial distribution $\overline{\rho}_\mu$ that will become of type ν at time h (see Fig. 1). Observe that the term $\mathcal{H}(q_{DN}\,|\,0)$ is zero if and only if $q_{DN}\equiv 0$ a.e., and ∞ otherwise; indeed no mass is allowed to change from D to N. Hence we omit the dependency on ρ_{DN}.

Theorem 11 below shows that for small h we have $\mathcal{J}^h_{DfDc}\approx\mathcal{K}^h_{DfDc}$, where

$$\mathcal{K}^h_{DfDc}(\rho_{NN},\rho_{ND},\rho_{DD}\,|\,\overline{\rho}_N,\overline{\rho}_D)$$
$$:= -\frac{1}{2}\mathcal{S}(\rho_{NN}+\rho_{ND})-\frac{1}{2}\mathcal{S}(\overline{\rho}_N)+\frac{1}{4h}d^2(\overline{\rho}_N,\rho_{NN}+\rho_{ND})+\frac{1}{2}\mathcal{S}(\rho_{DD})$$
$$-\frac{1}{2}\mathcal{S}(\overline{\rho}_D)+\frac{1}{4h}d^2(\overline{\rho}_D,\rho_{DD})+\mathcal{S}(\rho_{NN})+\mathcal{S}(\rho_{ND})$$
$$-|\rho_{NN}|\log r^h_{NN}-|\rho_{ND}|\log r^h_{ND}. \tag{47}$$

Let the admissible sets be:

$$B^0 := \{(\overline{\rho}_N,\overline{\rho}_D)\in\mathcal{M}^+(\mathbb{R}^d)^2 : \overline{\rho}_N+\overline{\rho}_D\in\mathcal{P}_2^{\mathcal{S}}(\mathbb{R}^d)\};$$
$$B(\overline{\rho}_N,\overline{\rho}_D) := \Big\{(\rho_{NN},\rho_{ND},\rho_{DD})\in\mathcal{M}^+(\mathbb{R}^d)^3 : \frac{1}{|\overline{\rho}_N|}(\rho_{NN}+\rho_{ND})\in\mathcal{P}_2^{\mathcal{S}}(\mathbb{R}^d)$$
$$\text{and } \frac{1}{|\overline{\rho}_D|}\rho_{DD}\in\mathcal{P}_2^{\mathcal{S}}(\mathbb{R}^d)\Big\},$$

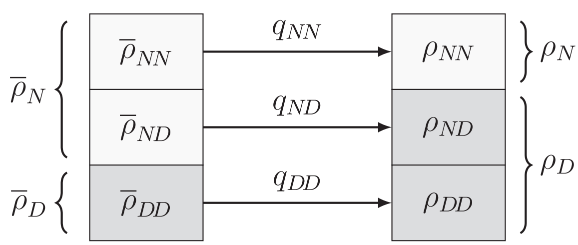

Fig. 1. Notation for the various measures in the diffusion-decay equation. The measures $q_{\mu\nu}$ are pair (coupled) measures, with first and second marginals indicated to the left and right of the arrows. The various marginals $\overline{\rho}_{\mu\nu}$ and $\rho_{\mu\nu}$ combine as indicated to form the observed normal ($\overline{\rho}_N$ and ρ_N) and dark matter ($\overline{\rho}_D$ and ρ_D) at the initial and final times.

equipped with the product of the weak or strong topologies from Sec. 2.3. We remark that $(\rho_{NN}, \rho_{ND}, \rho_{DD}) \in B(\overline{\rho}_N, \overline{\rho}_D)$ implies that $|\overline{\rho}_N| = |\rho_{NN} + \rho_{ND}|$ and $|\overline{\rho}_D| = |\rho_{DD}|$.

Theorem 11. *Assume that Conjecture* 3 *holds. Then for all* $(\overline{\rho}_N, \overline{\rho}_D) \in B^0$

$$\begin{aligned}&\mathcal{J}^h_{DfDc}(\cdot_{NN}, \cdot_{ND}, \cdot_{DD} \,|\, \overline{\rho}_N, \overline{\rho}_D) - \frac{1}{4h} d^2(\overline{\rho}_N\,, \cdot_{NN} + \cdot_{ND}) - \frac{1}{4h} d^2(\overline{\rho}_D\,, \cdot_{DD})\\ &\qquad + |\cdot_{ND}| \log r^h_{ND} + |\cdot_{NN}| \log r^h_{NN}\\ &\xrightarrow[h\to 0]{M} -\frac{1}{2}\mathcal{S}(\cdot_{NN} + \cdot_{ND}) - \frac{1}{2}\mathcal{S}(\overline{\rho}_N) + \frac{1}{2}\mathcal{S}(\cdot_{DD}) - \frac{1}{2}\mathcal{S}(\overline{\rho}_D) + \mathcal{S}(\cdot_{NN}) + \mathcal{S}(\cdot_{ND}),\end{aligned} \tag{48}$$

in $B(\overline{\rho}_N, \overline{\rho}_D)$.

Note that we have not only subtracted three singular terms from $\mathcal{J}^h_{DfDc}$, analogously to Theorem 9, but also the h-order term $-|\cdot_{NN}| \log r^h_{NN}$; the latter is for reasons of symmetry and to simplify calculations.

Finally, we show that the functional $\mathcal{K}^h_{DfDc}$ in (47) indeed defines a variational formulation of the diffusion-decay equation (15). In view of completeness, and of generalizations to diffusion–reaction equations that we will discuss in Sec. 4.5, we prove convergence of the full scheme, including the dark matter, to the system of Eqs. (45). We then derive the corresponding result for the single diffusion-decay equation (15) by minimizing over the dark matter (see Remark 13 below), a procedure essentially the same as the contraction principle (Sec. 2.2). Because we keep track of the dark matter, the matter that decays in a time step should be added to the dark matter already present from the previous iteration.

Theorem 12. *Let* $\rho^0 \in \mathcal{P}^a_2(\mathbb{R}^d)$ *and define the sequence* $\{(\rho^{h,k}_N, \rho^{h,k}_D)\}_{k\geq 0}$ *by*:

$$(\rho^{h,0}_N, \rho^{h,0}_D) := (\rho^0, 0),$$

and for $k \geq 1$:

$$(\rho^{h,k}_{NN}, \rho^{h,k}_{ND}, \rho^{h,k}_{DD}) \in \operatorname*{arg\,min}_{\rho_{NN}+\rho_{ND}+\rho_{DD}\in\mathcal{P}^a_2(\mathbb{R}^d)} \mathcal{K}^h_{DfDc}(\rho_{NN}, \rho_{ND}, \rho_{DD} \,|\, \rho^{h,k-1}_N, \rho^{h,k-1}_D), \tag{49a}$$

$$(\rho^{h,k}_N, \rho^{h,k}_D) := (\rho^{h,k}_{NN}, \rho^{h,k}_{ND} + \rho^{h,k}_{DD}). \tag{49b}$$

These minimizers exist uniquely, and as $h \to 0$ *the pair* $(\rho^{h,\lfloor t/h\rfloor}_N, \rho^{h,\lfloor t/h\rfloor}_D)$ *converges weakly in* $L^1(\mathbb{R}^d \times (0,T)) \times L^1(\mathbb{R}^d \times (0,T))$ *to the solution of* (45) *with initial condition* $(\rho^0, 0)$.

The proof of this theorem is based on [28], and can easily be extended to an additional drift term (see Sec. 4.5). Note that when we let $\lambda \to 0$ then $|\rho_{ND}|$ should vanish in (47) to prevent blow-up; indeed, in that case

$$\mathcal{K}^h_{DfDc}(\rho_{NN}, 0, \rho_{DD} \,|\, \rho^{k-1}_N, \rho^{k-1}_D) = \mathcal{K}^h_{Df}(\rho_{NN} \,|\, \rho^{k-1}_N) + \mathcal{K}^h_{Df}(\rho_{DD} \,|\, \rho^{k-1}_D).$$

Remark 13. A further contraction can be used to ignore the dark matter. We can then ignore the initial dark matter as well, so that the sequence $\frac{1}{n}\sum_{i=1:\nu_i^h=N}^{n}\delta_{Y_i^h}$ satisfies a large-deviation principle with rate n and rate functional

$$\rho_N \mapsto \inf_{\substack{0\le\overline{\rho}_{NN}\le\overline{\rho}_N\\ |\overline{\rho}_{NN}|=|\rho_N|}} \inf_{q\in\Gamma(\overline{\rho}_{NN},\rho_N)} \mathcal{H}(q_{NN}|\overline{\rho}_{NN} r^h_{NN}\theta^h).$$

The corresponding energy-dissipation functional is then:

$$\begin{aligned}\overline{K}^h_{DfDc}(\rho_N\,|\,\overline{\rho}_N) := \inf_{\rho_{ND}:|\rho_N+\rho_{ND}|=|\overline{\rho}_N|} & -\frac{1}{2}\mathcal{S}(\rho_N+\rho_{ND}) - \frac{1}{2}\mathcal{S}(\overline{\rho}_N)\\ &+\frac{1}{4h}d^2(\overline{\rho}_N,\rho_N+\rho_{ND}) + \mathcal{S}(\rho_N)+\mathcal{S}(\rho_{ND})\\ &-|\rho_N|\log r^h_{NN} - |\rho_{ND}|\log r^h_{ND},\end{aligned} \tag{50}$$

which matches the minimization problem (11). The corresponding version of Theorem 12 is the following.

Theorem 14. *Let $\rho^0 \in \mathcal{P}_2^a(\mathbb{R}^d)$ and define the sequence $\{\rho_N^{h,k}\}_{k\ge 0}$ by $\rho_N^{h,0} = \rho^0$ and for $k \ge 1$*

$$\rho_N^{h,k} \in \operatorname*{arg\,min}_{\rho\in\mathcal{M}^+(\mathbb{R}^d)} \overline{K}^h_{DfDc}(\rho\,|\,\rho_N^{h,k-1}).$$

These minimizers exist uniquely, and as $h \to 0$ the function $\rho_N^{h,\lfloor t/h\rfloor}$ converges weakly in $L^1(\mathbb{R}^d\times(0,T))$ to the solution of (15) with initial condition ρ^0.

Remark 15. If we restrict ourselves to measures of mass $|\rho_N| = r^h_{NN}|\overline{\rho}_N|$, thereby excluding the possible fluctuation in the decay process, then (50) further reduces to

$$\rho_N \mapsto \frac{1}{2}\mathcal{S}\left(\frac{1}{r^h_{NN}}\rho_N\right) - \frac{1}{2}\mathcal{S}(\overline{\rho}_N) + \frac{1}{4h}d^2\left(\overline{\rho}_N, \frac{1}{r^h_{NN}}\rho_N\right).$$

A similar scheme to deal with decaying mass can be found in [29].

4.3. *Proof of Theorem* 11

To reduce clutter we abbreviate $\rho_{NT} := \rho_{NN} + \rho_{ND}$ and $q_{NT} := q_{NN} + q_{ND}$. The sum over $\mu\nu = NN, ND$ in $\mathcal{J}^h_{DfDc}$ can be rewritten as:

$$\begin{aligned}&\inf_{\overline{\rho}_{NN}+\overline{\rho}_{ND}=\overline{\rho}_N} \sum_{\nu=N,D} \inf_{q_{N\nu}\in\Gamma(\overline{\rho}_{N\nu},\rho_{N\nu})} \mathcal{H}(q_{N\nu}\,|\,\overline{\rho}_N r^h_{N\nu}\theta^h)\\ &\quad= \inf_{\overline{\rho}_{NN}+\overline{\rho}_{ND}=\overline{\rho}_N} \sum_{\nu} \inf_{q_{N\nu}\in\Gamma(\overline{\rho}_{N\nu},\rho_{N\nu})}\\ &\qquad \times \iint \log\left(\frac{dq_{NT}}{d\overline{\rho}_N\theta^h}\cdot\frac{d\rho_{N\nu}}{d\rho_{NT}}\cdot\frac{1}{r^h_{N\nu}}\cdot\frac{dq_{N\nu}}{\frac{d\rho_{N\nu}}{d\rho_{NT}}dq_{NT}}\right) q_{N\nu}\end{aligned}$$

$$
\begin{aligned}
&= \inf_{q_{NT}\in\Gamma(\overline{\rho}_N,\rho_{NT})} \mathcal{H}(q_{NT}\,|\,\overline{\rho}_N\theta^h) + \mathcal{S}(\rho_{NN}) + \mathcal{S}(\rho_{ND}) - \mathcal{S}(\rho_{NT})\\
&\quad - |\rho_{NN}|\log r^h_{NN} - |\rho_{ND}|\log r^h_{ND}\\
&\quad + \inf_{\overline{\rho}_{NN}+\overline{\rho}_{ND}=\overline{\rho}_N}\ \inf_{\substack{q_{NN}+q_{ND}=q_{NT}\\ q_{NN}\in\Gamma(\overline{\rho}_{NN},\rho_{NN})}} \sum_\nu \mathcal{H}\left(q_{N\nu}\,\middle|\,\frac{d\rho_{N\nu}}{d\rho_{NT}}q_{NT}\right). \qquad (51)
\end{aligned}
$$

We now show that the last sum vanishes under the infima. Since $|q_{N\nu}| = |\rho_{N\nu}| = |\frac{d\rho_{N\nu}}{d\rho_{NT}}q_{NT}|$, we can apply Gibbs' inequality for $\nu = N, D$:

$$
\mathcal{H}\left(q_{N\nu}\,\middle|\,\frac{d\rho_{N\nu}}{d\rho_{NT}}q_{NT}\right) \geq 0.
$$

On the other hand, for any given q_{NT}, the measures

$$
\tilde{q}_{NN} := \frac{d\rho_{NN}}{d\rho_{NT}}q_{NT}, \quad \tilde{q}_{ND} := \frac{d\rho_{ND}}{d\rho_{NT}}q_{NT}
$$

and their first marginals $\overline{\rho}_{NN}(\cdot) = \tilde{q}_{NN}(\cdot\times\mathbb{R}^d)$ and $\overline{\rho}_{ND}(\cdot) = \tilde{q}_{ND}(\cdot\times\mathbb{R}^d)$ are admissible in the infima. It follows that

$$
\begin{aligned}
&\inf_{\overline{\rho}_{NN}+\overline{\rho}_{ND}=\overline{\rho}_N}\ \inf_{\substack{q_{NN}+q_{ND}=q_{NT}\\ q_{NN}\in\Gamma(\overline{\rho}_{NN},\rho_{NN})}} \sum_\nu \mathcal{H}\left(q_{N\nu}\,\middle|\,\frac{d\rho_{N\nu}}{d\rho_{NT}}q_{NT}\right)\\
&\quad \leq \sum_\nu \mathcal{H}\left(\tilde{q}_{N\nu}\,\middle|\,\frac{d\rho_{N\nu}}{d\rho_{NT}}q_{NT}\right) = 0. \qquad (52)
\end{aligned}
$$

Hence we can write:

$$
\begin{aligned}
\mathcal{J}^h_{DfDc}(\rho_{NN},\rho_{ND},\rho_{DD}\,|\,\overline{\rho}_N,\overline{\rho}_D) &= \inf_{q_{NT}\in\Gamma(\overline{\rho}_N,\rho_{NT})} \mathcal{H}(q_{NT}\,|\,\overline{\rho}_N\theta^h) + \mathcal{H}(q_{DD}\,|\,\overline{\rho}_D\theta^h)\\
&\quad + \mathcal{S}(\rho_{NN}) + \mathcal{S}(\rho_{ND}) - \mathcal{S}(\rho_{NT}) - |\rho_{NN}|\log r^h_{NN}\\
&\quad - |\rho_{ND}|\log r^h_{ND}. \qquad (53)
\end{aligned}
$$

Fix a $(\overline{\rho}_N,\overline{\rho}_D)\in B^0$. We first prove the lower bound of the Mosco convergence, and then the existence of a recovery sequence.

Lower Bound. Take any narrowly convergent sequence

$$
(\rho^h_{NN},\rho^h_{ND},\rho^h_{DD}) \rightharpoonup (\rho_{NN},\rho_{ND},\rho_{DD}) \quad \text{in } B(\overline{\rho}_N,\overline{\rho}_D).
$$

Again, we write $\rho^h_{NT} = \rho^h_{NN} + \rho^h_{ND}$. Combining (46), (48), and (53), we need to prove that:

$$
\begin{aligned}
&\liminf_{h\to 0} \inf_{q_{NT}\in\Gamma(\overline{\rho}_N,\rho^h_{NT})} \mathcal{H}(q_{NT}\,|\,\overline{\rho}_N\theta^h) - \frac{1}{4h}d^2(\overline{\rho}_N,\rho^h_{NT})\\
&\qquad + \inf_{q_{DD}\in\Gamma(\overline{\rho}_D,\rho^h_{DD})} \mathcal{H}(q_{DD}\,|\,\overline{\rho}_D\theta^h) - \frac{1}{4h}d^2(\overline{\rho}_D,\rho^h_{DD})\\
&\qquad + \mathcal{S}(\rho^h_{NN}) + \mathcal{S}(\rho^h_{ND}) - \mathcal{S}(\rho^h_{NT})\\
&\quad \geq -\frac{1}{2}\mathcal{S}(\rho_{NT}) - \frac{1}{2}\mathcal{S}(\overline{\rho}_N) + \frac{1}{2}\mathcal{S}(\rho_{DD}) - \frac{1}{2}\mathcal{S}(\overline{\rho}_D) + \mathcal{S}(\rho_{NN}) + \mathcal{S}(\rho_{ND}). \qquad (54)
\end{aligned}
$$

We will prove the lower bound for a number of terms separately.

- By assumption, $|\overline{\rho}_N|^{-1}\rho_{NT}$ lies in $\mathcal{P}_2^{\mathcal{S}}(\mathbb{R}^d)$. If Conjecture 3 is true for probability measures, it also holds for measures of different mass, so that:

$$\liminf_{h\to 0} \inf_{q_{NT}\in\Gamma(\overline{\rho}_N,\rho^h_{NT})} \mathcal{H}(q_{NT}\,|\,\overline{\rho}_N\theta^h) - \frac{1}{4h}d^2(\overline{\rho}_N,\rho^h_{NT}) \geq \frac{1}{2}\mathcal{S}(\rho_{NT}) - \frac{1}{2}\mathcal{S}(\overline{\rho}_N). \tag{55}$$

  Similarly, $|\overline{\rho}_D|^{-1}\rho_{DD} \in \mathcal{P}_2^{\mathcal{S}}(\mathbb{R}^d)$ and so:

$$\liminf_{h\to 0} \inf_{q_{DD}\in\Gamma(\overline{\rho}_D,\rho^h_{DD})} \mathcal{H}(q_{DD}\,|\,\overline{\rho}_D\theta^h) - \frac{1}{4h}d^2(\overline{\rho}_D,\rho^h_{DD}) \geq \frac{1}{2}\mathcal{S}(\rho_{DD}) - \frac{1}{2}\mathcal{S}(\overline{\rho}_D). \tag{56}$$

- Since the function $(x,y)\mapsto x\log x + y\log y - (x+y)\log(x+y)$ is convex, the functional

$$F : (\rho_{NN},\rho_{ND}) \mapsto \mathcal{S}(\rho_{NN}) + \mathcal{S}(\rho_{ND}) - \mathcal{S}(\rho_{NN}+\rho_{ND})$$

  is also convex, and lower semicontinuous in $B(\overline{\rho}_N,\overline{\rho}_D)$ with the narrow topology [25, Theorem 4.3]

$$\liminf_{h\to 0} \mathcal{S}(\rho^h_{NN}) + \mathcal{S}(\rho^h_{ND}) - \mathcal{S}(\rho^h_{NT}) \geq \mathcal{S}(\rho_{NN}) + \mathcal{S}(\rho_{ND}) - \mathcal{S}(\rho_{NT}). \tag{57}$$

The required lower bound (54) then follows from (55), (56) and (57).

Recovery Sequence. Fix $(\rho_{NN},\rho_{ND},\rho_{DD}) \in B(\overline{\rho}_N,\overline{\rho}_D)$ and take two recovery sequences $\rho^h_{DD}\to\rho_{DD}$ and $\rho^h_{NT}\to\rho_{NN}+\rho_{ND}$ in the strong topology from Conjecture 3 such that

$$\limsup_{h\to 0} \inf_{q_{DD}\in\Gamma(\overline{\rho}_D,\rho^h_{DD})} \mathcal{H}(q_{DD}\,|\,\overline{\rho}_D\theta^h) - \frac{1}{4h}d^2(\overline{\rho}_D,\rho^h_{DD}) = \frac{1}{2}\mathcal{S}(\rho_{DD}) - \frac{1}{2}\mathcal{S}(\overline{\rho}_D), \tag{58}$$

$$\limsup_{h\to 0} \inf_{q_{NT}\in\Gamma(\overline{\rho}_N,\rho^h_{NT})} \mathcal{H}(q_{NT}\,|\,\overline{\rho}_N\theta^h) - \frac{1}{4h}d^2(\overline{\rho}_N,\rho^h_{NT}) = \frac{1}{2}\mathcal{S}(\rho_{NN}+\rho_{ND}) - \frac{1}{2}\mathcal{S}(\overline{\rho}_N). \tag{59}$$

In contrast to the case of the lower bound we define ρ^h_{NN} and ρ^h_{ND} in terms of ρ^h_{NT}:

$$\rho^h_{NN} := \frac{d\rho_{NN}}{d(\rho_{NN}+\rho_{ND})}\rho^h_{NT}, \quad \rho^h_{ND} := \frac{d\rho_{ND}}{d(\rho_{NN}+\rho_{ND})}\rho^h_{NT}.$$

Here we define the Radon–Nikodym derivatives to be 1 on null sets of $\rho_{NN}+\rho_{ND}$. Observe that by definition of the strong topology $\mathcal{S}(\rho^h_{NT}) \to S(\rho_{NN}+\rho_{ND})$. By

Lemma 7, this implies that

$$\rho^h_{NT} \to \rho_{NN} + \rho_{ND} \quad \text{and} \quad \rho^h_{NT} \log \rho^h_{NT} \to (\rho_{NN} + \rho_{ND}) \log(\rho_{NN} + \rho_{ND})$$

strongly in $L^1(\mathbb{R}^d)$, if we redefine the sequence by its convergent subsequence. Therefore, with $0 \le \alpha(x) := \frac{d\rho_{NN}}{d(\rho_{NN}+\rho_{ND})}(x) \le 1$

$$\begin{aligned}
|\mathcal{S}(\rho^h_{NN}) - \mathcal{S}(\rho_{NN})| &= \left| \int \alpha\rho^h_{NT} \log \alpha\rho^h_{NT} - \int \alpha \cdot (\rho_{NN} + \rho_{ND}) \log \alpha \cdot (\rho_{NN} + \rho_{ND}) \right| \\
&\le \left| \int \alpha\rho^h_{NT} \log \rho^h_{NT} - \int \alpha(\rho_{NN} + \rho_{ND}) \log(\rho_{NN} + \rho_{ND}) \right| \\
&\quad + \left| \int \rho^h_{NT} \alpha \log \alpha - \int (\rho_{NN} + \rho_{ND}) \alpha \log \alpha \right| \\
&\le \int \left| \rho^h_{NT} \log \rho^h_{NT} - (\rho_{NN} + \rho_{ND}) \log(\rho_{NN} + \rho_{ND}) \right| \\
&\quad + \frac{1}{e} \int \left| \rho^h_{NT} - (\rho_{NN} - \rho_{ND}) \right| \\
&\to 0,
\end{aligned}$$

and analogously for ρ^h_{ND}. Collecting the convergence results:

$$\mathcal{S}(\rho^h_{NN}) \to \mathcal{S}(\rho_{NN}), \quad \mathcal{S}(\rho^h_{ND}) \to \mathcal{S}(\rho_{ND}) \quad \text{and} \quad \mathcal{S}(\rho^h_{NT}) \to \mathcal{S}(\rho_{NN} + \rho_{ND}). \quad (60)$$

Then it follows from (58), (59), and (60) that $(\rho^h_{NN}, \rho^h_{ND}, \rho^h_{DD})$ is a recovery sequence, i.e.

$$\begin{aligned}
&\limsup_{h\to 0} \inf_{q_{NT} \in \Gamma(\overline{\rho}_N, \rho^h_{NT})} \mathcal{H}(q_{NT} \,|\, \overline{\rho}_N \theta^h) - \frac{1}{4h} d^2(\overline{\rho}_N, \rho^h_{NT}) \\
&\qquad + \inf_{q_{DD} \in \Gamma(\overline{\rho}_D, \rho^h_{DD})} \mathcal{H}(q_{DD} \,|\, \overline{\rho}_D \theta^h) - \frac{1}{4h} d^2(\overline{\rho}_D, \rho^h_{DD}) \\
&\qquad + \mathcal{S}(\rho^h_{NN}) + \mathcal{S}(\rho^h_{ND}) - \mathcal{S}(\rho^h_{NT}) \\
&\quad \le -\frac{1}{2}\mathcal{S}(\rho_{NN} + \rho_{ND}) - \frac{1}{2}\mathcal{S}(\overline{\rho}_N) + \frac{1}{2}\mathcal{S}(\rho_{DD}) - \frac{1}{2}\mathcal{S}(\overline{\rho}_D) + \mathcal{S}(\rho_{NN}) + \mathcal{S}(\rho_{ND}).
\end{aligned}$$

This concludes the proof of Theorem 11. □

4.4. *Proof of Theorem* 12

Theorem 12 contains two main results: existence and uniqueness of minimizers, and the convergence of time-discrete solutions. We first discuss the existence and uniqueness of minimizers. By slightly rewriting (49) we can minimize, for fixed $(\rho^{h,k-1}_N, \rho^{h,k-1}_D) \in \mathcal{P}^a_2(\mathbb{R}^d)$, the functional

$$\begin{aligned}
(\rho_{NN}, \rho_{NT}, \rho_{DD}) &\mapsto \mathcal{K}^h_{DfDc}(\rho_{NN}, \rho_{NT} - \rho_{NN}, \rho_{DD} \,|\, \rho^{h,k-1}_N, \rho^{h,k-1}_D) \\
&= -\frac{1}{2}\mathcal{S}(\rho_{NT}) - \frac{1}{2}\mathcal{S}(\rho^{h,k-1}_N) + \frac{1}{4h} d^2(\rho^{h,k-1}_N, \rho_{NT})
\end{aligned}$$

$$+\frac{1}{2}\mathcal{S}(\rho_{DD}) - \frac{1}{2}\mathcal{S}(\rho_D^{h,k-1}) + \frac{1}{4h}d^2(\rho_D^{h,k-1}, \rho_{DD})$$
$$+\mathcal{S}(\rho_{NN}) + \mathcal{S}(\rho_{NT} - \rho_{NN}) - |\rho_{NN}|\log r_{NN}^h$$
$$-|\rho_{NT} - \rho_{NN}|\log r_{ND}^h. \tag{61}$$

The negative sign of the term $-\frac{1}{2}\mathcal{S}(\rho_{NT})$ makes this minimization problem slightly non-trivial. We therefore proceed in steps. For fixed ρ_{NT}, the functional

$$F^h(\rho_{NN}) := \mathcal{S}(\rho_{NN}) + \mathcal{S}(\rho_{NT} - \rho_{NN}) - |\rho_{NN}|\log r_{NN}^h - |\rho_{NT} - \rho_{NN}|\log r_{ND}^h$$

is convex and has a unique stationary point that satisfies

$$0 = \log\rho_{NN} - \log(\rho_{NT} - \rho_{NN}) - \log r_{NN}^h + \log r_{ND}^h,$$

implying that $\rho_{NN} := r_{NN}^h \rho_{NT}$ is the unique global minimizer of F. Therefore, at every step k, we have (see Fig. 1)

$$\rho_N^{h,k} = \rho_{NN}^{h,k} = r_{NN}^h \rho_{NT}^{h,k} \quad \text{and} \quad \rho_{ND}^{h,k} = r_{ND}^h \rho_{NT}^{h,k}. \tag{62}$$

The problem of minimizing (61) can now be reduced to the minimization of

$$(\rho_{NT}, \rho_{DD}) \mapsto \mathcal{K}_{DfDc}^h(r_{NN}^h\rho_{NT}, r_{ND}^h\rho_{NT}, \rho_{DD} \mid \rho_N^{h,k-1}, \rho_D^{h,k-1})$$
$$= \frac{1}{2}\mathcal{S}(\rho_{NT}) - \frac{1}{2}\mathcal{S}(\rho_N^{h,k-1}) + \frac{1}{4h}d^2(\rho_N^{h,k-1}, \rho_{NT})$$
$$+\frac{1}{2}\mathcal{S}(\rho_{DD}) - \frac{1}{2}\mathcal{S}(\rho_D^{h,k-1}) + \frac{1}{4h}d^2(\rho_D^{h,k-1}, \rho_{DD}), \tag{63}$$

which consists of two decoupled minimization problems, for which existence and uniqueness of minimizers are proved in [28, Proposition 4.1].

The compactness of the sequence $(\rho_N^{h,\lfloor t/h\rfloor}, \rho_D^{h,\lfloor t/h\rfloor})$ is based on the same principle as in [28], but with a twist. The central observation is again that $(\rho_N^{h,k-1}, \rho_D^{h,k-1})$ is admissible in (63), leading to the estimate

$$\frac{1}{2h}d^2(\rho_N^{h,k-1}, \rho_{NT}^{h,k}) + \frac{1}{2h}d^2(\rho_D^{h,k-1}, \rho_{DD}^{h,k})$$
$$\le -\mathcal{S}(\rho_{NT}^{h,k}) + \mathcal{S}(\rho_N^{h,k-1}) - \mathcal{S}(\rho_{DD}^{h,k}) + \mathcal{S}(\rho_D^{h,k-1}). \tag{64}$$

However, the migration of mass from normal to dark matter means that upon summing this estimate over k, terms in the right-hand side do not cancel. Below we establish the *a priori* estimates

$$M_2(\rho_N^{h,k} + \rho_D^{h,k}) := \int |x|^2 d(\rho_N^{h,k} + \rho_D^{h,k}) \le C, \tag{65}$$

$$\sum_{k=1}^{\lfloor T/h\rfloor} d^2(\rho_N^{h,k-1}, \rho_{NT}^{h,k}) + d^2(\rho_D^{h,k-1}, \rho_{DD}^{h,k}) \le Ch, \tag{66}$$

where the constant C only depends on the initial data and on the maximal time T. As in [28] these estimates provide the appropriate tightness in space (by (65)) and continuity in time (by (66)) to conclude that there exists a subsequence such that $(\rho_N^{h,\lfloor t/h\rfloor},\rho_D^{h,\lfloor t/h\rfloor}) \to (u_N,u_D)$, weakly in $L^1(\mathbb{R}^d\times(0,T))\times L^1(\mathbb{R}^d\times(0,T))$.

We now prove (65) and (66). Recall from [28] the estimates

$$\begin{aligned} -\mathcal{S}(\rho) &\le C(M_2(\rho)+1)^\alpha && \text{for some } 0<\alpha<1 \text{ and for all } \rho\in\mathcal{M}^+(\mathbb{R}^d),\\ M_2(\rho_1) &\le 2M_2(\rho_0)+2d^2(\rho_0,\rho_1) && \text{for all } \rho_0,\rho_1\in\mathcal{M}^+(\mathbb{R}^d) \text{ with } |\rho_0|=|\rho_1|. \end{aligned} \tag{67}$$

This allows us to estimate, for $n\in\mathbb{N}$ such that $nh\le T$,

$$M_2(\rho_N^{h,n}+\rho_D^{h,n}) \le 2M_2(\rho_N^0+\rho_D^0)+2d^2(\rho_N^{h,n}+\rho_D^{h,n},\rho_N^0+\rho_D^0). \tag{68}$$

The second term above we then estimate by

$$\begin{aligned} d^2(\rho_N^{h,n}+\rho_D^{h,n},\rho_N^0+\rho_D^0) &\le \left[\sum_{k=1}^n d(\rho_N^{h,k}+\rho_D^{h,k},\rho_N^{h,k-1}+\rho_D^{h,k-1})\right]^2\\ &\le n\sum_{k=1}^n d^2(\rho_N^{h,k}+\rho_D^{h,k},\rho_N^{h,k-1}+\rho_D^{h,k-1})\\ &= n\sum_{k=1}^n d^2(\rho_{NT}^{h,k}+\rho_{DD}^{h,k},\rho_N^{h,k-1}+\rho_D^{h,k-1})\\ &\overset{(16)}{\le} n\sum_{k=1}^n d^2(\rho_{NT}^{h,k},\rho_N^{h,k-1})+d^2(\rho_{DD}^{h,k},\rho_D^{h,k-1}). \end{aligned} \tag{69}$$

We also observe some properties of $\mathcal{S}$:

$$\begin{aligned} \mathcal{S}(\alpha\rho+\beta\rho) = \mathcal{S}(\alpha\rho)+\mathcal{S}(\beta\rho)-\alpha|\rho|\log\frac{\alpha}{\alpha+\beta}&\\ -\beta|\rho|\frac{\beta}{\alpha+\beta},\quad \text{for all } \alpha,\beta>0 \text{ and } \rho\in\mathcal{M}^+(\mathbb{R}^d),& \end{aligned}$$

and in general

$$\begin{aligned} \mathcal{S}(\rho_1+\rho_2) \le \mathcal{S}(\rho_1)+\mathcal{S}(\rho_2)-|\rho_1|\log\frac{|\rho_1|}{|\rho_1+\rho_2|}&\\ -|\rho_2|\log\frac{|\rho_2|}{|\rho_1+\rho_2|}\quad \text{for any } \rho_1,\rho_2\in\mathcal{M}^+(\mathbb{R}^d).& \end{aligned}$$

The first follows from a simple calculation, and the second can be proved by writing $\rho_1+\rho_2=\lambda(\rho_1/\lambda)+(1-\lambda)(\rho_2/(1-\lambda))$, applying the convexity of $\mathcal{S}$, and optimizing with respect to λ. Combining these with (62) we then have

$$\mathcal{S}(\rho_{NT}^{h,k}) = \mathcal{S}(\rho_{NN}^{h,k})+\mathcal{S}(\rho_{ND}^{h,k})-|\rho_{NN}^{h,k}|\log r_{NN}^h-|\rho_{ND}^{h,k}|\log r_{ND}^h, \tag{70}$$

$$\mathcal{S}(\rho_D^{h,k}) \le \mathcal{S}(\rho_{ND}^{h,k})+\mathcal{S}(\rho_{DD}^{h,k})-|\rho_{ND}^{h,k}|\log\frac{|\rho_{ND}^{h,k}|}{|\rho_D^{h,k}|}-|\rho_{DD}^{h,k}|\log\frac{|\rho_{DD}^{h,k}|}{|\rho_D^{h,k}|}. \tag{71}$$

Now, putting the ingredients together:

$$\begin{aligned}
M_2(\rho_N^{h,n}+\rho_D^{h,n}) &\overset{(68)}{\le} 2M_2(\rho_N^0+\rho_D^0)+2d^2(\rho_N^{h,n}+\rho_D^{h,n},\rho_N^0+\rho_D^0)\\
&\overset{(69)}{\le} C+2n\sum_{k=1}^n d^2(\rho_N^{h,k-1},\rho_{NT}^{h,k})+d^2(\rho_D^{h,k-1},\rho_{DD}^{h,k})\\
&\overset{(64)}{\le} C+4nh\sum_{k=1}^n \mathcal{S}(\rho_N^{h,k-1})-\mathcal{S}(\rho_{NT}^{h,k})+\mathcal{S}(\rho_D^{h,k-1})-\mathcal{S}(\rho_{DD}^{h,k})\\
&\overset{(70),(71)}{\le} C+4T\sum_{k=1}^n \mathcal{S}(\rho_N^{h,k-1})-\mathcal{S}(\rho_N^{h,k})+\mathcal{S}(\rho_D^{h,k-1})-\mathcal{S}(\rho_D^{h,k})\\
&\qquad \underbrace{\begin{aligned}&+4T\sum_{k=1}^n |\rho_{NN}^{h,k}|\log r_{NN}^h+|\rho_{ND}^{h,k}|\log r_{ND}^h-|\rho_{ND}^{h,k}|\log\frac{|\rho_{ND}^{h,k}|}{|\rho_D^{h,k}|}\\ &\quad -|\rho_{DD}^{h,k}|\log\frac{|\rho_{DD}^{h,k}|}{|\rho_D^{h,k}|}\end{aligned}}_{\le 0 \text{ (see below)}}\\
&\overset{(67)}{\le} C+4T[\mathcal{S}(\rho_N^0)+\mathcal{S}(\rho_D^0)+C(M_2(\rho_N^{h,n})+1)^\alpha\\
&\qquad +C(M_2(\rho_D^{h,n})+1)^\alpha]\\
&\le C+4T[\mathcal{S}(\rho_N^0)+\mathcal{S}(\rho_D^0)+2^\alpha C(M_2(\rho_N^{h,n}+\rho_D^{h,n})+2)^\alpha].
\end{aligned}$$

Therefore $M_2(\rho_N^{h,n}+\rho_D^{h,n})$ is bounded on finite time intervals, which proves (65), and the boundedness of the second line above implies (66).

The sign of the brace above can be shown as follows: setting $r := r_{NN}^h$ and therefore by (62), we have

$$|\rho_N^{h,k}|=r^k,\quad |\rho_D^{h,k}|=1-r^k,\quad |\rho_{ND}^{h,k}|=r^k-r^{k-1},\quad \text{and}\quad |\rho_{DD}^{h,k}|=1-r^{k-1}.$$

Then

$$\begin{aligned}
&\sum_{k=1}^n |\rho_{NN}^{h,k}|\log r_{NN}^h+|\rho_{ND}^{h,k}|\log r_{ND}^h-|\rho_{ND}^{h,k}|\log\frac{|\rho_{ND}^{h,k}|}{|\rho_D^{h,k}|}-|\rho_{DD}^{h,k}|\log\frac{|\rho_{DD}^{h,k}|}{|\rho_D^{h,k}|}\\
&\quad=\sum_{k=1}^n r^k\log r+(r^{k-1}-r^k)\log(1-r)-(r^{k-1}-r^k)\log\frac{r^{k-1}-r^k}{1-r^k}\\
&\qquad-(1-r^{k-1})\log\frac{1-r^{k-1}}{1-r^k}\\
&\quad=\sum_{k=1}^n r^k\log r^k-r^{k-1}\log r^{k-1}+(1-r^k)\log(1-r^k)\\
&\qquad-(1-r^{k-1})\log(1-r^{k-1})\\
&\quad=r^n\log r^n+(1-r^n)\log(1-r^n)\le 0.
\end{aligned}$$

This concludes the proof of the compactness and therefore the convergence of a subsequence.

We now determine the equation satisfied by the time-discrete minimizers using the method introduced in [28]. After perturbing the minimizers $\rho_{NT}^{h,k}$ and $\rho_{DD}^{h,k}$ by a push-forward, we find that for all $\xi \in C_0^\infty(\mathbb{R}^d;\mathbb{R}^d)$,

$$\begin{aligned}\iint (y-x)\cdot\xi(y)q_{NT}(dx\,dy) - h\int \operatorname{div}\xi(y)\rho_{NT}^{h,k}(y)dy = 0,\\ \iint (y-x)\cdot\xi(y)q_{DD}(dx\,dy) - h\int \operatorname{div}\xi(y)\rho_{DD}^{h,k}(y)dy = 0,\end{aligned} \tag{72}$$

where q_{NT} and q_{DD} are the optimal transport plans in $d(\rho_N^{h,k-1},\rho_{NT}^{h,k})$ and $d(\rho_D^{h,k-1},\rho_{DD}^{h,k})$. Using $\rho_N^{h,k} = \rho_{NN}^{h,k} = r_{NN}^h\rho_{NT}^{h,k}$ and $\rho_D^{h,k} = r_{ND}^h\rho_{NT}^{h,k} + \rho_{DD}^{h,k}$ as prescribed by (49b) and (62), we add up the equations above to find for all ξ,

$$\begin{aligned}\iint (y-x)\cdot\xi(y)r_{NN}^h q_{NT}(dx\,dy) - h\int \operatorname{div}\xi(y)\rho_N^{h,k}(y)dy = 0,\\ \iint (y-x)\cdot\xi(y)(r_{ND}^h q_{NT} + q_{DD})(dx\,dy) - h\int \operatorname{div}\xi(y)\rho_D^{h,k}(y)dy = 0.\end{aligned} \tag{73}$$

As $r_{NN}^h q_{NT} \in \Gamma(r_{NN}^h\rho_N^{h,k-1},\rho_N^{h,k})$ and $r_{ND}^h q_{NT} + q_{DD} \in \Gamma(r_{ND}^h\rho_N^{h,k-1} + \rho_D^{h,k-1},\rho_D^{h,k})$ (although the second may not be optimal), we have the following bounds for any $\zeta \in C_0^\infty(\mathbb{R}^d)$:

$$\begin{aligned}&\left|\int(\rho_N^{h,k} - r_{NN}^h\rho_N^{h,k-1})\zeta - \iint (y-x)\cdot\nabla\zeta(y)r_{NN}^h q_{NT}(dx\,dy)\right|\\ &\quad= \left|\iint (\zeta(y)-\zeta(x)+(x-y)\cdot\nabla\zeta(y))r_{NN}^h q_{NT}(dx\,dy)\right|\\ &\quad\le \frac{1}{2}\sup|\Delta\zeta| r_{NN}^h \iint |y-x|^2 q_{NT}(dx\,dy)\\ &\quad= \frac{1}{2}\sup|\Delta\zeta| d^2(\rho_N^{h,k-1},\rho_{NT}^{h,k}),\end{aligned}$$

and similarly,

$$\begin{aligned}&\left|\int(\rho_D^{h,k} - (r_{ND}^h\rho_N^{h,k-1} + \rho_D^{h,k-1}))\zeta - \iint (y-x)\cdot\nabla\zeta(y)(r_{ND}^h q_{NT} + q_{DD})(dx\,dy)\right|\\ &\quad\le \frac{1}{2}\sup|\Delta\zeta|(d^2(\rho_N^{h,k-1},\rho_{NT}^{h,k}) + d^2(\rho_D^{h,k-1},\rho_{DD}^{h,k})).\end{aligned}$$

After applying these bounds to (73), taking $\xi = \nabla\zeta$, we find for all ζ:

$$\left|\int\left(\frac{1}{h}(\rho_N^{h,k} - r_{NN}^h\rho_N^{h,k-1})\zeta - \rho_N^{h,k}\Delta\zeta\right)dy\right| \le \frac{1}{2h}\sup|\Delta\zeta| d^2(\rho_N^{h,k-1},\rho_{NT}^{h,k}),$$

and

$$\left| \int \left(\frac{1}{h} (\rho_D^{h,k} - r_{ND}^h \rho_N^{h,k-1} - \rho_D^{h,k-1}) \zeta - \rho_D^{h,k} \Delta\zeta \right) dy \right|$$
$$\leq \frac{1}{2h} \sup |\Delta\zeta| (d^2(\rho_N^{h,k-1}, \rho_{NT}^{h,k}) + d^2(\rho_D^{h,k-1}, \rho_{DD}^{h,k})).$$

Using the convergence of a subsequence (not relabeled) $(\rho_N^{h,\lfloor t/h \rfloor}, \rho_D^{h,\lfloor t/h \rfloor}) \to (u_N, u_D)$ weakly in $L^1(\mathbb{R}^d \times (0,T)) \times L^1(\mathbb{R}^d \times (0,T))$, we find that for all $\zeta \in C_0^\infty(\mathbb{R}^d \times [0,T])$,

$$\left| \int_0^T \int u_N \left(-\partial_t \zeta + \left(\lim_{h \to 0} \frac{1 - r_{NN}^h}{h} \right) \zeta - \Delta\zeta \right) dy\, dt \right|$$
$$\xleftarrow{h \to 0} \left| \int_0^T \int \left(\frac{1}{h} (\rho_N^{h,\lfloor t/h \rfloor} - \rho_N^{h,\lfloor t/h \rfloor - 1}) \zeta + \frac{1 - r_{NN}^h}{h} \rho_N^{h,\lfloor t/h \rfloor - 1} \zeta \right.\right.$$
$$\left.\left. - \rho_N^{h,\lfloor t/h \rfloor} \Delta\zeta \right) dx\, dt \right|$$
$$\leq \sum_{k=1}^{\lfloor T/h \rfloor} \frac{1}{2} \sup \left| \Delta \int_0^T \zeta \right| d^2(\rho_N^{h,k-1}, \rho_{NT}^{h,k})$$
$$\overset{(66)}{\leq} Ch \xrightarrow{h \to 0} 0,$$

and for the dark matter:

$$\left| \int_0^T \int \left(-u_D \partial_t \zeta - \left(\lim_{h \to 0} \frac{r_{ND}^h}{h} \right) u_N \zeta - u_D \Delta\zeta \right) dy\, dt \right|$$
$$\xleftarrow{h \to 0} \left| \int_0^T \int \left(\frac{1}{h} (\rho_D^{h,\lfloor t/h \rfloor} - \rho_D^{h,\lfloor t/h \rfloor - 1}) \zeta - \frac{r_{ND}^h}{h} \rho_N^{h,\lfloor t/h \rfloor - 1} \zeta \right.\right.$$
$$\left.\left. - \rho_D^{h,\lfloor t/h \rfloor} \Delta\zeta \right) dx\, dt \right|$$
$$\leq \sum_{k=1}^{\lfloor T/h \rfloor} \frac{1}{2} \sup \left| \Delta \int_0^T \zeta \right| (d^2(\rho_N^{h,k-1}, \rho_{NT}^{h,k}) + d^2(\rho_D^{h,k-1}, \rho_{DD}^{h,k}))$$
$$\overset{(66)}{\leq} Ch \xrightarrow{h \to 0} 0.$$

From this we see that the limit (u_N, u_D) indeed solves (45) (weakly in $L^1(\mathbb{R}^d \times (0,T))$). This concludes the proof of Theorem 12.

4.5. *Drift with decay and reactions*

Diffusion with drift and decay. The results from Secs. 3 and 4 can be easily combined in the following way. A microscopic model for the Fokker–Planck equation with decay (1) is obtained by replacing the spatial transition probability θ^h in the micro model from Sec. 4.1 by the fundamental solution η^h of the Fokker–Planck equation from Definition 8. The corresponding large-deviation rate functional then simply becomes (46) with that transition probability. By the same arguments of Theorems 9 and 11, the large-deviation rate functional is related to the following energy-dissipation functional in a Mosco-convergence sense:

$$\begin{aligned}\mathcal{K}^h_{FPDc}&(\rho_{NN},\rho_{ND},\rho_{DD}\,|\,\overline{\rho}_N,\overline{\rho}_D)\\ &:= -\frac{1}{2}\mathcal{S}(\rho_{NN}+\rho_{ND}) - \frac{1}{2}\mathcal{S}(\overline{\rho}_N) + \frac{1}{4h}d^2(\overline{\rho}_N,\rho_{NN}+\rho_{ND}) + \frac{1}{2}\mathcal{S}(\rho_{DD})\\ &\quad - \frac{1}{2}\mathcal{S}(\overline{\rho}_D) + \frac{1}{4h}d^2(\overline{\rho}_D,\rho_{DD}) + \mathcal{S}(\rho_{NN}) + \mathcal{S}(\rho_{ND}) - |\rho_{NN}|\log r^h_{NN}\\ &\quad - |\rho_{ND}|\log r^h_{ND} + \frac{1}{2}\mathcal{E}(\rho_{NN}+\rho_{ND}+\rho_{DD}) - \frac{1}{2}\mathcal{E}(\overline{\rho}_N+\overline{\rho}_D). \end{aligned} \tag{74}$$

Indeed, as our main result this functional defines a variational formulation for the Fokker–Planck equation with decay (1).

Theorem 16. *Let $\rho^0 \in \mathcal{P}^a_2(\mathbb{R}^d)$ and define the sequence $\{(\rho_N^{h,k},\rho_D^{h,k})\}_{k\geq 0}$ by:*

$$(\rho_N^{h,0},\rho_D^{h,0}) = (\rho^0,0),$$

and for $k \geq 1$:

$$(\rho_{NN}^{h,k},\rho_{ND}^{h,k},\rho_{DD}^{h,k}) \in \operatorname*{arg\,min}_{\rho_{NN}+\rho_{ND}+\rho_{DD}\in\mathcal{P}^a_2(\mathbb{R}^d)} \mathcal{K}^h_{FPDc}(\rho_{NN},\rho_{ND},\rho_{DD}\,|\,\rho_N^{h,k-1},\rho_D^{h,k-1}),$$

$$(\rho_N^{h,k},\rho_D^{h,k}) = (\rho_{NN}^{h,k},\rho_{ND}^{h,k}+\rho_{DD}^{h,k}).$$

These minimizers exist uniquely, and as $h \to 0$ the pair $(\rho_N^{h,\lfloor t/h\rfloor},\rho_D^{h,\lfloor t/h\rfloor})$ converges weakly in $L^1(\mathbb{R}^d\times(0,T))$ to the solution of (45) *with initial condition $(\rho^0,0)$.*

The proof is a slight adaptation of the proof of Theorem 12, with the observation that after perturbing with a push-forward, the continuity equations (72) include the additional terms $h\int\xi(y)\cdot\nabla\Psi(y)\,\rho_{NT}^{h,k}(y)\,dy$ and $h\int\xi(y)\cdot\nabla\Psi(y)\,\rho_{DD}^{h,k}(y)\,dy$ for the potential energy. Following the proof of Theorem 12, these extra terms will result in the convection term in Eq. (1).

Diffusion–reaction equations. Another useful generalization is a system of equations that describe the transition between a set of states ν in some index set I:

$$\partial_t u_\nu = \Delta u_\nu - \sum_{\mu\neq\nu} s_{\mu\nu}u_\nu + \sum_{\mu\neq\nu} s_{\nu\mu}u_\mu, \quad \nu\in I. \tag{75}$$

We should then choose the transition probabilities $r^h_{\mu\nu}$ of the microscopic system in such a way that $\lim_{h\to 0}\frac{r^h_{\mu\nu}}{h} = s_{\mu\nu}$ and $r^h_{\mu\mu} = 1-\sum_{\nu\neq\mu} r^h_{\mu\nu}$. The large-deviation rate functional corresponding to this micro model is:

$$(\{\rho_{\mu\nu}\}_{\mu,\nu\in I};\{\overline{\rho}_\mu\}_{\mu\in I}) \mapsto \sum_{\mu\in I} \inf_{\substack{\overline{\rho}_{\mu\nu}:\mu,\nu\in I\\ \sum_{\nu\in I}\overline{\rho}_{\mu\nu}=\overline{\rho}_\mu}} \sum_{\nu\in I} \inf_{q_{\mu\nu}\in\Gamma(\overline{\rho}_{\mu\nu},\rho_{\mu\nu})} \mathcal{H}(q_{\mu\nu}\,|\,\overline{\rho}_\mu r^h_{\mu\nu}\theta^h),$$

which Mosco-converges, after subtracting singular terms, to the functional:

$$\sum_{\mu\in I}\Bigg[-\frac{1}{2}S\left(\sum_{\nu\in I}\rho_{\mu\nu}\right) - \frac{1}{2}\mathcal{S}(\overline{\rho}_\mu) + \frac{1}{4h}d^2\left(\overline{\rho}_\mu,\sum_{\nu\in I}\rho_{\mu\nu}\right) + \sum_{\nu\in I}(\mathcal{S}(\rho_{\mu\nu}) - |\rho_{\mu\nu}|\log r^h_{\mu\nu})\Bigg]. \tag{76}$$

In the same way as in Theorem 12, this functional defines a variational formulation for the system of diffusion–reaction equations (75).

5. Discussion

The work of [1] uncovered an intriguing link between the diffusion equation, the entropy-Wasserstein gradient-flow formulation of that equation, and a large-deviation principle for a stochastic particle system. The work of the present paper is motivated by the question whether this link can be generalized.

Equation (1) moves beyond [1] in two ways. The additional drift term represented by Ψ is compatible with the Wasserstein framework. The corresponding equation (7) is a Wasserstein gradient flow of the free energy functional $\mathcal{S}+\mathcal{E}$. In Sec. 3 we showed that also the large-deviation connection generalizes to this case, with only minor modification. Corresponding continuous-time large-deviations results for instance in [11] or [21, Theorem 13.37] mirror this.

The case of decay is different. The structure of the time-discretized gradient flow in Theorem 12 has some non-standard features:

- The iteration defined in Theorem 12 is special in that the minimization is taken over the pair (ρ_{NN},ρ_{ND}), and the result is *added* to the dark matter of the previous time step. Of course, when ignoring the dark matter, as in Remark 13, this is not visible, as is shown in the corresponding definition in Theorem 14.
- The functional $\mathcal{K}^h_{DfDc}$ in (47) is not that of a "standard" gradient flow. The discussion in Sec. 1.2 and the proof of Theorem 12 suggest to split it into three parts; two parts that represent the diffusion steps for normal and decayed matter, and a third part for the decay step. The fact that the operator can be split into terms for each driving force is related to the independence of the processes in the micro model, so that the transition probability is a product of two probabilities, which can then be split according to calculation (51). Pursuing the analogy with the diffusion step, and with metric-space gradient flows, one might interpret $\mathcal{S}(\rho_{NN})+\mathcal{S}(\rho_{NT}-\rho_{NN})-\mathcal{S}(\rho_{NT})$ as the driving energy behind the decay, by which the dissipation would then become the (linear!) terms $-|\rho_{NN}|\log r^h_{NN} - |\rho_{ND}|\log r^h_{ND}$. In which sense this interpretation is meaningful is as yet unknown.

The way we have set up the microscopic model in this paper restricts us to decay processes. The reason that we cannot generalize to "birth" processes (i.e. $\lambda < 0$) is that, in the microscopic model, linear birth rates depend on the amount of existing normal matter. Therefore, in contrast to exponential decay, exponential birth requires a system of particles with interdependence, which prevents the techniques in this paper to be extended to birth processes in a trivial way.

The exact choice of the microscopic transition probabilities may not influence the continuum limit, as the limit only depends on asymptotic behavior of the probabilities $r^h_{\mu\nu}$ as $h \to 0$. However, this choice will affect the discrete-time approximation (47). In general, different microscopic systems can lead to different variational formulations for the same equation. For instance, the minimization functional (76) that we derive for a system of diffusion–reaction equations differs from the L^2-gradient flow in [44] for that same equation, as the underlying microscopic model of that paper models reaction as diffusion in a chemical landscape.

One of the interesting suggestions of the connection between large-deviation principles and gradient flows is the possibility that *every* gradient-flow structure might correspond to a large-deviation principle for *some* stochastic process. For instance, there is of course a different gradient-flow formulation for the diffusion-decay equation without drift (15), with driving energy

$$\mathcal{E}(\rho) := \int \left[\frac{1}{2}|\nabla\rho|^2 + \frac{\lambda}{2}\rho^2\right] dx,$$

and with the L^2-metric as dissipation. This can be seen by using the fact that in the Hilbert space L^2 a gradient flow satisfies at each time $t > 0$

$$(\partial_t\rho, s)_{L^2} = -\langle\mathcal{E}'(\rho), s\rangle \quad \text{for all } s \in L^2,$$

which can be rewritten as a weak form of (15). Could this structure be related to a large-deviation principle of some stochastic process? At this point we have no idea.

Appendix. The Quenched Large-Deviation Principle

In this Appendix we derive the large-deviation principles that are used in this paper — in a slightly more general context. First we state the large-deviation principle of the pair empirical measure. The proof is mainly due to Léonard, but we include it here to provide the full details. In the following, Ω will denote a (separable metric) Radon space.

Theorem A.1 ([31, Proposition 3.2]). *Fix $\rho^0 \in \mathcal{P}(\Omega)$ and let $\{x_i\}_{i=1,\ldots,n,n\geq 1} \subset \Omega$ be so that*

$$L^0_n := \frac{1}{n}\sum_{i=1}^{n}\delta_{x_i} \rightharpoonup \rho^0 \quad \text{as } n \to \infty. \tag{A.1}$$

Let $\zeta : \Omega \to \mathcal{P}(\Omega)$ be continuous with respect to the narrow topology of $\mathcal{P}(\Omega)$, and let each random variable Y_i in Ω be distributed by ζ^{x_i}. Define the pair empirical measure $M_n := n^{-1}\sum_{i=1}^{n}\delta_{(x_i,Y_i)}$. Then the sequence $\{M_n\}_n$ satisfies the large-deviation *principle in* $\mathcal{P}(\Omega^2)$ *with rate* n *and rate functional*:

$$I(q) := \begin{cases} \mathcal{H}(q\,|\,p) & \text{if } q(\cdot\times\Omega) = \rho^0(\cdot), \\ \infty & \text{otherwise}, \end{cases} \tag{A.2}$$

with $p(dx\,dy) := \zeta^x(dy)\rho^0(dx)$.

Proof. We write $C_b(\Omega^2)$ for the space of continuous bounded functions on Ω^2, and $C_b(\Omega^2)^*$ and $C_b(\Omega^2)'$ for its topological and algebraic dual respectively, the latter being the space of all linear functionals on $C_b(\Omega^2)$ with the weakest topology that makes all these linear functionals continuous. We equip both $C_b(\Omega^2)^*$ and $C_b(\Omega^2)'$ with the topology induced by the duality with $C_b(\Omega^2)$, denoted by $\langle\cdot,\cdot\rangle$. Recall that the dual $C_b(\Omega)^*$ can be identified with the space of finite, finitely additive, and regular signed Borel measures [17, Theorem IV.6.2]. Moreover, since Ω^2 is Radon any probability measure is regular. Hence $\mathcal{P}(\Omega^2) \subset C_b(\Omega^2)^* \subset C_b(\Omega^2)'$, and the topologies on $\mathcal{P}(\Omega^2)$ and $C_b(\Omega^2)^*$ coincide with the induced topology as a subset of $C_b(\Omega^2)'$. Note, however, that $C_b(\Omega^2)^*$ is closed, while $\mathcal{P}(\Omega^2)$ is not.

We first consider M_n as random variables in $C_b(\Omega^2)'$. For an arbitrary $d \in \mathbb{N}$ and $\phi_1,\ldots,\phi_d$ in $C_b(\Omega^2)$, define the new random variables:

$$\begin{aligned} Z_{\phi_1,\ldots,\phi_d;n} &:= (\langle\phi_1, M_n\rangle,\ldots,\langle\phi_d, M_n\rangle) \\ &= \left(\frac{1}{n}\sum_{i=1}^n \langle\phi_1,\delta_{(x_i,Y_i)}\rangle,\ldots,\frac{1}{n}\sum_{i=1}^n\langle\phi_d,\delta_{(x_i,Y_i)}\rangle\right) \\ &= \left(\frac{1}{n}\sum_{i=1}^n \phi_1(x_i,Y_i),\ldots,\frac{1}{n}\sum_{i=1}^n \phi_d(x_i,Y_i)\right). \end{aligned}$$

First we prove the large-deviation principle of $\mathrm{law}(Z_{\phi_1,\ldots,\phi_d;n})$ in $\mathbb{R}^d$, using the Gärtner–Ellis theorem. For any $\lambda \in \mathbb{R}^d$:

$$\begin{aligned} \Lambda_{\phi_1,\ldots,\phi_d;n}(\lambda) &:= \frac{1}{n}\log(\mathbb{E}\exp(n\lambda\cdot Z_{\phi_1,\ldots,\phi_d;n})) \\ &= \frac{1}{n}\log\left(\mathbb{E}\exp\left(\sum_{j=1}^d\sum_{i=1}^n \lambda_j\phi_j(x_i,Y_i)\right)\right) \\ &\overset{(*)}{=} \frac{1}{n}\log\left(\prod_{i=1}^n \mathbb{E}\exp\left(\sum_{j=1}^d \lambda_j\phi_j(x_i,Y_i)\right)\right) \\ &= \frac{1}{n}\sum_{i=1}^n\log\left(\int\exp\left(\sum_{j=1}^d\lambda_j\phi_j(x_i,y)\right)\zeta^{x_i}(dy)\right) \\ &= \int\frac{1}{n}\sum_{i=1}^n\log\left(\int\exp\left(\sum_{j=1}^d\lambda_j\phi_j(x,y)\right)\zeta^x(dy)\right)\delta_{x_i}(dx) \end{aligned}$$

$$
\begin{aligned}
&= \int \log\left(\int \exp\left(\sum_{j=1}^{d} \lambda_j \phi_j(x,y)\right) \zeta^x(dy)\right) L_n^0(dx) \\
&= \int \log\langle e^{\lambda\cdot\phi^x}, \zeta^x\rangle L_n^0(dx),
\end{aligned} \tag{A.3}
$$

using the notation $\phi^x : y \mapsto (\phi_1(x,y), \ldots, \phi_d(x,y))$. In $(*)$ we have used the independence of (x_i, Y_i) to take the sum out of the expectation.

In order to use (A.1) to pass to the limit $n \to \infty$ in (A.3), we need to show that $x \mapsto \log\langle e^{\lambda\cdot\phi^x}, \zeta^x\rangle$ is a bounded and continuous function. The boundedness follows directly from the fact that all ϕ_j are bounded. To prove continuity, take any convergent sequence $x_m \to x$. As ζ^x is continuous as a function from $x \in \Omega$ to $\mathcal{P}(\Omega)$, Prokhorov's theorem gives tightness of the sequence ζ^{x_m}. Hence for each $\epsilon > 0$ there exists a compact set $K_\epsilon \subseteq \Omega$ such that:

$$
\zeta^{x_m}(\Omega\backslash K_\epsilon) < \epsilon \quad \text{for all } m \geq 1.
$$

Using that the sequence of functions $y \mapsto e^{\lambda\cdot\phi^{x_m}(y)}$ converges uniformly on compact sets as $m \to \infty$, we have

$$
\begin{aligned}
&|\langle e^{\lambda\cdot\phi^{x_m}}, \zeta^{x_m}\rangle - \langle e^{\lambda\cdot\phi^x}, \zeta^x\rangle| \\
&\quad = |\langle e^{\lambda\cdot\phi^{x_m}} - e^{\lambda\cdot\phi^x}, \zeta^{x_m}\rangle + \langle e^{\lambda\cdot\phi^x}, \zeta^x - \zeta^{x_m}\rangle| \\
&\quad \leq \int_{\Omega\backslash K_\epsilon} |e^{\lambda\cdot\phi^{x_m}(y)} - e^{\lambda\cdot\phi^x(y)}|\zeta^{x_m}(dy) \\
&\qquad + \int_{K_\epsilon} |e^{\lambda\cdot\phi^{x_m}(y)} - e^{\lambda\cdot\phi^x(y)}|\zeta^{x_m}(dy) + |\langle e^{\lambda\cdot\phi^x}, \zeta^x - \zeta^{x_m}\rangle| \\
&\quad \leq (\|e^{\lambda\cdot\phi^{x_m}}\|_{L^\infty(\Omega)} + \|e^{\lambda\cdot\phi^x}\|_{L^\infty(\Omega)}) \underbrace{\zeta^{x_m}(\Omega\backslash K_\epsilon)}_{<\epsilon} \\
&\qquad + \underbrace{\|e^{\lambda\cdot\phi^{x_m}} - e^{\lambda\cdot\phi^x}\|_{L^\infty(K_\epsilon)}}_{\to 0} \zeta^{x_m}(K_\epsilon) + \underbrace{|\langle e^{\lambda\cdot\phi^x}, \zeta^x - \zeta^{x_m}\rangle|}_{\to 0} \\
&\quad \xrightarrow{m\to\infty} 2\epsilon \|e^{\lambda\cdot\phi^x}\|_{L^\infty(\Omega)}
\end{aligned}
$$

for arbitrary small $\epsilon > 0$. Hence indeed $\langle e^{\lambda\cdot\phi^x}, \zeta^x\rangle$ is continuous in x, so we can apply (A.1) to find the limit:

$$
\Lambda_{\phi_1,\ldots,\phi_d}(\lambda) := \lim_{n\to\infty} \Lambda_{\phi_1,\ldots,\phi_d;n}(\lambda) = \int \log\langle e^{\lambda\cdot\phi^x}, \zeta^x\rangle \rho^0(dx).
$$

Since this function is continuously differentiable and finite throughout its whole domain ($\mathbb{R}^d$), the conditions of the Gärtner–Ellis theorem [14, Theorem 2.3.6c] are met, so that $Z_{\phi_1,\ldots,\phi_d;n}$ satisfies the large-deviation principle in $\mathbb{R}^d$ with rate n and rate function $\Lambda^*_{\phi_1,\ldots,\phi_d}$, the Fenchel–Legendre transform of $\Lambda_{\phi_1,\ldots,\phi_d}$.

Next we apply the Dawson–Gärtner theorem [14, Theorem 4.6.9] to find that the sequence $\{M_n\}_n$ satisfies the large-deviation principle in $C_b(\Omega^2)'$ with rate n

and rate functional:

$$\begin{aligned} I(q) &:= \sup_{d\geq 1} \sup_{\phi_1,\ldots,\phi_d\in C_b(\Omega^2)} \Lambda^*_{\phi_1,\ldots,\phi_d}((\langle\phi_1,q\rangle,\ldots,\langle\phi_d,q\rangle)) \\ &= \sup_{d\geq 1} \sup_{\phi_1,\ldots,\phi_d\in C_b(\Omega^2)} \sup_{\lambda\in\mathbb{R}^d} \lambda\cdot(\langle\phi_1,q\rangle,\ldots,\langle\phi_d,q\rangle) - \Lambda_{\phi_1,\ldots,\phi_d}(\lambda) \\ &= \sup_{\phi\in C_b(\Omega^2)} \langle\phi,q\rangle - \int \log\langle e^{\phi^x},\zeta^x\rangle\rho^0(dx), \end{aligned}$$

where as before we write $\phi^x : y \mapsto \phi(x,y)$.

We now show that this rate functional is indeed (A.2). Since $C_b(\Omega^2)^*$ is a closed subset of $C_b(\Omega^2)'$ containing $\mathcal{P}(\Omega^2)$, we have $I=\infty$ on $C_b(\Omega^2)'\backslash C_b(\Omega^2)^*$ [14, Theorem 4.1.5]. Therefore, we only need to consider $q\in C_b^*(\Omega^2)$.

- First, we show that $I(q)=\infty$ whenever $q\in C_b^*(\Omega^2)$ with first marginal $\pi^1 q\neq\rho^0$. This can be seen by restricting the supremum to ϕ's that depend on the first variable only:

$$\begin{aligned} I(q) &\geq \sup_{\phi\in C_b(\Omega)} \langle\phi,q\rangle - \int \log\langle e^{\phi^x},\zeta^x\rangle\rho^0(dx) \\ &= \sup_{\phi\in C_b(\Omega)} \langle\phi,\pi^1 q\rangle - \langle\phi,\rho^0\rangle \\ &= \begin{cases} 0 & \text{if } \pi^1 q = \rho^0, \\ +\infty & \text{otherwise.} \end{cases} \end{aligned}$$

- Next, we show that $I(q)=\infty$ for any $q\in C_b(\Omega^2)^*$ that is finitely, but not countably additive. By the argument above, we only need to consider non-negative finitely additive measures with $q(\Omega^2)=1$. For such q, there exists a sequence of disjoint measurable sets $A_i\subset\Omega^2$ such that

$$\delta := q\left(\bigcup_{i=1}^\infty A_i\right) - \sum_{i=1}^\infty q(A_i) > 0.$$

  Without loss of generality, assume that $\bigcup_{i=1}^\infty A_i = \Omega^2$. Since q and p are regular, one can find for any $k\geq 1$, sequences of sets $K_i\subset A_i\subset O_i$ with K_i compact and O_i open, such that:

$$\sum_{i=1}^\infty q(O_i) \leq 1 - \frac{1}{2}\delta \quad\text{and}\quad \sum_{i=1}^\infty p(A_i\backslash K_i) \leq e^{-k}. \tag{A.4}$$

  Then for each $k,n\geq 1$ there exists a continuous function $\phi_{kn}:\Omega^2\to[-k,0]$ such that

$$\phi_{kn}(x,y) = \begin{cases} -k & \text{on } \bigcup_{i=1}^n K_i, \\ 0 & \text{on } \Omega^2 \Big\backslash \bigcup_{i=1}^n O_i. \end{cases}$$

For these functions we have, on the one hand (as O_i might not be disjoint)

$$\langle \phi_{kn}, q\rangle \geq -kq\left(\bigcup_{i=1}^n O_i\right) \geq -k\sum_{i=1}^n q(O_i), \tag{A.5}$$

and on the other hand

$$\langle e^{\phi_{kn}^x}, \zeta^x\rangle \leq \int (e^{-k}\chi_{\bigcup_{i=1}^n K_i}(x,y) + \chi_{\Omega^2\setminus\bigcup_{i=1}^n K_i}(x,y))\zeta^x(dy),$$

so that

$$\begin{aligned}
&\int \log\langle e^{\phi_{kn}^x}, \zeta^x\rangle \rho^0(dx)\\
&\quad\leq \int\left(-k + \log\int(\chi_{\bigcup_{i=1}^n K_i} + e^k\chi_{\Omega^2\setminus\bigcup_{i=1}^n K_i})\zeta^x\right)\rho^0(dx)\\
&\quad\overset{\text{Jensen}}{\leq} -k + \log\left(p\left(\bigcup_{i=1}^n K_i\right) + e^k p\left(\Omega^2\Big\backslash\bigcup_{i=1}^n K_i\right)\right).
\end{aligned} \tag{A.6}$$

Finally, we find for the rate functional:

$$\begin{aligned}
I(q) \quad &\geq \quad \limsup_{k\to\infty}\limsup_{n\to\infty}\langle\phi_{kn}, q\rangle - \int\log\langle e^{\phi_{kn}^x}, \zeta^x\rangle\rho^0(dx)\\
&\overset{(A.5),(A.6)}{\geq} \quad \limsup_{k\to\infty}\limsup_{n\to\infty} -k\sum_{i=1}^n q(O_i) + k\\
&\qquad - \log\left(p\left(\bigcup_{i=1}^n K_i\right) + e^k p\left(\Omega^2\Big\backslash\bigcup_{i=1}^n K_i\right)\right)\\
&= \quad \limsup_{k\to\infty} -k\sum_{i=1}^\infty q(O_i) + k - \log\left(p\left(\bigcup_{i=1}^\infty K_i\right) + e^k p\left(\Omega^2\Big\backslash\bigcup_{i=1}^\infty K_i\right)\right)\\
&\overset{(A.4)}{\geq} \quad \limsup_{k\to\infty} -k\left(1 - \frac{1}{2}\delta\right) + k - \log 2\\
&= \quad \limsup_{k\to\infty}\frac{1}{2}\delta k - \log 2 = \infty.
\end{aligned}$$

- Now assume that $q \in \mathcal{P}(\Omega^2)$ such that $\pi^1 q = \rho^0$. The Disintegration theorem then allows us to write

$$q(dxdy) = \rho^0(dx)q^x(dy)$$

for some family of measures $\{q^x : x \in \Omega\}$. In this case:

$$\begin{aligned}
I(q) &= \sup_{\phi\in C_b(\Omega^2)}\int(\langle\phi^x, q^x\rangle - \log\langle e^{\phi^x}, \zeta^x\rangle)\rho^0(dx)\\
&\leq \int\sup_{\phi^x\in C_b(\Omega)}\{\langle\phi^x, q^x\rangle - \log\langle e^{\phi^x}, \zeta^x\rangle\}\rho^0(dx)
\end{aligned}$$

$$= \int \mathcal{H}(q^x \,|\, \zeta^x)\rho^0(dx)$$

$$= \begin{cases} \displaystyle\iint \left(\log \frac{d(\rho^0 q^x)}{d(\rho^0 \zeta^x)}(x,y)\right) \rho^0(dx) q^x(dy) & \text{if } \rho^0 q^x \ll \rho^0 \zeta^x, \\ \infty & \text{otherwise} \end{cases}$$

$$= \mathcal{H}(q \,|\, p).$$

- We conclude the proof with the inequality in the other direction. Observe that I is the Fenchel–Legendre transform of

$$\Lambda : \phi \mapsto \int \log\langle e^{\phi^x}, \zeta^x\rangle \rho^0(dx)$$
$$\leq \log \int \langle e^{\phi^x}, \zeta^x\rangle \rho^0(dx) = \log\langle e^{\phi}, p\rangle,$$

  where the bound follows from Jensen's inequality. Hence

$$I(q) = \Lambda^*(q) \geq \sup_{\phi \in C(\Omega^2)} \{\langle \phi, q\rangle - \log\langle e^\phi, p\rangle\} = \mathcal{H}(q \,|\, p).$$

Since the large-deviation principle holds in $C_b(\Omega^2)^*$ with $D_I \subset \mathcal{P}(\Omega^2)$, it also holds in $\mathcal{P}(\Omega^2)$ with the same rate functional (i.e. restricted to $\mathcal{P}(\Omega^2)$) [14, Theorem 4.1.5]. □

The following corollary follows immediately from the contraction principle.

Corollary A.2. *The sequence* $\{n^{-1}\sum_{i=1}^n \delta_{Y_i}\}_n$ *satisfies the large-deviation principle in* $\mathcal{P}(\Omega)$ *with rate* n *and rate function*:

$$J(\rho) := \begin{cases} \displaystyle\inf_{q \in \Gamma(\rho^0,\rho)} \mathcal{H}(q \,|\, p) & \textit{if } q \in \Gamma(\rho^0, \rho), \\ \infty & \textit{otherwise}. \end{cases} \tag{A.7}$$

Remark A.3. A straightforward approach would be to look for a large-deviation principle in the set of probability measures:

$$A \mapsto \mathbb{P}(M_n \in A \,|\, L_n^0 = \rho^0). \tag{A.8}$$

However, these conditional probabilities are not well-defined: the events $\{L_n^0 = \rho^0\}$ typically have zero probability. One way to deal with this is to condition on small neighborhoods of ρ^0 of size δ instead, calculate the large-deviation rate functional for these conditional probabilities, and then take the limit for $\delta \to 0$. This is the approach taken in [1]. We note that because the limits $n \to \infty$ and $\delta \to 0$ cannot be interchanged, this approach does not *a priori* yield a large-deviation principle in the rigorous sense.

In the approach that we adopt from [31], we consider fixed initial positions so that there is no need to define the conditional probabilities above. This technique is sometimes called a *quenched large-deviation principle*.

Acknowledgments

The research of the first author has benefited from support by the Leverhulme Trust and by the Initial Training Network "FIRST" of the Seventh Framework Programme of the European Community (grant agreement number 238702).

References

[1] S. Adams, N. Dirr, M. A. Peletier and J. Zimmer, From a large-deviations principle to the Wasserstein gradient flow: A new micro-macro passage, *Commun. Math. Phys.* **307** (2011) 791–815.

[2] M. Agueh, Existence of solutions to degenerate parabolic equations via the Monge–Kantorovich theory, *Adv. Differential Equations* **10** (2005) 309–360.

[3] F. Almgren, J. E. Taylor and L. Wang, Curvature-driven flows: A variational approach, *SIAM J. Control Optim.* **31** (1993) 387–437.

[4] L. Ambrosio, N. Gigli and G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, 2nd edn., Lectures in Mathematics ETH Zürich (Birkhäuser, Basel, 2008).

[5] S. Arnrich, A. Mielke, M. A. Peletier, G. Savaré and M. Veneroni, Passing to the limit in a Wasserstein gradient flow: From diffusion to reaction, *Calc. Var. Partial Differential Equations* **44** (2012) 419–454.

[6] D. Aronson, Bounds for the fundamental solution of a parabolic equation, *Bull. Amer. Math. Soc.* **73** (1967) 890–896.

[7] V. I. Bogachev, *Measure Theory*, Vol. I (Springer, Berlin, 2007).

[8] H. Brezis and E. Lieb, A relation between pointwise convergence of functions and convergence of functionals, *Proc. Amer. Math. Soc.* **88** (1983) 486–490.

[9] J. A. Carrillo, R. J. McCann and C. Villani, Kinetic equilibration rates for granular media and related equations: Entropy dissipation and mass transportation estimates, *Rev. Mat. Iberoamericana* **19** (2003) 971–1018.

[10] G. Dal Maso, *An Introduction to* Γ*-Convergence*, 1st edn., Progress in Nonlinear Differential Equations and their Applications, Vol. 8 (Birkhäuser, Boston, 1993).

[11] D. A. Dawson and J. Gärtner, Multilevel large deviations and interacting diffusions, *Probab. Theory Related Fields* **98** (1994) 423–487.

[12] E. De Giorgi, New problems on minimizing movements, in *Ennio de Giorgi: Selected Papers* (Springer-Verlag, Berlin, 2006), pp. 699–713.

[13] A. De Masi and E. Presutti, *Mathematical Methods for Hydrodynamic Limits*, Lecture Notes in Mathematics, Vol. 1501 (Springer-Verlag, Berlin, 1991).

[14] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd edn., Stochastic Modelling and Applied Probability, Vol. 38 (Springer-Verlag, New York, 1987).

[15] F. den Hollander, *Large Deviations*, Fields Institute Monographs (American Mathematical Society, Providence, RI, 2000).

[16] R. Dudley, *Real Analysis and Probability* (Wadsworth and Brooks/Cole, Pacific Grove, CA, 1989).

[17] N. Dunford and J. Schwartz, *Linear Operators, Part I, General Theory* (Interscience, New York, 1957).

[18] M. Duong, V. Laschos and M. Renger, Wasserstein gradient flows from large deviations of many-particle limits, accepted for publication in *ESAIM Control Optim. Calc. Var.* (2013).

[19] M. A. Efendiev and A. Mielke, On the rate-independent limit of systems with dry friction and small viscosity, *J. Convex Anal.* **13** (2006) 151–167.
[20] A. Einstein, Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen, *Ann. Phys.* **17** (1905) 548–560.
[21] J. Feng and T. Kurtz, *Large Deviations for Stochastic Processes*, Mathematical Surveys and Monographs, Vol. 131 (American Mathematical Society, Providence, RI, 2006).
[22] A. Friedman, *Partial Differential Equations of Parabolic Type* (Prentice-Hall, Englewood Cliffs, NJ, 1964).
[23] L. Giacomelli and F. Otto, Variational formulation for the lubrication approximation of the Hele–Shaw flow, *Calc. Var. Partial Differential Equations* **13** (2001) 377–403.
[24] U. Gianazza, G. Savaré and G. Toscani, The Wasserstein gradient flow of the Fisher information and the quantum drift–diffusion equation, *Arch. Ration. Mech. Anal.* **194** (2009) 133–220.
[25] E. Giusti, *Direct Methods in the Calculus of Variations* (World Scientific, Singapore, 2003).
[26] K. Glasner, A diffuse-interface approach to Hele–Shaw flow, *Nonlinearity* **16**(1) (2003) 49–66.
[27] R. Jordan, D. Kinderlehrer and F. Otto, Free energy and the Fokker–Planck equation, *Phys. D* **107** (1997) 265–271.
[28] ———, The variational formulation of the Fokker–Planck equation, *SIAM J. Math. Anal.* **29** (1998) 1–17.
[29] D. Kinderlehrer and N. Walkington, Approximation of parabolic equations using the Wasserstein metric, *ESAIM*: *Math. Model. Numer. Anal.* **33** (1999) 837–852.
[30] M. Kočvara, A. Mielke and T. Roubíček, A rate-independent approach to the delamination problem, *Math. Mech. Solids* **11** (2006) 423–447.
[31] C. Léonard, A large deviation approach to optimal transport, preprint (2007); arXiv.org/abs/0710.1461v1.
[32] S. Luckhaus and T. Sturzenhecker, Implicit time discretization for the mean curvature flow equation, *Calc. Var. Partial Differential Equations* **3** (1995) 253–271.
[33] D. Matthes, R. J. McCann and G. Savaré, A family of nonlinear fourth order equations of gradient flow type, *Commun. Partial Differential Equations* **34** (2009) 1352–1397.
[34] A. Mielke, A new approach to elasto-plasticity using energy and dissipation functionals, in *Applied Mathematics Entering the 21st Century*, eds. J. M. Hill and R. Moore (SIAM, PA, 2004), pp. 315–335.
[35] ———, Evolution in rate-independent systems, *Handbook of Differential Equations*: *Evolutionary Differential Equations* (North-Holland, Amsterdam, 2005), pp. 461–559.
[36] ———, A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems, *Nonlinearity* **24** (2011) 1329–1346.
[37] A. Mielke, R. Rossi and G. Savaré, Nonsmooth analysis of doubly nonlinear evolution equations, *Calc. Var. Partial Differential Equations* **46**(1–2) (2013) 253–310.
[38] A. Mielke and F. Theil, On rate-independent hysteresis models, *Nonlinear Differential Equations Appl.* **11** (2004) 151–189.
[39] A. Mielke, F. Theil and V. I. Levitas, A variational formulation of rate-independent phase transformations using an extremum principle, *Arch. Ration. Mech. Anal.* **162** (2002) 137–177.
[40] U. Mosco, Convergence of convex sets and of solutions of variational inequalities, *Adv. Math.* **3** (1969) 510–585.

[41] F. Otto, Dynamics of labyrinthine pattern formation in magnetic fluids: A mean-field theory, *Arch. Ration. Mech. Anal.* **141** (1998) 63–103.

[42] ———, Lubrication approximation with prescribed nonzero contact angle, *Commun. Partial Differential Equations* **23** (1998) 63–103.

[43] ———, The geometry of dissipative evolution equations: The porous medium equation, *Commun. Partial Differential Equations* **26** (2001) 101–174.

[44] M. Peletier, G. Savaré and M. Veneroni, From diffusion to reaction via γ-convergence, *SIAM J. Math. Anal.* **42** (2010) 1805–1825.

[45] J. W. Portegies and M. A. Peletier, Well-posedness of a parabolic moving-boundary problem in the setting of Wasserstein gradient flows, *Interfaces Free Boundaries* **12** (2010) 121–150.

[46] R. Rossi, A. Mielke and G. Savare, A metric approach to a class of doubly nonlinear evolution equations and applications, *Ann. Scuole Norm. Sup. Pisa Cl. Sci.* (*5*) **7** (2008) 97–169.

[47] E. Sandier and S. Serfaty, Gamma-convergence of gradient flows with applications to Ginzburg–Landau, *Commun. Pure Appl. Math.* **57**(12) (2004) 1627–1672.

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

[49] ———, *Optimal Transport*: *Old and New* (Springer-Verlag, Berlin, 2009).