ESAIM: COCV 19 (2013) 1166–1188
DOI: 10.1051/cocv/2013049

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

WASSERSTEIN GRADIENT FLOWS FROM LARGE DEVIATIONS OF MANY-PARTICLE LIMITS

Manh Hong Duong[1], Vaios Laschos[1] and Michiel Renger[2]

Abstract. We study the Fokker–Planck equation as the many-particle limit of a stochastic particle system on one hand and as a Wasserstein gradient flow on the other. We write the path-space rate functional, which characterises the large deviations from the expected trajectories, in such a way that the free energy appears explicitly. Next we use this formulation via the contraction principle to prove that the discrete time rate functional is asymptotically equivalent in the Gamma-convergence sense to the functional derived from the Wasserstein gradient discretization scheme.

Mathematics Subject Classification. 35A15, 5Q84.

Received March 19, 2012. Revised January 4, 2013.
Published online August 13, 2013.

1. Introduction

Since the seminal work of Jordan, Otto and Kinderlehrer [12], it has become clear that there are many more partial differential equations that can be written as a gradient flow than previously known. Two important insights have contributed to this: the generalisation of gradient flows to metric spaces and the specific choice of the Wasserstein metric as the dissipation mechanism. The paper by Jordan, Kinderlehrer and Otto introduced a gradient-flow structure by approximation in discrete time. More recent work have shown how these ideas can be studied in continuous time [14], and how they can be generalised to any metric space [3]. This paper is mainly concerned with the time-discrete scheme, which we shall first explain.

In the space $L^2(\mathbb{R}^d)$, a gradient flow is an evolution equation of the form

$$\partial_t \rho_t = -\mathrm{grad}_{L^2} \mathcal{F}(\rho_t), \tag{1.1}$$

for some functional $\mathcal{F}$, where $\mathrm{grad}_{L^2}\mathcal{F}(\rho)$ denotes its functional derivative. For a gradient flow it is natural to use the following time-discrete variational scheme. If ρ_0 is the solution at time $t = 0$, then the solution at time $\tau > 0$ is approximated by the minimiser of the functional

$$\rho \mapsto \mathcal{F}(\rho) + \frac{1}{2\tau}\|\rho - \rho_0\|^2_{L^2(\mathbb{R}^d)}.$$

Keywords and phrases. Wasserstein, gradient flows, Fokker–Planck, Gamma-convergence, large deviations.

[1] Department of Mathematical sciences, University of Bath, Claverton Down, Bath BA2 7AY, United Kingdom.
[2] ICMS and Dep. of Math. and Comp. Sciences, TU Eindhoven, Den Dolech 2, 5612 AZ Eindhoven, the Netherlands.
d.r.m.renger@tue.nl

Article published by EDP Sciences

© EDP Sciences, SMAI 2013

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

Indeed, the Euler–Lagrange equation is then $\frac{\rho_\tau - \rho_0}{\tau} = -\mathrm{grad}_{L^2}\mathcal{F}(\rho_\tau)$, which approximates (1.1) as $\tau \to 0$. In the same manner, one can define a variational scheme by minimising the functional

$$\rho \mapsto \mathcal{F}(\rho) + \frac{1}{2\tau} W_2^2(\rho, \rho_0), \tag{1.2}$$

where W_2 is the Wasserstein metric. Convergence of this variational scheme was first proven in [12] with the choice of $\mathcal{F}(\rho) := \mathcal{S}(\rho) + \mathcal{E}(\rho)$, where

$$\mathcal{E}(\rho) = \int_{\mathbb{R}^d} \Psi(\mathbf{x})\rho(\mathrm{d}\mathbf{x}) \qquad \text{and} \qquad \mathcal{S}(\rho) := \begin{cases} \int_{\mathbb{R}^d} \rho(\mathbf{x}) \log \rho(\mathbf{x})\, \mathrm{d}\mathbf{x}, & \text{for } \rho(\mathrm{d}\mathbf{x}) = \rho(\mathbf{x})\, \mathrm{d}\mathbf{x} \\ \infty, & \text{otherwise,} \end{cases} \tag{1.3}$$

for some potential Ψ. In this case, the minimisers converge to the solution of the Fokker–Planck equation

$$\partial_t \rho_t = \Delta \rho_t + \mathrm{div}(\rho_t \nabla \Psi). \tag{1.4}$$

Later, in [14], this result was extended to more general $\mathcal{F}$, but we will be concerned with the specific choice (1.3). Physically, $\mathcal{S}$ can be interpreted as entropy, $\mathcal{E}$ as internal energy, and $\mathcal{F}$ as the corresponding Helmholtz free energy (if the temperature effects are hidden in Ψ); hence it is not surprising that this free energy should decay along solutions of (1.4). However, it is not intuitively clear why the dissipation of free energy must be described by the Wasserstein metric.

As we will explain in Section 3, for systems in equilibrium, the stochastic fluctuations around the equilibrium are characterised by a free energy similar to (1.3). Recent developments suggest a similar principle for systems away from equilibrium [1, 2, 7, 13, 15]. To explain this, consider N independent random particles in $\mathbb{R}^d$ with positions $X_k(t)$, initially distributed by some $\rho_0 \in \mathcal{P}(\mathbb{R}^d)$, where the probability distribution of each particle evolves according to (1.4). Define the corresponding *empirical process*

$$L_N : t \mapsto \frac{1}{N} \sum_{k=1}^{N} \delta_{X_k(t)}.$$

Then, as a consequence of the Law of Large Numbers, at each $\tau \geq 0$ the *empirical measure* $L_N(\tau)$ converges almost surely in the narrow topology as $N \to \infty$ to the solution of the Fokker–Planck equation (1.4) with initial condition ρ_0 [8]. The rate of this convergence is characterised by a large deviation principle. Roughly speaking, this means that there exists a unique $J_\tau : \mathcal{P}(\mathbb{R}^d) \times \mathcal{P}(\mathbb{R}^d) \to [0, \infty]$ such that (see Sect. 3)

$$\mathrm{Prob}\left(L_N(\tau) \approx \rho \,|\, L_N(0) \approx \rho_0\right) \sim \exp\left(-N\, J_\tau(\rho|\rho_0)\right) \qquad \text{as } N \to \infty.$$

In [13], Proposition 3.2 and [15], Corollary 13, it was found that

$$J_\tau(\rho|\rho_0) = \inf\left\{\mathcal{H}(\gamma|\rho_0\, p_\tau) : \gamma \in \mathit{\Pi}(\rho_0, \rho)\right\} \tag{1.5}$$

where $\mathcal{H}$ is the relative entropy (discussed in Sect. 3), p_τ is the fundamental solution of the Fokker–Planck equation (1.4) and $\mathit{\Pi}(\rho_0, \rho)$ is the set of all Borel measures in $\mathbb{R}^d \times \mathbb{R}^d$ that have first and second marginal ρ_0 and ρ respectively. In this paper, we characterise a class of potentials Ψ and initial data ρ_0 for which (1.5) is equal to

$$\inf_{\rho_{(\cdot)} \in C_{W_2}(\rho_0, \rho)} \left\{ \frac{1}{4\tau} \int_0^1 \|\partial_t \rho_t\|^2_{-1,\rho_t}\, \mathrm{d}t + \frac{\tau}{4} \int_0^1 \|\Delta \rho_t + \mathrm{div}(\rho_t \nabla \Psi)\|^2_{-1,\rho_t}\, \mathrm{d}t + \frac{1}{2}\mathcal{F}(\rho_1) - \frac{1}{2}\mathcal{F}(\rho_0) \right\}. \tag{1.6}$$

where the $\|\cdot\|_{-1,\rho}$ norm and the exact meaning of $\partial_t \rho_t$, $\Delta \rho_t$ and $\mathrm{div}(\rho_t \nabla \Psi)$ will be defined in the sequel. In the main theorem, by using the above equality, we show that the Wasserstein scheme (1.2) has the same asymptotic behavior with J_τ for $\tau \to 0$, in terms of Gamma-convergence (see [5] for an exposition of Gamma-convergence).

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

Theorem 1.1. *Let $\rho_0 = \rho_0(x)\mathrm{d}x \in \mathcal{P}_2(\mathbb{R})$ be absolutely continuous with respect to the Lebesgue measure and with density $\rho_0(x)$ being bounded from below by a positive constant in every compact set. Assume that $\int_{\mathbb{R}} |\nabla\Psi(x)|^2\, \rho_0(\mathrm{d}x)$ and the Fisher information $I(\rho_0)$ (introduced in Sect. 2) are finite, and that Ψ satisfies either Assumption 4.1 or 4.4 (introduced in Sect. 4). Then we have*

$$J_\tau(\cdot\,|\rho_0) - \frac{W_2^2(\rho_0,\,\cdot\,)}{4\tau} \xrightarrow[\tau\to 0]{\Gamma} \frac{1}{2}\mathcal{F}(\,\cdot\,) - \frac{1}{2}\mathcal{F}(\rho_0), \qquad \text{in } \mathcal{P}_2(\mathbb{R}). \tag{1.7}$$

Here $\mathcal{P}_2(\mathbb{R})$ denotes the space of probability measures on $\mathbb{R}$ having finite second moment. Actually, we will prove Mosco convergence of (1.7), *i.e.* in Theorem 5.1 we prove that the Gamma-convergence lower bound holds for any sequence in $\mathcal{P}_2(\mathbb{R})$, equipped with the narrow topology, and in Theorem 6.1 we prove the existence of the recovery sequence in the Wasserstein topology. This is equivalent to having Gamma convergence in both topologies.

In the Wasserstein topology, the Gamma-convergence (1.7) immediately implies:

$$\tau\, J_\tau(\cdot\,|\rho_0) \xrightarrow[\tau\to 0]{\Gamma} \frac{1}{4}W_2^2(\rho_0,\,\cdot\,) \qquad \text{in } \mathcal{P}_2(\mathbb{R}). \tag{1.8}$$

For a system of Brownian particles, i.e. $\Psi \equiv 0$, statement (1.8) can also be found in [13]. Together, the two statements (1.7) and (1.8) make up an asymptotic development of the rate J_τ for small τ, *i.e.*

$$J_\tau(\rho|\rho_0) \approx \frac{1}{2}\mathcal{F}(\rho) - \frac{1}{2}\mathcal{F}(\rho_0) + \frac{1}{4\tau}W_2^2(\rho_0,\rho).$$

Apart from the factor 1/2 and the constant $\mathcal{F}(\rho_0)$, which do not affect the minimisers, this approximation indeed corresponds to the functional defining the time-discrete variational scheme (1.2) from [12].

For $\Psi = 0$, the main statement (1.7) was proven in [1] in a subset of $\mathcal{P}_2(\mathbb{R})$ consisting of measures that are sufficiently close to a uniform distribution on a compact interval. In [15], it was proven that whenever (1.7) holds for $\Psi = 0$, then it also holds for any $\Psi \in C_b^2(\mathbb{R}^d)$. Both papers make use of the specific form of the fundamental solution of (1.4). In [7], (1.7) was shown for Gaussian measures on the real line. The novelty of the present paper lies in the use of large deviations in the space of trajectories rather than conditional large deviations of the form (1.5). The conditional large deviations are obtained by a contraction. This provides us with an alternative formulation of the rate functional from which, formally, the Gamma lower bound follows immediately. Moreover, this approach allows us to prove the Gamma convergence in a much more general context.

All theorems in this paper are valid in higher dimensions except for the existence of the recovery sequence. There are a number of reasons why, at least by the approach of this paper, the argument fails in higher dimensions. First of all, in the proof of Lemma 6.3 we use an explicit formula of optimal transport maps in terms of cumulative distribution functions. Secondly, the proof of the same lemma in higher dimensions would require regularity and global estimates of derivates of the transport map, which are still unknown today (see for example [17], p. 141).

The required concepts of this paper are introduced in Section 2. In Section 3, we explain the concept of large deviations in the case of an equilibrium system, introduce the dynamical particle system that we study more precisely, and discuss the conditional large deviations for this system. The alternative form of the functional J_τ is proven in Section 4 *via* the path-wise large deviation principles. Finally, in Section 5 we prove the Gamma-convergence lower bound, and in Section 6 the existence of the recovery sequence.

2. Preliminaries

By the nature of this study, we need a combination of techniques from probability theory, mostly from the theory of large deviations, and from functional analysis, mostly from the gradient flow calculus as set out in [3]. Let us introduce these concepts here.

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

2.1. Topological measure spaces

Unless otherwise stated, the space of probability measures $\mathcal{P}(\mathbb{R}^d)$ will be endowed with the narrow topology, defined by convergence against continuous bounded test functions:

$$\rho_t \to \rho \text{ as } t \to 0 \quad \text{ if and only if } \quad \int_{\mathbb{R}^d} \phi \, \mathrm{d}\rho_t \to \int_{\mathbb{R}^d} \phi \, \mathrm{d}\rho \quad \text{ for all } \phi \in C_b(\mathbb{R}^d).$$

We sometimes identify measures with densities when possible, which is typically the case if a measure has finite entropy. The space $\mathcal{P}_2(\mathbb{R}^d) = \left\{\rho \in \mathcal{P}(\mathbb{R}^d) : \int |\mathbf{x}|^2 \, \rho(\mathrm{d}\mathbf{x}) < \infty\right\}$ will be endowed with the topology generated by the Wasserstein metric W_2. The Wasserstein distance of two measures $\rho_0, \rho \in \mathcal{P}_2(\mathbb{R}^d)$ is defined via

$$W_2^2(\rho_0, \rho) = \inf_{\gamma \in \Pi(\rho_0, \rho)} \left\{ \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |\mathbf{x} - \mathbf{y}|^2 \mathrm{d}\gamma \right\}.$$

We note that convergence in the Wasserstein topology is equivalent to narrow convergence together with convergence of second moments [17], Theorem 7.12.

2.2. Continuous and absolutely continuous curves

We write $C([0,1], \mathcal{P}(\mathbb{R}^d))$ for the space of narrowly continuous curves $[0,1] \to \mathcal{P}(\mathbb{R}^d)$, and $C(\rho_0, \rho)$ for the space of narrowly continuous curves $[0,1] \to \mathcal{P}(\mathbb{R}^d)$ starting in ρ_0 and ending in ρ. Similarly, for Wasserstein-continuous curves in $\mathcal{P}_2(\mathbb{R}^d)$ we write $C_{W_2}([0,1], \mathcal{P}_2(\mathbb{R}^d))$ and $C_{W_2}(\rho_0, \rho)$.

Furthermore, we use two different notions of absolutely continuous curves. The first notion is taken from [6], Definition 4.1. Let $\mathcal{D} = C_c^\infty(\mathbb{R}^d)$ be the space of test functions with the corresponding topology (see [16], Sect. 6.2), let $\mathcal{D}'$ be its dual, consisting of the associated distributions, and let $\langle \, , \, \rangle$ be the dual pairing between $\mathcal{D}'$ and $\mathcal{D}$. We will identify a measure $\rho \in \mathcal{P}(\mathbb{R}^d)$ with a distribution by setting $\langle \rho, f \rangle := \int f \, \mathrm{d}\rho$. Denote by $\mathcal{D}_K \subset \mathcal{D}$ the subspace of all Schwartz functions with compact support $K \subset \mathbb{R}^d$. Then a curve $\rho_{(\cdot)} : [0,1] \to \mathcal{D}'$ is said to *be absolutely continuous in the distributional sense* if for each compact set $K \subset \mathbb{R}^d$ there is a neighborhood U_K of 0 in $\mathcal{D}_K$ and an absolutely continuous function $G_K : [0,1] \to \mathbb{R}$ such that

$$|\langle \rho_{t_2}, f \rangle - \langle \rho_{t_1}, f \rangle| \leq |G_K(t_2) - G_K(t_1)|,$$

for all $0 < t_1, t_2 < 1$ and $f \in U_K$. Note that if a curve $\rho_{(\cdot)} : [0,1] \to \mathcal{D}'$ is absolutely continuous then the derivative in the distributional sense $\partial_t \rho_t = \lim_{\tau \to 0} \frac{1}{\tau}(\rho_{t+\tau} - \rho_t)$ exists for almost all $t \in [0,1]$.

Secondly, we say a curve $\rho_{(\cdot)} : [0,1] \to \mathcal{P}_2(\mathbb{R}^d)$ is *absolutely continuous in the Wasserstein sense* if there exists a $g \in L^1(0,1)$ such that

$$W_2(\rho_{t_1}, \rho_{t_2}) \leq \int_{t_1}^{t_2} g(t) \, \mathrm{d}t$$

for all $0 < t_1 \leq t_2 < 1$ (see for example [3], Def. 1.1.1).

Sufficient conditions for absolute continuity in the Wasserstein sense are given by the following useful lemma:

Lemma 2.1 ([3], Theorem 8.3.1)**.** *Let $\rho_{(\cdot)} : (0,\tau) \to \mathcal{P}(\mathbb{R}^d)$ be a narrowly continuous curve and let $\mathbf{v}_{(\cdot)}$ be a vector field such that the continuity equation holds:*

$$\partial_t \rho_t + \operatorname{div}(\rho_t \, \mathbf{v}_t) = 0 \qquad \text{in distributional sense.} \tag{2.1}$$

If

$$\rho_0 \in \mathcal{P}_2(\mathbb{R}^d) \text{ and } \int_0^\tau \|\mathbf{v}_t\|^2_{L^2(\rho_t)} \mathrm{d}t < \infty \tag{2.2}$$

then $\rho_t \in \mathcal{P}_2(\mathbb{R}^d)$ for all $0 < t < \tau$ and $\rho_{(\cdot)}$ is absolutely continuous in the Wasserstein sense.

Remark 2.2. We point out that the hypothesis in Lemma 2.1 requires a priori that the curve $\rho_{(\cdot)}$ lies in $\mathcal{P}_2(\mathbb{R}^d)$, but the proof actually shows that the condition (2.2) implies the whole curve to be in $\mathcal{P}_2(\mathbb{R}^d)$ (and it is absolutely continuous in the Wasserstein sense).

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

2.3. The tangent space

For an absolutely continuous curve in the Wasserstein sense $\rho_{(\cdot)}$ there is a unique Borel field $\mathbf{v}_t \in V(\rho_t) := \overline{\{\nabla f : f \in \mathcal{D}\}}^{L^2(\rho_t)}$ such that the continuity equation (2.1) holds [3], Theorem 8.3.1. This motivates the identification of the tangent space[3] of $\mathcal{P}_2(\mathbb{R}^d)$ at ρ with all $s \in \mathcal{D}'$ for which there exists a $\mathbf{v} \in V(\rho)$ such that

$$s + \operatorname{div}(\rho\,\mathbf{v}) = 0 \qquad \text{in distributional sense.} \tag{2.3}$$

The following inner product on the tangent space at ρ is the metric tensor corresponding to the Wasserstein metric [14]

$$(s_1, s_2)_{-1,\rho} := \int_{\mathbb{R}^d} \mathbf{v}_1 \cdot \mathbf{v}_2 \,\mathrm{d}\rho,$$

where $\mathbf{v}_1$ and $\mathbf{v}_2$ are associated with s_1 and s_2 through (2.3). The corresponding norm coincides with the dual operator norm on $\mathcal{D}'$

$$\|s\|^2_{-1,\rho} := \sup_{f\in\mathcal{D}} \left\{ 2\langle s, f\rangle - \int_{\mathbb{R}^d} |\nabla f|^2 \mathrm{d}\rho \right\}. \tag{2.4}$$

This norm is closely related to the Wasserstein metric through the Benamou–Brenier formula [4]

$$W_2(\rho_0,\rho_1)^2 = \min\left\{ \int_0^1 \|\partial_t\rho_t\|^2_{-1,\rho_t}\,\mathrm{d}t : \rho_t|_{t=0} = \rho_0 \text{ and } \rho_t|_{t=1} = \rho_1 \right\}. \tag{2.5}$$

2.4. Relevant functionals

We sometimes write $\Delta\rho$ and $\operatorname{div}(\rho\nabla\Psi)$ for the functionals in $\mathcal{D}'$ defined by

$$\langle \Delta\rho, f\rangle := \int_{\mathbb{R}^d} \Delta f \,\mathrm{d}\rho \qquad \text{and} \qquad \langle \operatorname{div}(\rho\nabla\Psi), f\rangle := -\int \nabla\Psi\cdot\nabla f \,\mathrm{d}\rho.$$

For $\rho \in \mathcal{P}_2(\mathbb{R}^d)$, we define the *Fisher information*

$$I(\rho) := \begin{cases} \int_{\mathbb{R}^d} \frac{|\nabla\rho(x)|^2}{\rho(x)}\,\mathrm{d}x & \text{if } \rho(\mathrm{d}x) = \rho(x)\,\mathrm{d}x \text{ and } \sqrt{\rho} \in H^1(\mathbb{R}^d), \\ \infty & \text{otherwise,} \end{cases} \tag{2.6}$$

where $\nabla\rho$ is the distributional derivative of ρ. By using (2.4), it is straightforward to see that $\|\Delta\rho\|^2_{-1,\rho} \le I(\rho)$, where the inequality turns to equality when the right hand is finite. Similarly we have that $\|\operatorname{div}(\rho\nabla\Psi)\|^2_{-1,\rho} \le \int |\nabla\Psi|^2\mathrm{d}\rho$. Here equality holds whenever $\int_{\mathbb{R}}|\nabla\Psi|^2\,\mathrm{d}\rho < \infty$, which is certainly true if Ψ satisfies assumptions (4.1) and $\rho \in \mathcal{P}_2(\mathbb{R}^d)$.

Observe that in Theorem 1.1 we assume finiteness of $I(\rho_0)$ and $\int_{\mathbb{R}}|\nabla\Psi|^2\,\mathrm{d}\rho_0$. As a consequence of the HWI inequality [18], Corollary 20.13, these conditions, together with Assumptions 4.1 and 4.4, imply that the free energy $\mathcal{S}(\rho_0) + \mathcal{E}(\rho_0)$ is also finite.

2.5. Chain rule

We conclude this section with a chain rule for the free energy (1.3) on absolutely continuous curves.

Lemma 2.3. *Let $\Psi \in C^2(\mathbb{R}^d)$ be bounded from below and λ-convex for some $\lambda \in \mathbb{R}$ (see [17], Sect. 2.1.3). Assume also that $\rho_{(\cdot)}\colon (0,\tau) \to \mathcal{P}_2(\mathbb{R}^d)$ is an absolutely continuous curve in the Wasserstein sense, that satisfies the conditions $\mathcal{E}(\rho_t), \mathcal{S}(\rho_t) < \infty \quad \forall t \in [0,\tau]$ and*

$$\int_0^\tau \left(\int_{\mathbb{R}^d} |\nabla\Psi(x)|^2 \rho_t(x)\,\mathrm{d}x + I(\rho_t) \right) \mathrm{d}t < \infty. \tag{2.7}$$

[3]Here we like to point out that in [3] the tangent space is identified with the set of velocity fields $V(\rho)$.

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

Then $t \to \mathcal{F}(\rho_t)$ is absolutely continuous and for a.e $t \in [0,\tau]$ we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}(\rho_t) = -\left(\Delta\rho_t + \mathrm{div}(\rho_t \nabla\Psi), \partial_t\rho_t\right)_{-1,\rho_t}. \tag{2.8}$$

Proof. This lemma is a direct consequence of [3], Theorem 10.3.18. All conditions of this theorem are easily checked; the only non-trivial condition may be $\int_0^\tau |\partial\mathcal{F}|(\rho_t)\,|\dot\rho_t|\,\mathrm{d}t < \infty$, where $|\partial\mathcal{F}|(\rho_t)$ is the metric slope and $|\dot\rho_t|$ is the metric derivative (see [3], Sects. 1.1 and 1.2). This is true, since by [3], Theorems 10.4.13 and 8.3.1 and (2.7):

$$\begin{aligned}
\int_0^\tau |\partial\mathcal{F}|(\rho_t)\,|\dot\rho_t|\,\mathrm{d}t &\le \frac{1}{2}\int_0^\tau |\partial\mathcal{F}|^2(\rho_t)\,\mathrm{d}t + \frac{1}{2}\int_0^\tau |\dot\rho_t|^2\,\mathrm{d}t \\
&\le \frac{1}{2}\int_0^\tau\int \left|\nabla\Psi(x) + \tfrac{\nabla\rho_t(x)}{\rho_t(x)}\right|^2 \rho_t(x)\,\mathrm{d}x\,\mathrm{d}t + \frac{1}{2}\int_0^\tau \|\partial_t\rho_t\|^2_{-1,\rho_t}\,\mathrm{d}t \\
&\le \int_0^\tau\int |\nabla\Psi(x)|^2\,\rho_t(x)\,\mathrm{d}x\,\mathrm{d}t + \int_0^\tau I(\rho_t)\,\mathrm{d}t + \frac{1}{2}\int_0^\tau \|\partial_t\rho_t\|^2_{-1,\rho_t}\,\mathrm{d}t < \infty.
\end{aligned}$$

□

3. Particle system and conditional large deviations

In this section we first explain the concept of large deviations with a simple model particle system. Then, we introduce the dynamic particle system that we study more precisely, and discuss the large deviation principle for this system.

Consider a system of N independent random particles in $\mathbb{R}^d$ (without dynamics), where the positions $\mathbf{X}_1,\dots,\mathbf{X}_N$ are identically distributed with law ρ_0. Then, as a consequence of the Law of Large Numbers, the empirical measure $L_N = \frac{1}{N}\sum_{k=1}^N \delta_{\mathbf{X}_k}$ converges almost surely to ρ_0 in the narrow topology as $N\to\infty$ ([8], Thm. 11.4.1). Naturally, this also implies weak convergence:

$$\lim_{N\to\infty} \mathrm{Prob}(L_N \in C) = \delta_{\rho_0}(C)$$

for all narrow continuity sets $C \subset \mathcal{P}(\mathbb{R}^d)$. A large deviation principle quantifies the exponential rate of convergence to 0 (or 1). More precisely, we say *the system satisfies a large deviation principle in $\mathcal{P}(\mathbb{R}^d)$ with (unique) rate $J : \mathcal{P}(\mathbb{R}^d) \to [0,\infty]$* if J is lower semicontinuous, and for all sets $U \subset \mathcal{P}(\mathbb{R}^d)$ there holds (see, for example [9])

$$-\inf_{U^\circ} J \le \liminf_{N\to\infty} \frac{1}{N}\log \mathrm{Prob}(L_N \in U^\circ) \le \limsup_{N\to\infty} \frac{1}{N}\log \mathrm{Prob}(L_N \in \overline{U}) \le -\inf_{\overline{U}} J.$$

In addition, we say a rate functional is *good* if it has compact sub-level sets. By Sanov's Theorem ([9], Thm. 6.2.10), our model example indeed satisfies a large deviation principle, where the good rate functional $J(\rho)$ is the relative entropy

$$\mathcal{H}(\rho|\rho^0) := \begin{cases} \int \log(\frac{\mathrm{d}\rho}{\mathrm{d}\rho^0})\,\mathrm{d}\rho, & \text{if } \rho \ll \rho^0, \\ \infty, & \text{otherwise.} \end{cases}$$

In this example we see the (relative) entropy appearing naturally from a limit of a simple particle system.

Let us now consider our particle system with dynamics, and study its Sanov-type large deviations. To define the system more precisely, let $\mathbf{X}_1(t), \mathbf{X}_2(t), \dots$ be a sequence of independent random processes in $\mathbb{R}^d$. Assume that the initial values are fixed deterministically by some $\mathbf{X}_1(0) = \mathbf{x}_1, \mathbf{X}_2(0) = \mathbf{x}_2, \dots$ in such a way that[4]

$$L_N(0) \to \rho_0 \qquad \text{narrowly for some given } \rho_0 \in \mathcal{P}(\mathbb{R}^d). \tag{3.1}$$

[4]The reason behind this specific initial condition is that we want to somehow condition on $L_N = \rho$, which is a measure-0 set.

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

The evolution of the system is prescribed by the same transition probability for each particle $\mathrm{Prob}(\mathbf{X}_k(t) \in \mathrm{d}\mathbf{y}|\mathbf{X}_k(0) = \mathbf{x}) = p_t(\mathrm{d}\mathbf{y}|\mathbf{x})$. Naturally, for such probability there must hold $p_t(\mathrm{d}\mathbf{y}|\mathbf{x}) \to \delta_{\mathbf{x}}(\mathrm{d}\mathbf{y})$ narrowly as $t \to 0$, and it should evolve according to (1.4). We thus define p_t to be the fundamental solution of (1.4)[5].

Again by the law of large numbers $L_N(\tau) \to \rho_\tau$ almost surely in $\mathcal{P}(\mathbb{R}^d)$, where $\rho_\tau = \rho_0 * p_\tau$, the solution of (1.4) at time τ with initial condition ρ_0. In addition, the empirical measure $L_N(\tau)$ satisfies a large deviation principle

$$\mathrm{Prob}\left(L_N(\tau) \approx \rho\right) \sim \exp\left(-N J_\tau(\rho|\rho_0)\right) \qquad \text{as } N \to \infty$$

with good rate functional (1.5). Observe that $J_\tau(\,\cdot\,|\rho_0) \geq 0$ is minimised by $\rho_0 * p_\tau$.

4. Large deviations of trajectories

In this section we prove, under suitable assumptions for ρ_0 and Ψ, the equivalence of the rate functionals (1.5) and (1.6). The latter form will be used to prove the main Gamma convergence theorem. First, the large deviations of the empirical process is derived. To this aim we will need to distinguish between two different types of potentials Ψ. Next, we transform these large deviation principles back to the large deviations of the empirical measure $L_N(\tau)$ by a contraction principle, and finally show that the resulting rate functionals are the same for both cases.

In the first case we consider potentials that satisfy the following

Assumption 4.1 (The subquadratic case)**.** Let $\Psi \in C^2(\mathbb{R}^d)$ such that

1. Ψ is bounded from below,
2. there is a $C_0 > 0$ such that $|\mathbf{x}||\nabla\Psi(\mathbf{x})| \leq C_0(1 + |\mathbf{x}|^2)$ for all $\mathbf{x} \in \mathbb{R}^d$,
3. Ψ is λ-convex for some $\lambda \in \mathbb{R}$,
4. there exists constants $0 \leq C_1 < \frac{1}{4}$ and $C_2, C_3 \in \mathbb{R}^+$ such that $|\Delta\Psi(\mathbf{x})| \leq C_1|\nabla\Psi(\mathbf{x})|^2 + C_2\Psi(\mathbf{x}) + C_3$.

Note that the second assumption indeed implies $|\Psi(\mathbf{x})| \leq C_0(1 + |\mathbf{x}|^2)$. Under Assumption 4.1, combined with initial condition (3.1), the empirical process $\{L_N(t)\}_{0\leq t\leq\tau}$ satisfies a large deviation principle in $C([0,\tau], \mathcal{P}(\mathbb{R}^d))$ with good rate functional [6], Theorem 4.5

$$\tilde{J}_\tau(\rho_{(\cdot)}) = \frac{1}{4}\int_0^\tau \|\partial_t\rho_t - \Delta\rho_t - \mathrm{div}(\rho_t\nabla\Psi)\|^2_{-1,\rho_t}\mathrm{d}t, \tag{4.1}$$

if the curve $\rho_{(\cdot)}$ is absolutely continuous in the distributional sense; else we set $\tilde{J}_\tau$ to ∞. It follows from a contraction principle [9], Theorem 4.2.1 and a change of variables $t \mapsto t/\tau$ that the conditional rate functional (1.5) can also be written as

$$J_\tau(\rho|\rho_0) = \inf_{\rho_{(\cdot)}\in C(\rho_0,\rho)} \frac{1}{4\tau}\int_0^1 \|\partial_t\rho_t - \tau(\Delta\rho_t + \mathrm{div}(\rho_t\nabla\Psi))\|^2_{-1,\rho_t}\,\mathrm{d}t. \tag{4.2}$$

Remark 4.2. The first assumption guarantees that the functional $\mathcal{E} : \mathcal{P}(\mathbb{R}^d) \to (-\infty, \infty]$ is well defined. The last two assumptions are not necessary to derive (4.1); however we will need them in the sequel.

Remark 4.3. In (4.2) we implicitly set $\frac{1}{4\tau}\int_0^1 \|\partial_t\rho_t - \tau(\Delta\rho_t + \mathrm{div}(\rho_t\nabla\Psi))\|^2_{-1,\rho_t}\,\mathrm{d}t = \infty$ if the curve is not absolutely continuous in distributional sense. Therefore, from now on, we shall only consider curves in $C(\rho_0,\rho)$ or $C_{W_2}(\rho_0,\rho)$ that are absolutely continuous in distributional sense.

[5]Equivalently, we can define the dynamics of $\mathbf{X}_1, \ldots, \mathbf{X}_N$ by the Itō stochastic equations

$$\mathrm{d}\mathbf{X}_k(t) = -\nabla\Psi(\mathbf{X}_k(t))\,\mathrm{d}t + \sqrt{2}\,d\mathbf{W}_k(t), \quad k = 1, \cdots, N$$

where $\mathbf{W}_1, \ldots, \mathbf{W}_N$ are independent Wiener processes.

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

In the second case we require a combination of assumptions on Ψ that were taken from [10,11]:

Assumption 4.4 (The superquadratic case)**.** Let $\Psi \in C^4(\mathbb{R}^d)$ such that:

1. Ψ is λ_Ψ-convex for some $\lambda_\Psi \in \mathbb{R}$;
2. $\int_{\mathbb{R}^d} \Psi(\mathbf{x}) e^{-2\Psi(\mathbf{x})}\, \mathrm{d}\mathbf{x} < \infty$;
3. Ψ has superquadratic growth at infinity, *i.e.* $\lim_{|\mathbf{x}|\to\infty} \frac{\Psi(\mathbf{x})}{|\mathbf{x}|^2} = +\infty$;
4. There exists an $\omega \in C(\mathbb{R}_+ \cup \{0\})$ with $\omega(0) = 0$ and an $\alpha \geq 1$ such that $\lim_{x\to\infty} \frac{\omega(\mathbf{x})}{|\mathbf{x}|^\alpha} \to 0$, and for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$
$$\Psi(\mathbf{y}) - \Psi(\mathbf{x}) \leq \omega(|\mathbf{y}-\mathbf{x}|)(1+\Psi(\mathbf{x})),$$
$$|\Psi(\mathbf{y}) - \Psi(\mathbf{x})|^2 \leq \omega(|\mathbf{y}-\mathbf{x}|)(1+|\nabla\Psi(\mathbf{x})|^2 + \Psi(\mathbf{x}));$$
5. $2\Delta\Psi \leq A|\nabla\Psi|^2 + B$ for some $0 < A < 1$ and $B > 0$;
6. $\zeta := |\nabla\Psi|^2 - 2\Delta\Psi$ has superquadratic growth at infinity, *i.e.* $\lim_{|\mathbf{x}|\to\infty} \frac{\zeta(\mathbf{x})}{|\mathbf{x}|^2} = +\infty$;
7. ζ is λ_ζ-convex for some $\lambda_\zeta \in \mathbb{R}$.

Whenever Assumption 4.4 and initial condition (3.1) hold, then by ([10], Thm. 13.37) the process $\{L_N(t)\}_{0\leq t\leq\tau}$ satisfies a large deviation principle in $C_{W_2}([0,\tau], \mathcal{P}_2(\mathbb{R}^d))$ with good rate functional (4.1).

Remark 4.5. Contrary to the subquadratic case, the latter is actually a large deviation principle on the set of all continuous paths in $\mathcal{P}_2(\mathbb{R}^d)$ with respect to the Wasserstein topology. Although we strongly believe that this is also true for the subquadratic case, it is very difficult to prove due to the fact that the functional $\tilde{J}_\tau$ does not have Wasserstein-compact sub-level sets, and therefore it can not be a good rate functional in $C_{W_2}([0,\tau], \mathcal{P}_2(\mathbb{R}^d))$ when Ψ is subquadratic.

Again, by a contraction principle and a simple change of variables, it follows from (4.1) that (1.5) must be equal to:

$$J_\tau(\rho|\rho_0) = \inf_{\rho_{(\cdot)} \in C_{W_2}(\rho_0,\rho)} \frac{1}{4\tau} \int_0^1 \|\partial_t\rho_t - \tau(\Delta\rho_t + \mathrm{div}(\rho_t\nabla\Psi))\|^2_{-1,\rho_t}\, \mathrm{d}t.$$

Observe that in this case the infimum is taken over Wasserstein-continuous curves, while in the subquadratic case (4.2) the infimum was over narrowly continuous curves. However, we will prove that under the extra assumption that $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$ and $\mathcal{F}(\rho_0)$ is finite, even in the subquadratic case the infimum can be taken over $C_{W_2}(\rho_0,\rho)$. Actually, we will prove something even stronger, that we will need in the sequel, namely the following:

Proposition 4.6. *Let $\Psi \in C^2(\mathbb{R}^d)$ satisfy Assumption 4.1. Let $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$ with $\mathcal{F}(\rho_0) < \infty$ and assume $\rho_{(\cdot)} \in C(\rho_0,\rho)$ with $\tilde{J}_\tau(\rho_{(\cdot)})$ finite. Then we have that $\rho_t \in \mathcal{P}_2(\mathbb{R}^d)$ for every $t \in [0,1]$ and, furthermore, the curve $\rho_{(\cdot)}$ is absolutely continuous in the Wasserstein sense, and $\mathcal{F}(\rho_t)$ is absolutely continuous with respect to t. Finally, there holds:*

$$\frac{1}{4\tau}\int_0^1 \|\partial_t\rho_t - \tau(\Delta\rho_t + \mathrm{div}(\rho_t\nabla\Psi))\|^2_{-1,\rho_t}\, \mathrm{d}t$$
$$= \frac{1}{4\tau}\int_0^1 \|\partial_t\rho_t\|^2_{-1,\rho_t}\, \mathrm{d}t + \frac{\tau}{4}\int_0^1 \|\Delta\rho_t + \mathrm{div}(\rho_t\nabla\Psi)\|^2_{-1,\rho_t}\, \mathrm{d}t + \frac{1}{2}\mathcal{F}(\rho_1) - \frac{1}{2}\mathcal{F}(\rho_0).$$

Before we prove this theorem we prove two auxiliary lemmas.

Lemma 4.7. *Assume that*

1. $\mathcal{E}(\rho_0) < \infty$

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

2. $\Psi \in C^2(\mathbb{R}^d)$ *satisfies Assumption 4.1,*
3. $\rho_{(\cdot)} \in C(\rho_0, \rho)$,
4. $\tilde{J}_\tau(\rho_{(\cdot)}) < \infty$.

Then

$$\int_0^\tau \int_{\mathbb{R}^d} |\nabla \Psi(\mathbf{x})|^2 \rho_t(\mathrm{d}\mathbf{x})\,\mathrm{d}t < \infty. \tag{4.3}$$

Proof. For simplicity we take $\tau = 1$. We will prove the following statement: there exist $0 < \delta \leq 1$ and $\alpha, \beta > 0$ that depend only on Ψ such that

$$\alpha \sup_{t\in[0,\delta]} \int_{\mathbb{R}^d} |\Psi|\,\mathrm{d}\rho_t + \beta \int_0^\delta \int_{\mathbb{R}^d} |\nabla \Psi|^2\,\mathrm{d}\rho_t\,\mathrm{d}t \leq 8\tilde{J}_1(\rho_{(\cdot)}) + 4\,|\inf \Psi| + 2\int_{\mathbb{R}^d} \Psi\,\mathrm{d}\rho_0 + 2\delta C_3. \tag{4.4}$$

Obviously (4.3) follows from (4.4) by repeating it $1/\delta$ times.

By [6], Lemma 4.8, for any $0 \leq s \leq 1$ we have

$$4\tilde{J}_1(\rho_{(\cdot)}) \geq 4\tilde{J}_s(\rho_{(\cdot)}) = \sup_{f\in C_c^2(\mathbb{R}^d)} \int_{\mathbb{R}^d} f\,\mathrm{d}\rho_s - \int_{\mathbb{R}^d} f\,\mathrm{d}\rho_0 - \int_0^s \int_{\mathbb{R}^d} \left(\Delta f - \nabla\Psi\cdot\nabla f + \frac{1}{2}|\nabla f|^2\right)\mathrm{d}\rho_t\,\mathrm{d}t. \tag{4.5}$$

It is worth highlighting that in the above equality, the supremum is taken over $C_c^2(\mathbb{R}^d)$ functions instead of $\mathcal{D}$.

The idea is to use two approximations of Ψ so that it can be chosen as a test function f in (4.5). The first approximation is used to show that this inequality still holds if we take replace $C_c^2(\mathbb{R}^d)$ by

$$A := \left\{ f \in C^2(\mathbb{R}^d) : f, \nabla f, \Delta f, \mathbf{x}f, |\nabla f|\,|\mathbf{x}| \text{ are all bounded} \right\}.$$

Take an arbitrary $f \in A$. Define the bump function

$$\zeta(\mathbf{x}) := \begin{cases} \exp\left(1 - \frac{1}{1-|\mathbf{x}|^2}\right), & |\mathbf{x}| < 1, \\ 0, & |\mathbf{x}| \geq 1, \end{cases}$$

and set $\zeta_k(\mathbf{x}) := \zeta(\mathbf{x}/k)$. Then surely $\zeta_k f \in C_c^2(\mathbb{R}^d)$. It is easy to check that

$$|\zeta_k(\mathbf{x})| \leq 1, \qquad |\nabla\zeta_k(\mathbf{x})| \leq \frac{1}{k} \qquad \text{and} \qquad |\Delta\zeta_k(\mathbf{x})| \leq \frac{1}{k^2}.$$

By the Dominated Convergence Theorem, as $k \to \infty$

$$\begin{aligned}
&\int_{\mathbb{R}^d} \zeta_k f\,\mathrm{d}\rho_s \to \int_{\mathbb{R}^d} f\,\mathrm{d}\rho_s, \\
&\int_{\mathbb{R}^d} \zeta_k f\,\mathrm{d}\rho_0 \to \int_{\mathbb{R}^d} f\,\mathrm{d}\rho_0, \\
&\int_0^s \int_{\mathbb{R}^d} \Delta(\zeta_k f)\,\mathrm{d}\rho_t\,\mathrm{d}t = \int_0^s \int_{\mathbb{R}^d} (f\Delta\zeta_k + 2\nabla\zeta_k\cdot\nabla f + \zeta_k\Delta f)\,\mathrm{d}\rho_t\,\mathrm{d}t \to \int_0^s \int_{\mathbb{R}^d} \Delta f\,\mathrm{d}\rho_t\,\mathrm{d}t, \\
&\int_0^s \int_{\mathbb{R}^d} \nabla\Psi\cdot\nabla(\zeta_k f)\,\mathrm{d}\rho_t\,\mathrm{d}t = \int_0^s \int_{\mathbb{R}^d} \nabla\Psi\cdot(f\nabla\zeta_k + \zeta_k\nabla f)\,\mathrm{d}\rho_t\,\mathrm{d}t \to \int_0^s \int_{\mathbb{R}^d} \nabla\Psi\cdot\nabla f\,\mathrm{d}\rho_t\,\mathrm{d}t, \\
&\int_0^s \int_{\mathbb{R}^d} |\nabla(\zeta_k f)|^2\,\mathrm{d}\rho_t\,\mathrm{d}t = \int_0^s \int_{\mathbb{R}^d} |f\nabla\zeta_k + \zeta_k\nabla f|^2\,\mathrm{d}\rho_t\,\mathrm{d}t \to \int_0^s \int_{\mathbb{R}^d} |\nabla f|^2\,\mathrm{d}\rho_t\,\mathrm{d}t,
\end{aligned}$$

where the absolute finiteness of the right-hand integrals is guaranteed by the properties of the set A. Therefore (4.5) indeed becomes

$$4\tilde{J}_1(\rho_{(\cdot)}) \geq \sup_{f\in A} \int_{\mathbb{R}^d} f\,\mathrm{d}\rho_s - \int_{\mathbb{R}^d} f\,\mathrm{d}\rho_0 - \int_0^s \int_{\mathbb{R}^d} \left(\Delta f - \nabla\Psi\cdot\nabla f + \frac{1}{2}|\nabla f|^2\right)\mathrm{d}\rho_t\,\mathrm{d}t. \tag{4.6}$$

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

For the second approximation we take

$$\eta(\mathbf{x}) := \exp\left(1 - \sqrt{1+|\mathbf{x}|^2}\right),$$

and set $\eta_k(\mathbf{x}) := \eta(\mathbf{x}/k)$. Then the following estimates hold

$$|\eta_k(\mathbf{x})| \le 1, \qquad |\nabla\eta_k(\mathbf{x})| \le \frac{1}{k}\eta_k(\mathbf{x}) \qquad \text{and} \qquad |\Delta\eta_k(\mathbf{x})| \le \frac{1}{k^2}\eta_k(\mathbf{x}). \tag{4.7}$$

Since $\eta_k\Psi \in A$ by the subquadratic Assumption 4.1, we can substitute $\eta_k\Psi$ in (4.6):

$$4\tilde{J}_1(\rho_{(\cdot)}) \ge \int_{\mathbb{R}^d} \eta_k\Psi \,\mathrm{d}\rho_s - \int_{\mathbb{R}^d} \eta_k\Psi \,\mathrm{d}\rho_0 - \int_0^s \int_{\mathbb{R}^d} \Delta(\eta_k\Psi)\,\mathrm{d}\rho_t\,\mathrm{d}t + \int_0^s \int_{\mathbb{R}^d} \left(\nabla\Psi\cdot\nabla(\eta_k\Psi) - \frac{1}{2}|\nabla(\eta_k\Psi)|^2\right)\mathrm{d}\rho_t\,\mathrm{d}t. \tag{4.8}$$

for any $k \in \mathbb{N}$ and $s \in [0,1]$.

We now estimate each term in the right-hand side of (4.8). For the first two terms, we have

$$\int_{\mathbb{R}^d} \eta_k\Psi \,\mathrm{d}\rho_s - \int_{\mathbb{R}^d} \eta_k\Psi \,\mathrm{d}\rho_0 \ge \int_{\mathbb{R}^d} \eta_k|\Psi| \,\mathrm{d}\rho_s - 2\,|\inf\Psi| - \int_{\mathbb{R}^d} \eta_k\Psi \,\mathrm{d}\rho_0. \tag{4.9}$$

For the third term of (4.8), we find

$$\begin{aligned}
-\int_0^s \int_{\mathbb{R}^d} \Delta(\eta_k\Psi)\,\mathrm{d}\rho_t\,\mathrm{d}t &= -\int_0^s \int_{\mathbb{R}^d} (\Psi\Delta\eta_k + 2\nabla\eta_k\cdot\nabla\Psi + \eta_k\Delta\Psi)\,\mathrm{d}\rho_t\,\mathrm{d}t \\
&\ge -\int_0^s \int_{\mathbb{R}^d} \left(|\Delta\eta_k|\,|\Psi| + |\nabla\eta_k|\,(|\nabla\Psi|^2+1) + \eta_k|\Delta\Psi|\right)\mathrm{d}\rho_t\,\mathrm{d}t \\
&\overset{(4.7)}{\ge} -\int_0^s \int_{\mathbb{R}^d} \left(\frac{1}{k^2}\eta_k\Psi + \frac{\eta_k}{k}(|\nabla\Psi|^2+1) + \eta_k|\Delta\Psi|\right)\mathrm{d}\rho_t\,\mathrm{d}t \\
&\overset{\text{Ass.4.1(4)}}{\ge} -\int_0^s \int_{\mathbb{R}^d} \left(\frac{1}{k^2}\eta_k\Psi + \frac{\eta_k}{k}(|\nabla\Psi|^2+1) + \eta_k\left(C_1|\nabla\Psi|^2 + C_2|\Psi| + C_3\right)\right)\mathrm{d}\rho_t\,\mathrm{d}t \\
&\ge -s\left(\frac{1}{k^2} + C_2\right)\sup_{t\in[0,s]}\int_{\mathbb{R}^d}\eta_k|\Psi|\,\mathrm{d}\rho_t - \left(\frac{1}{k} + C_1\right)\int_0^s\int_{\mathbb{R}^d}\eta_k|\nabla\Psi|^2\,\mathrm{d}\rho_t\,\mathrm{d}t - \frac{s}{k} - sC_3.
\end{aligned} \tag{4.10}$$

Finally, for the last part of (4.8)

$$\begin{aligned}
&\int_0^s \int_{\mathbb{R}^d} \left(\nabla\Psi\cdot\nabla(\eta_k\Psi) - \frac{1}{2}|\nabla(\eta_k\Psi)|^2\right)\mathrm{d}\rho_t\,\mathrm{d}t \\
&\quad= \int_0^s \int_{\mathbb{R}^d} \left(-\frac{1}{2}|\nabla\eta_k|^2\Psi^2 + (1-\eta_k)\nabla\eta_k\cdot\Psi\nabla\Psi + (1-\frac{1}{2}\eta_k)\eta_k|\nabla\Psi|^2\right)\mathrm{d}\rho_t\,\mathrm{d}t \\
&\quad\overset{(4.7)}{\ge} \int_0^s \int_{\mathbb{R}^d} \left(-\frac{1}{2k^2}\eta_k\Psi^2 - 2\eta_k\left|\frac{1}{k}\Psi\right|\left|\frac{1}{2}\nabla\Psi\right| + \frac{3}{4}\eta_k|\nabla\Psi|^2\right)\mathrm{d}\rho_t\,\mathrm{d}t, \\
&\quad\ge \int_0^s \int_{\mathbb{R}^d} \left(-\frac{3}{2k^2}\eta_k\Psi^2 + \left(\frac{3}{4} - \frac{1}{4}\right)\eta_k|\nabla\Psi|^2\right)\mathrm{d}\rho_t\,\mathrm{d}t. \\
&\quad\ge \int_0^s \int_{\mathbb{R}^d} \left(-\frac{3C_0(1+k^2)}{2k^2}\eta_k|\Psi| + \frac{1}{2}\eta_k|\nabla\Psi|^2\right)\mathrm{d}\rho_t\,\mathrm{d}t \\
&\quad\ge -\frac{3sC_0(1+k^2)}{2k^2}\sup_{t\in[0,s]}\int_{\mathbb{R}^d}\eta_k|\Psi|\,\mathrm{d}\rho_t + \int_0^s\int_{\mathbb{R}^d}\left(\frac{1}{2}\eta_k|\nabla\Psi|^2\right)\mathrm{d}\rho_t\,\mathrm{d}t,
\end{aligned} \tag{4.11}$$

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

where the fourth line follows from Young's inequality, and in the fifth line we used subquadratic Assumption 4.1(2). Substituting (4.9), (4.10) and (4.11) into (4.8) we get

$$\int_{\mathbb{R}^d} \eta_k|\Psi|\,\mathrm{d}\rho_s + \int_0^s \int_{\mathbb{R}^d} \frac{1}{2}\eta_k|\nabla\Psi|^2\,\mathrm{d}\rho_t\,\mathrm{d}t \le 4\tilde{J}_1(\rho_{(\cdot)}) + 2\,|\inf\Psi| + \int_{\mathbb{R}^d}\eta_k\Psi\,\mathrm{d}\rho_0 + \frac{s}{k} + sC_3$$
$$+ s\left(\frac{1}{k^2} + C_2 + \frac{3C_0(1+k^2)}{2k^2}\right)\sup_{t\in[0,s]}\int_{\mathbb{R}^d}\eta_k|\Psi|\,\mathrm{d}\rho_t + \left(\frac{1}{k} + C_1\right)\int_0^s\int_{\mathbb{R}^d}\eta_k|\nabla\Psi|^2\,\mathrm{d}\rho_t\,\mathrm{d}t.$$

If we first discard the first term on the left-hand side and maximise the equation over $s \in [0,\delta]$ for some $0 < \delta \le 1$, then discard the second term and maximise, the sum of the inequalities can be written as

$$\left(1 - 2\delta\left(\frac{1}{k^2} + C_2 + \frac{3C_0(1+k^2)}{2k^2}\right)\right)\sup_{t\in[0,\delta]}\int_{\mathbb{R}^d}\eta_k|\Psi|\,\mathrm{d}\rho_t + \left(\frac{1}{2} - \frac{2}{k} - 2C_1\right)\int_0^\delta\int_{\mathbb{R}^d}|\nabla\Psi|^2\,\mathrm{d}\rho_t\,\mathrm{d}t$$
$$\le 8\tilde{J}_1(\rho_{(\cdot)}) + 4\,|\inf\Psi| + 2\int_{\mathbb{R}^d}\eta_k\Psi\,\mathrm{d}\rho_0 + \frac{2\delta}{k} + 2\delta C_3.$$

For δ such that $1 > 2\delta\left(C_2 + \frac{3C_0}{2}\right)$, we get $1 > 2\delta\left(\frac{1}{k^2} + C_2 + \frac{3C_0(1+k^2)}{2k^2}\right)$ for suffiently large k, and therefore from Fatou's Lemma

$$\alpha\sup_{t\in[0,\delta]}\int_{\mathbb{R}^d}|\Psi|\,\mathrm{d}\rho_t + \beta\int_0^\delta\int_{\mathbb{R}^d}|\nabla\Psi|^2\,\mathrm{d}\rho_t\,\mathrm{d}t \le 8\tilde{J}_1(\rho_{(\cdot)}) + 4\,|\inf\Psi| + 2\int_{\mathbb{R}^d}\Psi\,\mathrm{d}\rho_0 + 2\delta C_3,$$

with $\alpha := 1 - 2\delta(C_2 + \frac{3C_0}{2}) > 0$ and $\beta := \frac{1}{2} - 2C_1$. The latter is positive by Assumption 4.1(4). □

The second auxiliary lemma is:

Lemma 4.8. *Let $\epsilon > 0$ and $\rho(\mathbf{x})\,\mathrm{d}\mathbf{x} \in \mathcal{P}(\mathbb{R}^d)$ be given. Let $\theta(\mathbf{x}) := \left(\frac{1}{2\pi}\right)^{\frac{d}{2}} e^{\frac{-|\mathbf{x}|^2}{2}}$ be the density of the d-dimensional normal distribution. We define $\theta_\epsilon(\mathbf{x}) := \epsilon^{-d}\theta(\frac{\mathbf{x}}{\epsilon})$ and $\rho_\epsilon := \rho * \theta_\epsilon$. Then there exists a constant C_ϵ that depends only on ϵ such that $I(\rho_\epsilon) < C_\epsilon$.*

Proof. We have

$$\nabla\rho_\epsilon(\mathbf{x}) = (\rho * \nabla\theta_\epsilon)(\mathbf{x}) = \int_{\mathbb{R}^d}\rho(\mathbf{x}-\mathbf{y})\nabla\theta_\epsilon(\mathbf{y})\,\mathrm{d}\mathbf{y} = -\epsilon^{-2}\int_{\mathbb{R}^d}\rho(\mathbf{x}-\mathbf{y})\mathbf{y}\theta_\epsilon(\mathbf{y})\,\mathrm{d}\mathbf{y}.$$

Furthermore

$$|\nabla\rho_\epsilon(\mathbf{x})|^2 \le \epsilon^{-4}\int_{\mathbb{R}^d}\rho(\mathbf{x}-\mathbf{y})|\mathbf{y}|^2\theta_\epsilon(\mathbf{y})\,\mathrm{d}\mathbf{y}\int_{\mathbb{R}^d}\rho(\mathbf{x}-\mathbf{y})\theta_\epsilon(\mathbf{y})\,\mathrm{d}\mathbf{y} \le \epsilon^{-4}\rho_\epsilon(\mathbf{x})\int_{\mathbb{R}^d}\rho(\mathbf{x}-\mathbf{y})|\mathbf{y}|^2\theta_\epsilon(\mathbf{y})\,\mathrm{d}\mathbf{y}.$$

Now

$$I(\rho_\epsilon) = \int_{\mathbb{R}^d}\frac{|\nabla\rho_\epsilon(\mathbf{x})|^2}{\rho_\epsilon(\mathbf{x})}\,\mathrm{d}\mathbf{x} \le \epsilon^{-4}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\rho(\mathbf{x}-\mathbf{y})|\mathbf{y}|^2\theta_\epsilon(\mathbf{y})\,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{x}$$
$$= \epsilon^{-4}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\rho(\mathbf{x}-\mathbf{y})\,\mathrm{d}\mathbf{x}\,|\mathbf{y}|^2\theta_\epsilon(\mathbf{y})\,\mathrm{d}\mathbf{y}$$
$$\le \epsilon^{-4}\int_{\mathbb{R}^d}|\mathbf{y}|^2\theta_\epsilon(\mathbf{y})\,\mathrm{d}\mathbf{y} =: C_\epsilon.$$

□

We are now ready to proceed with the

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

Proof of Proposition 4.6. Let $\rho_{(\cdot)}$ satisfy the assumptions (of Prop. 4.6). By Lemma 4.7 we have

$$\int_0^1 \int_{\mathbb{R}^d} |\nabla \Psi(\mathbf{x})|^2 \rho_t(\mathrm{d}\mathbf{x})\, \mathrm{d}t < \infty$$

and therefore

$$\frac{1}{4\tau}\int_0^1 \|\partial_t \rho_t - \tau \Delta \rho_t\|^2_{-1,\rho_t}\, \mathrm{d}t \le \frac{1}{2\tau}\int_0^1 \|\partial_t \rho_t - \tau(\Delta \rho_t + \operatorname{div}(\rho_t \nabla \Psi))\|^2_{-1,\rho_t}\, \mathrm{d}t + \frac{\tau}{2}\int_0^1 \int_{\mathbb{R}^d} |\nabla \Psi|^2 \rho_t(\mathrm{d}\mathbf{x})\, \mathrm{d}t < \infty.$$

Take a $0 < s \le 1$. Since

$$\frac{1}{4\tau}\int_0^s \|\partial_t \rho_t - \tau \Delta \rho_t\|^2_{-1,\rho_t}\, \mathrm{d}t < \infty$$

we have that $\|\partial_t \rho_t - \tau \Delta \rho_t\|^2_{-1,\rho_t} < \infty$ for almost every t. By ([10], Lem. D.34) there is a $\mathbf{v}_t \in L^2(\rho_t)$ such that

$$\partial_t \rho_t - \tau \Delta \rho_t = -\operatorname{div}(\mathbf{v}_t\, \rho_t)$$

in distributional sense. Take the Gaussian $\theta_\epsilon(\mathbf{x})$ as in Lemma 4.8. Then we have

$$\partial_t \rho_{t,\epsilon} - \tau \Delta \rho_{t,\epsilon} = -\operatorname{div}(\mathbf{v}_{t,\epsilon}\, \rho_{t,\epsilon}),$$

where

$$\rho_{t,\epsilon} = \rho_t * \theta_\epsilon(\mathbf{x}), \qquad \mathbf{v}_{t,\epsilon} = \frac{(\mathbf{v}_t\, \rho_t) * \theta_\epsilon(\mathbf{x})}{\rho_{t,\epsilon}}.$$

By ([3], Thm. 8.1.9) we have

$$\frac{1}{4\tau}\int_0^s \|\partial_t \rho_{t,\epsilon} - \tau \Delta \rho_{t,\epsilon}\|^2_{-1,\rho_t}\, \mathrm{d}t \le \frac{1}{4\tau}\int_0^s \|\mathbf{v}_{t,\epsilon}\|^2_{L^2(\rho_{t,\epsilon})}\, \mathrm{d}t \le \frac{1}{4\tau}\int_0^s \|\mathbf{v}_t\|^2_{L^2(\rho_t)}\, \mathrm{d}t = \frac{1}{4\tau}\int_0^s \|\partial_t \rho_t - \tau \Delta \rho_t\|^2_{-1,\rho_t}\, \mathrm{d}t. \quad (4.12)$$

Furthermore by Lemma 4.8 we have that

$$\int_0^s \|\Delta \rho_{t,\epsilon}\|^2_{-1,\rho_{t,\epsilon}}\, \mathrm{d}t = \int_0^s I(\rho_{t,\epsilon})\, \mathrm{d}t \le C_\epsilon, \quad (4.13)$$

and therefore

$$\int_0^s \|\partial_t \rho_{t,\epsilon}\|^2_{-1,\rho_{t,\epsilon}}\, \mathrm{d}t < \infty. \quad (4.14)$$

From (4.13) and since $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$, by using ([10], Lem. D.34) and Lemma 2.1 we get that the curve $\rho_{t,\epsilon}$ is absolutely continuous in $\mathcal{P}_2(\mathbb{R}^d)$. In addition, it is straightforward that $\mathcal{S}(\rho_{t,\epsilon})$ is finite for every $0 < t \le s$. From (4.13), (4.14) and by Lemma 2.3, $\mathcal{S}(\rho_{t,\epsilon})$ is absolutely continuous with respect to t. Hence we obtain

$$\begin{aligned}
&\frac{1}{4\tau}\int_0^s \|\partial_t \rho_{t,\epsilon} - \tau \Delta \rho_{t,\epsilon}\|^2_{-1,\rho_t}\mathrm{d}t \\
&= \frac{1}{4\tau}\int_0^s \|\partial_t \rho_{t,\epsilon}\|^2_{-1,\rho_t}\mathrm{d}t + \frac{\tau}{4}\int_0^s \|\Delta \rho_{t,\epsilon}\|^2_{-1,\rho_t}\mathrm{d}t - \frac{1}{2}\int_0^s (\Delta \rho_{t,\epsilon}, \partial_t \rho_{t,\epsilon})_{-1,\rho_t}\, \mathrm{d}t \\
&= \frac{1}{4\tau}\int_0^s \|\partial_t \rho_{t,\epsilon}\|^2_{-1,\rho_t}\mathrm{d}t + \frac{\tau}{4}\int_0^s \|\Delta \rho_{t,\epsilon}\|^2_{-1,\rho_t} + \frac{1}{2}\mathcal{S}(\rho_{s,\epsilon}) - \frac{1}{2}\mathcal{S}(\rho_{0,\epsilon}).
\end{aligned}$$

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

It follows from this and (4.12) that

$$\frac{1}{4\tau}\int_0^s \|\partial_t\rho_{t,\epsilon}\|^2_{-1,\rho_t}\mathrm{d}t + \frac{\tau}{4}\int_0^s \|\Delta\rho_{t,\epsilon}\|^2_{-1,\rho_t}\mathrm{d}t + \frac{1}{2}\mathcal{S}(\rho_{s,\epsilon}) - \frac{1}{2}\mathcal{S}(\rho_{0,\epsilon}) \le \frac{1}{4\tau}\int_0^1 \|\partial_t\rho_t - \tau\Delta\rho_t\|^2_{-1,\rho_t}\mathrm{d}t.$$

Now letting ϵ go to zero and by the lower semicontinuity of the entropy and the Fisher information functionals we get $\mathcal{S}(\rho_s) < \infty$ and $\int_0^s \|\Delta\rho_t\|^2_{-1,\rho_t}\,\mathrm{d}t < \infty$. Therefore

$$\int_0^s \|\partial_t\rho_t\|^2_{-1,\rho_t}\mathrm{d}t \le 2\left(\int_0^s \|\partial_t\rho_t - \tau\Delta\rho_t\|^2_{-1,\rho_t}\mathrm{d}t + \tau^2\int_0^s \|\Delta\rho_t\|^2_{-1,\rho_t}\,\mathrm{d}t\right) < \infty.$$

and

$$\int_0^s \|\Delta\rho_t + \operatorname{div}\rho_t\nabla\Psi\|^2_{-1,\rho_t}\mathrm{d}t \le 2\left(\int_0^s \|\Delta\rho_t\|^2_{-1,\rho_t}\,\mathrm{d}t + \int_0^s\int_{\mathbb{R}^d} |\nabla\Psi(\mathbf{x})|^2\rho_t(\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}t\right) < \infty.$$

By Lemma 2.1, the curve ρ_t is in $AC_{W_2}\left([0,1];\mathcal{P}_2(\mathbb{R}^d)\right)$. Moreover, $t \mapsto \mathcal{F}(\rho_t)$ is absolutely continuous and (2.8) holds. Hence we have

$$\begin{aligned}\frac{1}{4\tau}\int_0^1 \|\partial_t\rho_t - \tau(\Delta\rho_t + \operatorname{div}(\rho_t\nabla\Psi))\|^2_{-1,\rho_t}\,\mathrm{d}t \\ = \frac{1}{4\tau}\int_0^1 \|\partial_t\rho_t\|^2_{-1,\rho_t}\,\mathrm{d}t + \frac{\tau}{4}\int_0^1 \|\Delta\rho_t + \operatorname{div}(\rho_t\nabla\Psi))\|^2_{-1,\rho_t}\,\mathrm{d}t + \frac{1}{2}\mathcal{F}(\rho_1) - \frac{1}{2}\mathcal{F}(\rho_0).\end{aligned}$$

This finishes the proof of the proposition. □

Remark 4.9. For the superquadratic case, the above Proposition was proved by Feng and Nguyen in [11] by using probabilistic tools. In addition, they obtain an estimate for the growth of $\mathcal{F}$ along the curves.

Now the following is a straightforward result:

Corollary 4.10. *Let $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$ with $\mathcal{F}(\rho_0) < \infty$. If $\Psi \in C^2(\mathbb{R}^d)$ satisfies either Assumption 4.1 or 4.4, then*

$$J_\tau(\rho|\rho_0) = \inf_{\rho_{(\cdot)}\in C_{W_2}(\rho_0,\rho)} \frac{1}{4\tau}\int_0^1 \|\partial_t\rho_t - \tau(\Delta\rho_t + \operatorname{div}(\rho_t\nabla\Psi))\|^2_{-1,\rho_t}\,\mathrm{d}t.$$

5. Lower bound

In this section we prove the lower bound of the Gamma convergence (1.7) in our main result, Theorem 1.1.

Theorem 5.1 (Lower bound)**.** *Under the assumptions of Theorem 1.1, we have for any $\rho_1 \in \mathcal{P}_2(\mathbb{R}^d)$ and all sequences $\rho_1^\tau \in \mathcal{P}_2(\mathbb{R}^d)$ narrowly converging to ρ_1*

$$\liminf_{\tau\to 0}\left(J_\tau(\rho_1^\tau|\rho_0) - \frac{W_2^2(\rho_0,\rho_1^\tau)}{4\tau}\right) \ge \frac{1}{2}\mathcal{F}(\rho_1) - \frac{1}{2}\mathcal{F}(\rho_0). \tag{5.1}$$

Proof. Take any sequence $\rho_1^\tau \in \mathcal{P}_2(\mathbb{R}^d)$ narrowly converging to a $\rho_1 \in \mathcal{P}_2(\mathbb{R}^d)$. We only need to consider those ρ_1^τ for which $J_\tau(\rho_1^\tau|\rho_0) < \infty$. For each such ρ_1^τ, by the definition of infimum there exists a curve $\rho_t^\tau \in C(\rho_0,\rho_1^\tau)$ satisfying

$$\frac{1}{4\tau}\int_0^1 \left\|\partial_t\rho_t^\tau - \tau(\Delta\rho_t^\tau + \operatorname{div}(\rho_t^\tau\,\nabla\Psi))\right\|^2_{-1,\rho_t^\tau}\mathrm{d}t \le J_\tau(\rho_1^\tau|\rho_0) + \tau < \infty.$$

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

By Proposition 4.6 for the subquadratic case and [11], Lemma 2.6 for the superquadratic case, we have

$$\begin{aligned}
J_\tau(\rho_1^\tau|\rho_0)+\tau &\geq \frac{1}{4\tau}\int_0^1 \left\|\partial_t\rho_t^\tau - \tau(\Delta\rho_t^\tau + \operatorname{div}(\rho_t^\tau\,\nabla\Psi))\right\|^2_{-1,\rho_t^\tau}\,\mathrm{d}t\\
&= \frac{1}{4\tau}\left(\int_0^1 \left\|\partial_t\rho_t^\tau\right\|^2_{-1,\rho_t^\tau}\,\mathrm{d}t + 2\tau(\mathcal{F}(\rho_1^\tau)-\mathcal{F}(\rho_0)) + \tau^2\int_0^1 \|\Delta\rho_t^\tau + \operatorname{div}(\rho_t^\tau\,\nabla\Psi)\|^2_{-1,\rho_t^\tau}\,\mathrm{d}t\right)\\
&= \frac{1}{2}(\mathcal{F}(\rho_1^\tau)-\mathcal{F}(\rho_0)) + \frac{1}{4\tau}\int_0^1 \left\|\partial_t\rho_t^\tau\right\|^2_{-1,\rho_t^\tau}\,\mathrm{d}t + \frac{\tau}{4}\int_0^1 \|\Delta\rho_t^\tau + \operatorname{div}(\rho_t^\tau\,\nabla\Psi)\|^2_{-1,\rho_t^\tau}\,\mathrm{d}t\\
&\geq \frac{1}{2}(\mathcal{F}(\rho_1^\tau)-\mathcal{F}(\rho_0)) + \frac{1}{4\tau}\int_0^1 \left\|\partial_t\rho_t^\tau\right\|^2_{-1,\rho_t^\tau}\,\mathrm{d}t\\
&\geq \frac{1}{2}(\mathcal{F}(\rho_1^\tau)-\mathcal{F}(\rho_0)) + \frac{1}{4\tau}W_2^2(\rho_0,\rho_1^\tau).
\end{aligned}$$

In the last inequality above we have used the Benamou–Brenier formula (2.5) for the Wasserstein distance. Finally, using $\rho_1^\tau \to \rho_1$ narrowly with the narrow lower semi-continuity of $\mathcal{F}$, we find that

$$\liminf_{\tau\to 0}\left(J_\tau(\rho_1^\tau|\rho_0) - \frac{W_2^2(\rho_0,\rho_1^\tau)}{4\tau}\right) \geq \frac{1}{2}\mathcal{F}(\rho_1) - \frac{1}{2}\mathcal{F}(\rho_0). \qquad \square$$

6. Recovery sequence

In this section we prove the upper bound of the Gamma convergence (1.7). This will conclude the proof of Theorem 1.1.

Theorem 6.1 (Recovery sequence)**.** *Under the assumptions of Theorem 1.1, for any $\rho_1 \in \mathcal{P}_2(\mathbb{R})$ there exists a sequence $\rho_1^\tau \in \mathcal{P}_2(\mathbb{R})$ converging to ρ_1 in the Wasserstein metric such that*

$$\limsup_{\tau\to 0}\left(J_\tau(\rho_1^\tau|\rho_0) - \frac{W_2^2(\rho_0,\rho_1^\tau)}{4\tau}\right) \leq \frac{1}{2}S(\rho_1) - \frac{1}{2}S(\rho_0).$$

As mentioned in Section 1, our approach for the recovery sequence only works for $d = 1$. Hence throughout this section, we will consider $d = 1$.

The existence of the recovery sequence is proven by making use of the following denseness argument, which is also interesting in its own[6]:

Proposition 6.2. *Let (X, d) be a metric space and let Q be a dense subset of X. If $\{K_n, n \in \mathbb{N}\}$ and K_∞ are functions from X to $\mathbb{R}$ such that:*

(a) $K_n(q) \to K_\infty(q)$ for all $q \in Q$,
(b) for every $x \in X$ there exists a sequence $q_n \in Q$ with $q_n \to x$ and $K_\infty(q_n) \to K_\infty(x)$,

then for every $x \in X$ there exists a sequence $r_n \in Q$, with $r_n \to x$ such that $K_n(r_n) \to K_\infty(x)$.

Proof. The proof is by a diagonal argument. Take any $x \in X$ and take the corresponding sequence $q_n \to x$ such that $K_\infty(q_n) \to K_\infty(x)$. By assumption, for any $q \in Q$ and $L > 0$ there exists a $n_{L,q}$ such that for any $n \geq n_{L,q}$ there holds $d(K_n(q), K_\infty(q)) < 1/L$. Define

$$l_n := \begin{cases} 1, & 1 \leq n < n_{2,q_2},\\ 2, & n_{2,q_2} \leq n < \max\{n_{2,q_2}, n_{3,q_3}\},\\ \dots \end{cases}$$

[6] A more or less similar idea can be found in ([5], Rem. 1.29); Proposition 6.2 is slightly stronger.

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

Take the subsequence $r_n := q_{l_n}$. Observe that $l_n \to \infty$ as $n \to \infty$ such that indeed $q_{l_n} \to x$, and:

$$d(K_n(q_{l_n}), K_\infty(x)) \leq \underbrace{d(K_n(q_{l_n}), K_\infty(q_{l_n}))}_{\leq \frac{1}{l_n}} + d(K_\infty(q_{l_n}), K_\infty(x)) \to 0. \qquad \square$$

For a fixed ρ_0 satisfying the assumptions of Theorem 1.1, we want to apply Proposition 6.2 to the situation where

$$X = \mathcal{P}_2(\mathbb{R}),$$
$$Q = Q(\rho_0) = \Big\{ \rho(x)\,\mathrm{d}x \in \mathcal{P}_2(\mathbb{R}) : \rho(x) \text{ is bounded from below by a positive constant in every compact set,}$$
$$I(\rho), \int_{\mathbb{R}} |\Psi'(x)|^2 \rho(x)\,\mathrm{d}x < \infty, \text{ and there exists a } M > 0 \text{ such that } \rho_0(x) = \rho(x) \text{ for all } |x| > M \Big\},$$
$$K_n(\rho) = J_{\tau_n}(\rho\,|\rho_0) - \frac{W_2^2(\rho_0, \rho)}{4\tau_n}, \text{ for an arbitrary sequence } \tau_n \text{ converging to zero,}$$
$$K_\infty(\rho) = \frac{1}{2}\mathcal{F}(\rho) - \frac{1}{2}\mathcal{F}(\rho_0).$$

Assumption *(a)* of Proposition 6.2, *i.e.* pointwise convergence for every $\rho_1 \in Q(\rho_0)$, can be proven as follows. Take $\rho_1 \in Q(\rho_0)$ and let ρ_t be the geodesic that connects ρ_0 and ρ_1. In the following Lemma 6.3, we will prove that $I(\rho_t)$ and $\int_{\mathbb{R}} |\Psi'(x)|^2 \rho_t(x)\mathrm{d}x$ are uniformly bounded, so that we have

$$\int_0^1 \|\partial_t \rho_t - \tau(\partial_{xx}\rho_t + \partial_x(\rho_t \Psi'))\|^2_{-1,\rho_t}\,\mathrm{d}t \leq 3\int_0^1 \|\partial_t \rho_t\|^2_{-1,\rho_t}\,\mathrm{d}t + 3\tau^2 \int_0^1 \|\partial_{xx}\rho_t\|^2_{-1,\rho_t}\,\mathrm{d}t + 3\tau^2 \int_0^1 \|\partial_x(\rho_t\Psi')\|^2_{-1,\rho_t}\,\mathrm{d}t < \infty.$$

By Proposition 4.6 for the subquadratic case or [11], Lemma 2.6 for the superquadratic case, together with Young's inequality:

$$\lim_{\tau\to 0}\left(J_\tau(\rho_1|\rho_0) - \frac{W_2^2(\rho_0,\rho_1)}{4\tau}\right) \leq \lim_{\tau\to 0}\left[\frac{\tau}{2}\int_0^1 \left(\int_{\mathbb{R}} \left(\frac{(\rho_t'(x))^2}{\rho_t(x)} + |\Psi'(x)|^2\rho_t(x)\right)\mathrm{d}x\right)\mathrm{d}t + \frac{1}{2}\mathcal{F}(\rho_1) - \frac{1}{2}\mathcal{F}(\rho_0)\right] = \frac{1}{2}\mathcal{F}(\rho_1) - \frac{1}{2}\mathcal{F}(\rho_0).$$

The pointwise convergence then follows from this together with the lower bound (5.1).

To prove the uniform bounds:

Lemma 6.3. *Let $\Psi \in C^2(\mathbb{R})$ with $\Psi(x) > -A - B|x|^2$ for some positive constants (this includes both our cases). Let $\rho_0 = \rho_0(x)\mathrm{d}x \in \mathcal{P}_2(\mathbb{R})$ be absolutely continuous with respect to the Lesbegue measure, where $\rho_0(x)$ is bounded from below by a positive constant in every compact set. Let $\rho_1 \in Q(\rho_0)$ and ρ_t be the geodesic that connects ρ_0 and ρ_1. Assume that $\mathcal{E}(\rho_0)$, $I(\rho_0)$ and $\int_{\mathbb{R}} |\Psi'(x)|^2\rho_0(x)\mathrm{d}x$ are all finite. Then $\mathcal{F}(\rho_t)$, $I(\rho_t)$ and $\int_{\mathbb{R}} |\Psi'(x)|^2\rho_t(x)\,\mathrm{d}x$ are uniformly bounded with respect to t.*

Proof. Let $T(x)$ be the optimal map that transports $\rho_0(\mathrm{d}x)$ to $\rho_1(\mathrm{d}x)$. The geodesic that connects ρ_0 and ρ_1 is defined by

$$\rho_t(x) = ((1-t)x + tT(x))_\sharp \rho_0(x).$$

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

First we prove that $I(\rho_t)$ is uniformly bounded with respect to t. In the real line, the map $T(x)$ can be determined *via* the cumulative distribution functions as follows ([17], Sect. 2.2). Let $F(x)$ and $G(x)$ be respectively the cumulative distribution functions of $\rho_0(\mathrm{d}x)$ and $\rho_1(\mathrm{d}x)$, *i.e.*

$$F(x) = \int_{-\infty}^{x} \rho_0(x)\,\mathrm{d}x; \quad G(x) = \int_{-\infty}^{x} \rho_1(x)\,\mathrm{d}x.$$

Then $T = G^{-1} \circ F$. We have

$$F(M) + \int_M^{+\infty} \rho_0(x)\,\mathrm{d}x = G(M) + \int_M^{+\infty} \rho_1(x)\,\mathrm{d}x = 1. \tag{6.1}$$

From (6.1) and by the assumption that $\rho_0(x) = \rho_1(x)$ for all $|x| > M$ we find that $F(M) = G(M)$. Hence for all x such that $|x| > M$ we have

$$F(x) = F(M) + \int_M^{x} \rho_0(x)\,\mathrm{d}x = G(M) + \int_M^{x} \rho_1(x)\,\mathrm{d}x = G(x).$$

Consequently, for all x with $|x| > M$ we have $T(x) = (G^{-1} \circ F)(x) = x$. Therefore $T'(x) = 1$ for all $|x| > M$. Also since the densities ρ_0, ρ_1 are absolutely continuous (by assumption) we get that $F(x), G(x)$ are differentiable everywhere with $G'(x) = \rho_1(x) > 0$. We deduce that $T(x)$ has a classical derivative everywhere and moreover since $G(T(x)) = F(x)$, by differentiating we get that $T(x)$ satisfies the Monge–Ampère equation.

$$\rho_0(x) = \rho_1(T(x))T'(x).$$

or equivalently (since $\rho_1(x) > 0$),

$$T'(x) = \frac{\rho_0(x)}{\rho_1(T(x))}. \tag{6.2}$$

Because of (6.2) we have that $T'(x)$ is absolutely continuous and strictly positive. Therefore the derivative of T' exists almost everywhere. Now for the derivative of T' we have

$$\begin{aligned}
\frac{T''(x)}{T'(x)} &= (\log(T'(x)))' \\
&= (\log(\rho_0(x)) - \log(\rho_1(T(x)))' \\
&= \frac{\rho_0'(x)}{\rho_0(x)} - \frac{\rho_1'(T(x))T'(x)}{\rho_1(T(x))}.
\end{aligned}$$

Set $T_t(x) = tx + (1-t)T(x)$. For $0 \le t \le 1$ we have

$$\rho_t(x) = \rho_1(T_t(x))T_t'(x), \tag{6.3}$$

Since $\rho_1(T_t(x))$ and $T_t'(x)$ are both absolutely continuous so is $\rho_t(x)$. Hence the derivative appeared in (2.6) for $I(\rho_t)$ is the classical derivative. Substituting (6.3) into (2.6) we get

$$\begin{aligned}
\int_{\mathbb{R}} \frac{(\rho_t'(x))^2}{\rho_t(x)}\,\mathrm{d}x &= \int_{\mathbb{R}} \frac{[(\rho_1(T_t(x))T_t'(x))']^2}{\rho_1(T_t(x))T_t'(x)}\,\mathrm{d}x \\
&= \int_{\mathbb{R}} \frac{[\rho_1'(T_t(x))T_t'(x)^2 + \rho_1(T_t(x))T_t''(x)]^2}{\rho_1(T_t(x))T_t'(x)}\,\mathrm{d}x \\
&\le 2\int_{\mathbb{R}} \frac{(\rho_1'(T_t(x)))^2(T_t'(x))^4}{\rho_1(T_t(x))T_t'(x)}\,\mathrm{d}x + 2\int_{\mathbb{R}} \frac{(\rho_1(T_t(x))T_t''(x))^2}{\rho_1(T_t(x))T_t'(x)}\,\mathrm{d}x \\
&= 2\int_{\mathbb{R}} \frac{(\rho_1'(T_t(x)))^2}{\rho_1(T_t(x))}(T_t'(x))^3\,\mathrm{d}x + 2\int_{\mathbb{R}} \rho_1(T_t(x))\frac{(T_t''(x))^2}{T_t'(x)}\,\mathrm{d}x
\end{aligned} \tag{6.4}$$

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

Note that in the inequality above we have used $(a+b)^2 \le 2(a^2+b^2)$. To proceed we will estimate each term in the right-hand side of (6.4) using the fact that $|T'(x)|$ is bounded and $I(\rho_0), I(\rho_1) < \infty$. For the first part we have

$$\begin{aligned}\int_{\mathbb{R}} \frac{(\rho_1'(T_t(x)))^2}{\rho_1(T_t(x))}(T_t'(x))^3\,\mathrm{d}x &= \int_{\mathbb{R}} \frac{(\rho_1'(T_t(x)))^2}{\rho_1(T_t(x))}(T_t'(x))(T_t'(x))^2\,\mathrm{d}x\\ &\le C^2\int_{\mathbb{R}} \frac{(\rho_1'(T_t(x)))^2}{\rho_1(T_t(x))}(T_t'(x))\,\mathrm{d}x\\ &= C^2\int_{\mathbb{R}} \frac{(\rho_1'(x))^2}{\rho_1(x)}\,\mathrm{d}x\\ &= C^2 I(\rho_1). \end{aligned} \tag{6.5}$$

Let B be the ball of radius M centered at the origin. Since $T''(x) = 0$ for all $|x| > M$ we can restrict our calculation for the second part in the ball B.

$$\begin{aligned}\int_{\mathbb{R}} \rho_1(T_t(x))\frac{(T_t''(x))^2}{T_t'(x)}\,\mathrm{d}x &= \int_B \rho_1(T_t(x))\frac{(T_t''(x))^2}{T_t'(x)}\,\mathrm{d}x\\ &= \int_B \rho_1(T_t(x))\frac{((1-t)T''(x))^2}{T_t'(x)}\,\mathrm{d}x\\ &= \int_B \rho_1(T_t(x))T_t'(x)\left(\frac{T'(x)(1-t)}{T_t'(x)}\right)^2\left(\frac{T''(x)}{T'(x)}\right)^2\,\mathrm{d}x\\ &= \int_B \rho_1(T_t(x))T_t'(x)\left(\frac{T'(x)(1-t)}{t+(1-t)T'(x)}\right)^2\left(\frac{\rho_0'(x)}{\rho_0(x)} - \frac{\rho_1'(T(x))T'(x)}{\rho_1(T(x))}\right)^2\,\mathrm{d}x\\ &\le 2\int_B \rho_1(T_t(x))T_t'(x)\left(\frac{\rho_0'(x)}{\rho_0(x)}\right)^2\,\mathrm{d}x\\ &\quad + 2\int_B \rho_1(T_t(x))T_t'(x)\left(\frac{\rho_1'(T(x))T'(x)}{\rho_1(T(x))}\right)^2\,\mathrm{d}x\\ &= 2\int_B \frac{\rho_1(T_t(x))T_t'(x)}{\rho_0(x)}\left(\frac{(\rho_0'(x))^2}{\rho_0(x)}\right)\,\mathrm{d}x\\ &\quad + 2\int_B \frac{\rho_1(T_t(x))T_t'(x)T'(x)}{\rho_1(T(x))}\left(\frac{(\rho_1'(T(x)))^2}{\rho_1(T(x))}T'(x)\right)\,\mathrm{d}x\\ &\le C\left(\int_B \frac{(\rho_0'(x))^2}{\rho_0(x)}\,\mathrm{d}x + \int_B \frac{(\rho_1'(T(x)))^2}{\rho_1(T(x))}T'(x)\,\mathrm{d}x\right)\\ &\le C(I(\rho_0)+I(\rho_1)). \end{aligned} \tag{6.6}$$

From (6.4), (6.5) and (6.6) we find that

$$I(\rho_t) = \int_{\mathbb{R}} \frac{(\rho_t'(x))^2}{\rho_t(x)}\,\mathrm{d}x \le C(I(\rho_0)+I(\rho_1)).$$

Next we are going to prove the boundedness of the functional $\int_{\mathbb{R}} |\Psi'(x)|^2\rho_t(x)\mathrm{d}x$.

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

Since $T(x) = x$ for $|x| > M$ we have $\rho_t(x) = \rho_1(x)$ for $|x| > M$. Hence

$$\begin{aligned}\int_{\mathbb{R}} |\Psi'(x)|^2 \rho_t(x)\,\mathrm{d}x &= \int_B |\Psi'(x)|^2 \rho_t(x)\,\mathrm{d}x + \int_{|x|>M} |\Psi'(x)|^2 \rho_t(x)\,\mathrm{d}x \\ &= \int_B |\Psi'(x)|^2 \rho_t(x)\,\mathrm{d}x + \int_{|x|>M} |\Psi'(x)|^2 \rho_1(x)\,\mathrm{d}x \\ &\le C \int_B \rho_t(x)\,\mathrm{d}x + \int_{|x|>M} |\Psi'(x)|^2 \rho_1(x)\,\mathrm{d}x \\ &\le C + \int |\Psi'(x)|^2 \rho_1(x)\,\mathrm{d}x < \infty.\end{aligned}$$

Now we repeat the same argument for $\mathcal{E}(\rho_t)$. Finally by [18], Corollary 20.13 we get that $\mathcal{S}(\rho_0), \mathcal{S}(\rho_1)$ are finite and the result for $\mathcal{S}(\rho_t)$ comes from the fact that $\mathcal{S}$ is geodesically convex. □

Finally we prove that for ρ_0 satisfying the assumptions in the main Theorem 1.1, the set $Q(\rho_0)$ is dense in $\mathcal{P}_2(\mathbb{R})$, thus satisfying assumption *(b)* of Proposition 6.2. The idea behind the lemma is a simple modification of a cut and glue argument (see Fig. 1). For a given measure $\rho_1 \in \mathcal{P}_2(\mathbb{R})$, we construct a measure that is in some sense nice and close to ρ_1 in a compact set, and equal to ρ_0 outside of it. To do so, we first find an interval such that the contribution of both measures ρ_0, ρ_1 to the functionals $\mathcal{S}$ and $\mathcal{E}$ is small outside that interval. We cut out the part of ρ_0 that lies outside the interval, mollify it to ensure both positivity and smoothness, and then add a quadratic decay to get finiteness of the Fisher information functional[7]. For ρ_0 we just keep the tails and add a quadratic decay. The approximating probability measure is then produced by a linear combination of the above constructed measures.

Lemma 6.4. *Assume that $\rho_0 \in \mathcal{P}_2(\mathbb{R})$ such that the density is bounded from below by a positive constant in every compact set, and $\mathcal{F}(\rho_0), \int |\Psi'|^2\,\mathrm{d}\rho_0$ and $I(\rho_0)$ are all finite. Let $\Psi \in C^2(\mathbb{R})$ satisfy either Assumption 4.1 or 4.4. Then for any $\rho_1 \in \mathcal{P}_2(\mathbb{R})$ there exists a sequence ρ^τ in $Q(\rho_0)$ such that $\rho^\tau \to \rho_1$ in the Wasserstein topology, and $\mathcal{F}(\rho^\tau) \to \mathcal{F}(\rho_1)$.*

Proof. Take a $\rho_1 \in \mathcal{P}_2(\mathbb{R})$ with $\mathcal{F}(\rho_1) < \infty$ (otherwise the construction is trivial). First observe that, because $\int x^2 \rho_1(\mathrm{d}x), \int x^2 \rho_0(\mathrm{d}x)$ and $\mathcal{S}(\rho_0), \mathcal{S}(\rho_1)$ are all finite, $\int |\rho_1 \log \rho_1|, \int |\rho_0 \log \rho_0|$ are also finite [12], equation (15). Secondly, $\int |\Psi|\,\mathrm{d}\rho_1$ and $\int |\Psi|\,\mathrm{d}\rho_0$ are also finite since Ψ is bounded from below in both Assumptions 4.1 and 4.4. Therefore, for any $\tau > 0$ there exist Lebesgue points $M_\tau^- < -1$ and $M_\tau^+ > 1$ of ρ_1 such that (to ease notation we assume that $-M_\tau^- = M_\tau^+ =: M_\tau$)

$$\rho_0(-M_\tau), \rho_1(-M_\tau) < \min\left\{\frac{\tau}{|\Psi(-M_\tau)|}, \frac{\tau}{M_\tau^2}\right\} \text{ and } \rho_0(M_\tau), \rho_1(M_\tau) < \min\left\{\frac{\tau}{|\Psi(M_\tau)|}, \frac{\tau}{M_\tau^2}\right\} \tag{6.7a}$$

$$\int_{|x|>M_\tau} \left(\rho_0 + |\rho_0 \log \rho_0| + x^2 \rho_0 + |\Psi|\rho_0\right) < \tau, \tag{6.7b}$$

$$\int_{|x|>M_\tau} \left(\rho_1 + |\rho_1 \log \rho_1| + x^2 \rho_1 + |\Psi|\rho_1\right) < \tau. \tag{6.7c}$$

Define a new density by cutting the tails of ρ_1, and mollifying it by the Gaussian θ_t from Lemma 4.8:

$$\mu^\tau := (\rho_1 \mathbb{1}_{[-M_\tau, M_\tau]}) * \theta_{t_\tau},$$

[7] It is easy to check that a linear decay, which would have been a simpler choice, is not enough to keep the Fisher information functional finite.

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

where t_τ is chosen sufficiently small such that

$$\int_{-M_\tau}^{M_\tau} |\mu^\tau - \rho_1| < \tau, \tag{6.8a}$$

$$\int_{-M_\tau}^{M_\tau} |\Psi\mu^\tau - \Psi\rho_1| < \tau \text{ and } \int_{-M_\tau}^{M_\tau} |x^2\mu^\tau - x^2\rho_1| < \tau, \tag{6.8b}$$

$$\left| \int_{-M_\tau}^{M_\tau} \left(\mu^\tau \log \mu^\tau - \rho_1 \log \rho_1\right) \right| < \tau, \tag{6.8c}$$

$$\mu^\tau(-M_\tau) < \min\left\{\tfrac{\tau}{|\Psi(-M_\tau)|}, \tfrac{\tau}{M_\tau^2}\right\} \text{ and } \mu^\tau(M_\tau) < \min\left\{\tfrac{\tau}{|\Psi(M_\tau)|}, \tfrac{\tau}{M_\tau^2}\right\}, \tag{6.8d}$$

$$\mu^\tau(x) > 0 \quad \text{whenever } |x| \le M_\tau. \tag{6.8e}$$

Observe that property (6.8d) is feasible, because $-M_\tau$ and M_τ are Lebesgue points of ρ_1 and

$$\mu^\tau(M_\tau) \le (\rho_1 * \theta_{t_\tau})(M_\tau) \qquad \text{and} \qquad \mu^\tau(-M_\tau) \le (\rho_1 * \theta_{t_\tau})(-M_\tau).$$

In order to construct a suitable approximating sequence for ρ_1, small intervals around $-M_\tau$ and M_τ are needed where bounds of the type (6.7a) and (6.8d) still hold. Indeed, because of (6.8d) and the continuity of Ψ, there exists $0 < a_\tau < 1$ such that for all $x \in [-M_\tau - a_\tau, -M_\tau + a_\tau]$:

$$\rho_0(x) < \tau, \qquad \text{and} \qquad \rho_0(-M_\tau) < \min\left\{\frac{\tau}{|\Psi(x)|}, \frac{\tau}{x^2}\right\}, \tag{6.9a}$$

$$\mu^\tau(x) < \tau, \qquad \text{and} \qquad \mu^\tau(-M_\tau) < \min\left\{\frac{\tau}{|\Psi(x)|}, \frac{\tau}{x^2}\right\}, \tag{6.9b}$$

and for all $x \in [M_\tau - a_\tau, M_\tau + a_\tau]$:

$$\rho_0(x) < \tau, \qquad \text{and} \qquad \rho_0(M_\tau) < \min\left\{\frac{\tau}{|\Psi(x)|}, \frac{\tau}{x^2}\right\}, \tag{6.9c}$$

$$\mu^\tau(x) < \tau, \qquad \text{and} \qquad \mu^\tau(M_\tau) < \min\left\{\frac{\tau}{|\Psi(x)|}, \frac{\tau}{x^2}\right\}. \tag{6.9d}$$

Note that by assumption $M_\tau > 1$, so that the two intervals can not overlap.

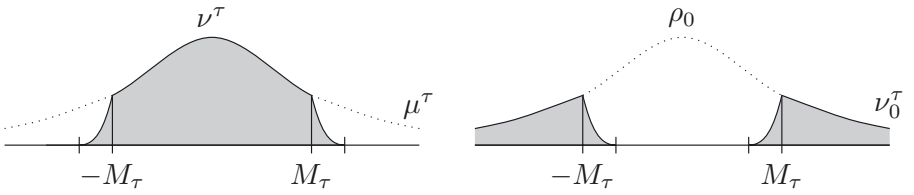
Now, using these intervals, replace the tails of μ^τ, which were introduced by the mollification, by quadratically decaying tails (see Fig. 1)

$$\nu^\tau(x) = \begin{cases} \mu^\tau(x), & |x| \le M_\tau, \\ \left(\frac{x - M_\tau - a_\tau}{a_\tau}\right)^2 \mu^\tau(M_\tau), & M_\tau < x < M_\tau + a_\tau, \\ \left(\frac{x + M_\tau + a_\tau}{a_\tau}\right)^2 \mu^\tau(-M_\tau), & -M_\tau - a_\tau < x < -M_\tau, \\ 0, & |x| \ge M_\tau + a_\tau. \end{cases}$$

On the other hand, the approximation sequence for ρ_1 requires the same tails as ρ_0; these tails are captured by (see Fig. 1)

$$\nu_0^\tau(x) = \begin{cases} 0, & |x| \le M_\tau - a_\tau, \\ \left(\frac{x - M_\tau + a_\tau}{a_\tau}\right)^2 \rho_0(M_\tau), & M_\tau - a_\tau < x < M_\tau, \\ \left(\frac{x - a_\tau + M_\tau}{a_\tau}\right)^2 \rho_0(-M_\tau), & -M_\tau < x < -M_\tau + a_\tau, \\ \rho_0(x), & |x| \ge M_\tau. \end{cases}$$

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press



(a) Cut the tails, add quadratic decay. (b) Crop the tails, add quadratic decay.

FIGURE 1. The construction of ν^τ and ν_0^τ.

Finally, the approximating sequence is defined as a normalised sum of ν and ν_0:

$$\rho^\tau(x) := \alpha_\tau\, \nu^\tau(x) + \nu_0^\tau(x),$$

where $\|\cdot\|_1$ abbreviates the $L^1(\mathbb{R})$ norm, and $\alpha_\tau := \frac{1-\|\nu_0^\tau\|_1}{\|\nu^\tau\|_1}$.

Now we check that the sequence ρ^τ indeed lies in $Q(\rho_0)$. By construction, ρ^τ has the same tails as ρ_0, and it is bounded from below a positive constant on compact sets. Moreover, it is straight-forward that $\int x^2\, d\nu_0^\tau$, $\int x^2\, \nu^\tau$, $\int |\Psi'|^2\, d\nu_0^\tau$, $\int |\Psi'|^2\, d\nu^\tau$ and, $I(\nu_0^\tau)$ are all finite; $I(\nu^\tau)$ is finite by Lemma 4.8. Then the functionals $\int x^2\, \mathrm{d}\rho^\tau$, $\int |\Psi'|^2\, \mathrm{d}\rho^\tau$ are also finite. To check that the Fisher information remains finite:

$$\begin{aligned}
I(\rho^\tau) &= \int_{\mathbb{R}} \frac{\left(\alpha_\tau\, \nu^{\tau\prime} + \nu_0^{\tau\prime}\right)^2}{\alpha_\tau\, \nu^\tau + \nu_0^\tau}\, \mathrm{d}x \\
&\le 2\int_{\mathbb{R}} \frac{\left(\alpha_\tau\, \nu^{\tau\prime}\right)^2}{\alpha_\tau\, \nu^\tau + \nu_0^\tau}\, \mathrm{d}x + 2\int_{\mathbb{R}} \frac{\left(\nu_0^{\tau\prime}\right)^2}{\alpha_\tau\, \nu^\tau + \nu_0^\tau}\, \mathrm{d}x \\
&\le 2\int_{\mathbb{R}} \frac{\left(\alpha_\tau\, \nu^{\tau\prime}\right)^2}{\alpha_\tau\, \nu^\tau}\, \mathrm{d}x + 2\int_{\mathbb{R}} \frac{\left(\nu_0^{\tau\prime}\right)^2}{\nu_0^\tau}\, \mathrm{d}x \\
&= 2\alpha_\tau I(\nu^\tau) + 2I(\nu_0^\tau) < \infty,
\end{aligned}$$

so that indeed $\rho^\tau \in Q(\rho_0)$.

Next, the convergence properties of the sequence ρ^τ are checked. First we show that $\rho^\tau \to \rho_1$ in $L^1(\mathbb{R})$. Since $\|\nu_0^\tau\|_1 \to 0$ and $\|\nu^\tau\|_1 \to 1$, the normalisation constant also converges: $\alpha_\tau \to 1$. Therefore,

$$\begin{aligned}
\int_{\mathbb{R}} |\rho^\tau - \rho_1| &= \int_{\mathbb{R}} |\alpha_\tau \nu^\tau + \nu_0^\tau - \rho_1| \\
&\le \int_{-M_\tau}^{M_\tau} |\alpha_\tau \mu^\tau - \mu^\tau| + \int_{-M_\tau}^{M_\tau} |\mu_\tau - \rho_1| + \int_{|x|>M_\tau} \alpha_\tau \nu^\tau + \int_{|x|>M_\tau} \rho_1 + \int_{\mathbb{R}} \nu_0^\tau \\
&\le \int_{-M_\tau}^{M_\tau} |\alpha_\tau \mu^\tau - \mu^\tau| + \int_{-M_\tau}^{M_\tau} |\mu_\tau - \rho_1| + \alpha_\tau \mu^\tau(-M_\tau) \int_{-M_\tau - a_\tau}^{-M_\tau} \left(\tfrac{x+M_\tau+a_\tau}{a_\tau}\right)^2 \\
&\quad + \alpha_\tau \mu^\tau(M_\tau) \int_{M_\tau}^{M_\tau + a_\tau} \left(\tfrac{x-M_\tau-a_\tau}{a_\tau}\right)^2 + \int_{|x|>M_\tau} \rho_1 + \|\nu_0^\tau\|_1 \\
&\le |\alpha_\tau - 1| \int_{-M_\tau}^{M_\tau} \mu^\tau + \int_{-M_\tau}^{M_\tau} |\mu_\tau - \rho_1| + \alpha_\tau a_\tau\, \mu^\tau(-M_\tau) + \alpha_\tau a_\tau\, \mu^\tau(M_\tau) + \int_{|x|>M_\tau} \rho_1 + \|\nu_0^\tau\|_1 \\
&\le |\alpha_\tau - 1| + \tau + \alpha_\tau \mu^\tau(-M_\tau) + \alpha_\tau \mu^\tau(M_\tau) + \tau + \|\nu_0^\tau\|_1 \xrightarrow[\tau\to 0]{} 0,
\end{aligned} \tag{6.10}$$

where the last line follows from $a_\tau < 1$ together with (6.8a) and (6.7c).

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

Secondly, we check the convergence of the second moments $\int_{\mathbb{R}} x^2\rho^\tau \to \int_{\mathbb{R}} x^2\rho_1$. Observe that there is a uniform bound on

$$\int_{-M_\tau}^{M_\tau} x^2\mu^\tau \le \int_{-M_\tau}^{M_\tau} |x^2\mu^\tau - x^2\rho_1| + \int_{-M_\tau}^{M_\tau} x^2\rho_1 \overset{(6.8b)}{<} \tau + \int_{\mathbb{R}} x^2\rho_1 \le 1 + \int_{\mathbb{R}} x^2\rho_1 \tag{6.11}$$

for $\tau \le 1$. Moreover, for the right-side quadratic tail of ν^τ:

$$\int_{M_\tau}^{M_\tau+a_\tau} x^2\nu^\tau = \int_{M_\tau}^{M_\tau+a_\tau} x^2 \left(\frac{x-M_\tau-a_\tau}{a_\tau}\right)^2 \mu^\tau(M_\tau)\,\mathrm{d}x \le \int_{M_\tau}^{M_\tau+a_\tau} x^2\,\mu^\tau(M_\tau)\,\mathrm{d}x \overset{(6.9d)}{<} \tau\, a_\tau \le \tau, \tag{6.12}$$

and similarly for the other quadratically decaying parts of ν^τ and ν_0^τ. Therefore

$$\begin{aligned}
\int_{\mathbb{R}} |x^2\rho^\tau - x^2\rho_1| &\le \int_{\mathbb{R}} |\alpha_\tau x^2\nu_\tau - x^2\nu^\tau| + \int_{\mathbb{R}} |x^2\nu^\tau - x^2\rho_1| + \int_{\mathbb{R}} x^2\nu_0^\tau \\
&\le |\alpha_\tau - 1| \int_{-M_\tau}^{M_\tau} x^2\mu^\tau + |\alpha_\tau - 1| \int_{|x|>M_\tau} x^2\nu^\tau + \int_{-M_\tau}^{M_\tau} |x^2\mu^\tau - x^2\rho_1| + \int_{|x|>M_\tau} x^2\nu^\tau \\
&\quad + \int_{|x|>M_\tau} x^2\rho_1 + \int_{-M_\tau}^{M_\tau} x^2\nu_0^\tau + \int_{|x|>M_\tau} x^2\rho_0 \\
&\le |\alpha_\tau - 1| \left(1 + \int_{\mathbb{R}} x^2\rho_1 + 2\tau\right) + 7\tau \to 0
\end{aligned}$$

as $\tau \to 0$, where the last line follows from (6.11), (6.12), (6.7b), (6.7c) and (6.8b). Since the sequence ρ^τ converges strongly in $L^1(\mathbb{R})$ to ρ_1 by (6.10), it also converges narrowly. Together with the convergence of the second moments, this implies convergence in the Wasserstein distance [17], Theorem 7.12, which was to be shown.

Thirdly, we need to check that $\mathcal{E}(\rho^\tau) \to \mathcal{E}(\rho_1)$; this is proven in the same way as the convergence of the second moments above, where x^2 is replaced by the potential $\Psi(x)$.

Finally, we prove the convergence of the entropies $\mathcal{S}(\rho^\tau) \to \mathcal{S}(\rho_1)$. Because of

$$|\mathcal{S}(\rho^\tau) - \mathcal{S}(\rho_1)| \le |\mathcal{S}(\nu^\tau) - \mathcal{S}(\rho_1)| + |\mathcal{S}(\rho^\tau) - \mathcal{S}(\nu^\tau)|, \tag{6.13}$$

it suffices to show that both differences on the right-hand side vanish. For the first difference:

$$\mathcal{S}(\nu^\tau) = \int_{-M_\tau}^{M_\tau} \left(\nu^\tau \log \nu^\tau - \rho_1 \log \rho_1\right) + \int_{-M_\tau}^{M_\tau} \rho_1 \log \rho_1 + \int_{M_\tau<|x|<M_\tau+a_\tau} \nu^\tau \log \nu^\tau \to \mathcal{S}(\rho_1).$$

Here, the first term vanishes by (6.8c), and the third term, containing the quadratically decaying tails, vanishes because $\mu^\tau(-M_\tau)$ and $\mu^\tau(M_\tau)$ vanish. For the second difference in (6.13):

$$\begin{aligned}
|\mathcal{S}(\rho^\tau) - \mathcal{S}(\nu^\tau)| \le{}& \underbrace{\int_{-M_\tau+a_\tau}^{M_\tau-a_\tau} |\alpha_\tau\mu^\tau \log \alpha_\tau\mu^\tau - \mu^\tau \log \mu^\tau|}_{(I)} \\
&+ \underbrace{\int_{M_\tau-a_\tau\le|x|\le M_\tau+a_\tau} |\rho^\tau \log \rho^\tau - \nu^\tau \log \nu^\tau|}_{(II)} \\
&+ \underbrace{\int_{|x|>M_\tau+a_\tau} |\rho_0 \log \rho_0|}_{(III)}.
\end{aligned}$$

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

It will now be shown that each of the three parts convergence to 0 as $\tau \to 0$. For the first part, because of (6.8c),

$$\begin{aligned}(I) &= \int_{-M_\tau+a_\tau}^{M_\tau-a_\tau} |\alpha_\tau \mu^\tau \log(\alpha_\tau \mu^\tau) - \mu^\tau \log \mu^\tau| \\ &\leq |\alpha_\tau - 1| \int_{-M_\tau+a_\tau}^{M_\tau-a_\tau} |\mu^\tau \log \mu^\tau| + \alpha_\tau \log \alpha_\tau \int_{-M_\tau+a_\tau}^{M_\tau-a_\tau} |\mu^\tau| \to 0.\end{aligned}$$

For the second part, observe that by assumptions (6.9a), (6.9b), (6.9c), (6.9b), there holds for $M_\tau - a_\tau \leq |x| \leq M_\tau$:

$$\nu_0^\tau(x) \leq \rho_0(M_\tau) < \tau, \qquad \nu^\tau(x) = \mu^\tau(x) < \tau,$$

and for $M_\tau \leq |x| \leq M_\tau + a_\tau$:

$$\nu_0^\tau(x) = \rho_0(x) < \tau, \qquad \nu^\tau(x) \leq \mu^\tau(M_\tau) < \tau.$$

Therefore, since $a_\tau < 1$:

$$\begin{aligned}(II) &\leq \int_{M_\tau-a_\tau\leq|x|\leq M_\tau+a_\tau} (|\rho^\tau \log \rho^\tau| + |\nu^\tau \log \nu^\tau|) \\ &= \int_{M_\tau-a_\tau\leq|x|\leq M_\tau+a_\tau} |(\alpha_\tau \nu^\tau + \nu_0^\tau) \log(\alpha_\tau \nu^\tau + \nu_0^\tau)| + \int_{M_\tau-a_\tau\leq|x|\leq M_\tau+a_\tau} |\nu^\tau \log \nu^\tau| \to 0.\end{aligned}$$

Finally, part (III) converges to 0 by assumption (6.7b). □

Acknowledgements. We would like to thank Nicolas Dirr, Mark Peletier and Johannes Zimmer for their initial suggestion and support during the project. The current proof of Lemma 4.7 without probabilistic tools was done after a discussion with Mark Peletier. We also thank Jin Feng, Truyen Nguyen for their helpful discussion and comments. Manh Hong Duong has received funding from the ITN "FIRST" of the Seventh Framework Programme of the European Community's (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] S. Adams, N. Dirr, M.A. Peletier and J. Zimmer, Large deviations and gradient flows. Philosophical Transactions of the Royal Society A. To appear (2013).
[3] L. Ambrosio, N. Gigli and G. Savaré, Gradient flows in metric spaces and in the space of probability measures. In *Lect. Math.*, ETH Zürich. Birkhauser, Basel, 2nd edition (2008).
[4] J.D. Benamou and Y. Brenier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.* **84** (2000) 375–393.
[5] A. Braides, *Gamma convergence for beginners*. Oxford University Press, Oxford (2002).
[6] D.A. Dawson and J. Gärtner, Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics* **20** (1987) 247–308.
[7] N. Dirr, V. Laschos and J. Zimmer, Upscaling from particle models to entropic gradient flows (submitted) (2010).
[8] R.M. Dudley, *Real analysis and probability*. Wadsworth and Brooks/Cole, Pacific Grove, CA, USA (1989).
[9] A. Dembo and O. Zeitouni, Large deviations techniques and applications, in *Stoch. Model. Appl. Probab.*, vol. 38. Springer, New York, NY, USA, 2nd edition (1987).
[10] J. Feng and T.G. Kurtz, Large deviations for stochastic processes, in *Mathematical surveys and monographs* of vol. 131. AMS, Providence, RI, USA (2006).
[11] J. Feng and T. Nguyen, Hamilton-Jacobi equations in space of measures associated with a system of convervations laws. *J. Math. Pures Appl.* **97** (2011) 318–390.

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press

[12] R. Jordan, D. Kinderlehrer and F. Otto, The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.* **29** (1998) 1–17.
[13] C. Léonard, A large deviation approach to optimal transport. arXiv:org/abs/0710.1461v1 (2007).
[14] F. Otto, The geometry of dissipative evolution equations: the porous medium equation. *Commun. Partial Differ. Equ.* **26** (2001) 101–174.
[15] M.A. Peletier, M. Renger and M. Veneroni, Variational formulation of the Fokker-Planck equation with decay: a particle approach. arXiv:org/abs/1108.3181 (2012).
[16] W. Rudin, Functional Analysis. McGraw-Hill, New York, NY, USA (1973).
[17] C. Villani, Topics in optimal transportation, *Graduate Studies in Mathematics*, vol. 58. AMS, Providence (2003).
[18] C. Villani, Optimal transport, *Grundlehren der Mathematischen Wissenschaften*, vol. 338. Springer-Verlag, Berlin (2009).

https://doi.org/10.1051/cocv/2013049 Published online by Cambridge University Press