HAL
open science

An entropic interpolation proof of the HWI inequality

Ivan Gentil, Christian Léonard, Luigia Ripani, Luca Tamanini

To cite this version:

Ivan Gentil, Christian Léonard, Luigia Ripani, Luca Tamanini. An entropic interpolation proof of the HWI inequality. Stochastic Processes and their Applications, 2020, 130 (2), pp.907-923. 10.1016/j.spa.2019.04.002 . hal-01840629v2

HAL Id: hal-01840629

https://hal.science/hal-01840629v2

Submitted on 15 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

AN ENTROPIC INTERPOLATION PROOF OF THE HWI INEQUALITY

IVAN GENTIL, CHRISTIAN LÉONARD, LUIGIA RIPANI, AND LUCA TAMANINI

Abstract. The HWI inequality is an "interpolation" inequality between the *Entropy* H, the *Fisher information* I and the *Wasserstein distance* W. We present a pathwise proof of the HWI inequality which is obtained through a zero noise limit of the Schrödinger problem. Our approach consists in making rigorous the Otto-Villani heuristics in [23] taking advantage of the entropic interpolations, which are regular both in space and time, rather than the displacement ones.

Contents

Introduction 1
1. Statement of the HWI inequality 2
2. Entropic interpolations 4
3. Proof of Proposition 1.5 9
4. Final remarks and comments 15
References 16

Introduction

In a seminal paper [23], Otto and Villani obtained a powerful functional inequality relating the relative entropy $H(\cdot\,|\,\mathfrak{m})$ with respect to some reference measure $\mathfrak{m} \in \mathcal{P}_2(\mathcal{X})$, the quadratic transport cost $W_2^2(\cdot, \mathfrak{m})$ and the Fisher information $I(\cdot\,|\,\mathfrak{m})$. This so-called HWI inequality roughly states that: $H \leq W\sqrt{I} - \kappa W^2/2$, where the real parameter κ is a curvature lower bound associated to $\mathfrak{m}$, see (1.2), (1.4) and Theorem 1.6 below for the exact statement and its well-known consequences in terms of Talagrand and logarithmic Sobolev inequalities.

The first part of Otto and Villani's article is dedicated to a heuristic proof based on Otto calculus, see [22, 25], where one formally equips the set of probability measures with the Riemannian-like distance W_2 and where McCann displacement interpolations are interpreted as geodesics. Since these interpolations suffer from a lack of regularity, the first and second order time derivatives along them are only formal. Consequently, although heuristics led to the right conjecture, the authors presented an alternative rigorous proof based on a significantly different approach.

In the present article, a new proof of the HWI inequality is proposed. The main idea is to replace the 'irregular' McCann interpolation (μ_t) between two probability measures

Date: July 15, 2018.
2010 *Mathematics Subject Classification.* 60E15,60J60.
Key words and phrases. Entropic interpolations, Schrödinger problem, relative entropy, Fisher information, Wasserstein distance.

μ_0 and μ_1 by a family of 'smooth' curves of measures (μ_t^ε), called 'entropic interpolations' (see Definition 2.4 below), where $\varepsilon > 0$ is a small fluctuation parameter such that μ_t^ε converges to μ_t narrowly as $\varepsilon \downarrow 0$. Otto and Villani's heuristics apply rigorously to (μ_t^ε), so that it remains to let ε tend down to zero to obtain the desired result.

The paper is structured as follows. Section 1 is dedicated to the statement of the HWI inequality. Basic material about entropic interpolations which is needed for the proof is gathered at Section 2. Finally the proof of the inequality is done at Section 3; its core is Lemma 3.11 which is the analogue of Otto and Villani's heuristic approach. In Section 4 we propose some comments about possible extensions and simplifications of our approach.

1. Statement of the HWI inequality

Before stating the HWI inequality at Proposition 1.5 and Theorem 1.6 below, we need to make clear the framework we shall work within and introduce the quantities H, W and I.

Setting 1. Let $(\mathcal{X}, \mathsf{d}, \mathfrak{m})$ be:

(a) either $(\mathbb{R}^n, |\cdot|, \mathfrak{m})$, where $|\cdot|$ is the Euclidean distance and the reference measure $\mathfrak{m}$ is defined as

$$\mathfrak{m} := e^{-V}\mathcal{L}^n \tag{1.1}$$

with $\mathcal{L}^n$ the n-dimensional Lebesgue measure and $V : \mathbb{R}^n \to [0, \infty)$ satisfying the following hypotheses: it belongs to $C^\infty(\mathbb{R}^n)$, is such that $\mathfrak{m}$ is a probability measure and

$$\mathrm{Hess}(V) \geq \kappa \mathrm{Id} \tag{1.2}$$

for some $\kappa \in \mathbb{R}$;

(b) or $(M, \mathsf{d}_g, \mathfrak{m})$, where M is a smooth, connected and complete Riemannian manifold without boundary and with metric tensor g, d_g is the induced distance and $\mathfrak{m}$ is given by

$$\mathfrak{m} := e^{-V}\mathrm{Vol} \tag{1.3}$$

with Vol the volume measure on M and $V : M \to [0, \infty)$ satisfying the following hypotheses: it belongs to $C^\infty(\mathbb{R}^n)$, is such that $\mathfrak{m}$ is a probability measure and, for some $\kappa \in \mathbb{R}$ and $N \geq n = \dim(M)$, the Bakry-Émery Ricci tensor $\mathrm{Ric}_{V,N}$ satisfies the lower bound

$$\mathrm{Ric}_{V,N} := \mathrm{Ric}_g - (N-n)\frac{\mathrm{Hess}(e^{-\frac{1}{N-n}V})}{e^{-\frac{1}{N-n}V}} \geq \kappa g. \tag{1.4}$$

Inequality (1.2) (resp. (1.3)) corresponds to the so-called Bakry-Émery $\mathsf{CD}(\kappa, \infty)$ curvature-dimension condition (resp. $\mathsf{CD}(\kappa, N)$), cf. [4, Sec. C.6].

Relative entropy. For any two probability measures p and r on a measurable space $\mathcal{Z}$ the relative entropy of p with respect to r is defined by

$$H(p\,|\,r) := \int_{\mathcal{Z}} \log\left(\frac{\mathrm{d}p}{\mathrm{d}r}\right)\mathrm{d}p \in [0, \infty],$$

where it is understood that this quantity is infinite when p is not absolutely continuous with respect to r. In our case, $\mathcal{Z}$ will be $\mathcal{X}$, $\mathcal{X} \times \mathcal{X}$ or $C([0,1], \mathcal{X})$.

Quadratic transport cost. By $\mathcal{P}(\mathcal{X})$ we shall denote the space of Borel probability measures on $\mathcal{X}$ and by $\mathcal{P}_2(\mathcal{X})$ the subclass of those with finite second moment, namely all $\mu \in \mathcal{P}(\mathcal{X})$ such that $\int_{\mathcal{X}} \mathsf{d}^2(\cdot, x)\mathrm{d}\mu < \infty$ for some (and thus all) $x \in \mathcal{X}$. With this said, the squared Wasserstein distance between $\mu, \nu \in \mathcal{P}_2(\mathcal{X})$ is defined as

$$W_2^2(\mu, \nu) := \inf_{\pi} \int_{\mathcal{X}\times\mathcal{X}} \mathsf{d}^2(x, y)\, \pi(\mathrm{d}x\mathrm{d}y)$$

where the infimum runs through all the couplings $\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{X})$ of μ and ν, that is $\pi(\mathrm{d}x \times \mathcal{X}) = \mu(\mathrm{d}x)$ and $\pi(\mathcal{X} \times \mathrm{d}y) = \nu(\mathrm{d}y)$.

Fisher information. The Fisher information of $\mu \in \mathcal{P}(\mathcal{X})$ with respect to $\mathfrak{m}$ is defined by

$$I(\mu \,|\, \mathfrak{m}) := 4 \int_{\mathcal{X}} |\nabla \sqrt{\rho}|^2 \,\mathrm{d}\mathfrak{m} = \int_{\{\rho>0\}} \frac{|\nabla \rho|^2}{\rho} \,\mathrm{d}\mathfrak{m} \qquad \text{if } \mu = \rho\mathfrak{m},\ \sqrt{\rho} \in \mathcal{W}^{1,2}(\mathcal{X})$$

and $+\infty$ otherwise. Up to identify μ with its density, the Fisher information is lower semicontinuous with respect to the weak topology of $L^1(\mathfrak{m})$ (see for instance [2]).

With this premise, the statement of the HWI inequality is

Proposition 1.5 (HWI inequality)**.** *Let $(\mathcal{X}, \mathsf{d}, \mathfrak{m})$ be as in Setting 1. Then for any $\mu_0, \mu_1 \in \mathcal{P}_2(\mathcal{X})$ such that $H(\mu_0 \,|\, \mathfrak{m}) < \infty$,*

$$H(\mu_1 \,|\, \mathfrak{m}) - H(\mu_0 \,|\, \mathfrak{m}) \le W_2(\mu_0, \mu_1)\sqrt{I(\mu_1 \,|\, \mathfrak{m})} - \frac{\kappa}{2} W_2^2(\mu_0, \mu_1).$$

In this case it is said that the reference measure $\mathfrak{m}$ satisfies the HWI inequality. Let us note that even if the parameter N, in the Setting 1-(b), doesn't appear in the HWI inequality, it is useful to obtain estimates like inequality (2.14) below.

As already shown by Otto and Villani in [23], different choices of μ_0 and μ_1 in Proposition 1.5 entail three important consequences collected here below.

Theorem 1.6. *Let $(\mathcal{X}, \mathsf{d}, \mathfrak{m})$ be as in Setting 1 with the further assumption that $\mathfrak{m} \in \mathcal{P}_2(\mathcal{X})$. Then the following inequalities are satisfied:*

(a) HWI inequality:

$$H(\nu \,|\, \mathfrak{m}) \le W_2(\nu, \mathfrak{m})\sqrt{I(\nu \,|\, \mathfrak{m})} - \frac{\kappa}{2} W_2^2(\nu, \mathfrak{m}), \qquad \forall \nu \in \mathcal{P}_2(\mathcal{X});$$

(b) Talagrand inequality:

$$\frac{\kappa}{2} W_2^2(\nu, \mathfrak{m}) \le H(\nu \,|\, \mathfrak{m}), \qquad \forall \nu \in \mathcal{P}_2(\mathcal{X});$$

(c) Logarithmic Sobolev inequality: assume that $\kappa > 0$, then

$$H(\nu \,|\, \mathfrak{m}) \le \frac{1}{2\kappa} I(\nu \,|\, \mathfrak{m}), \qquad \forall \nu \in \mathcal{P}(\mathcal{X}).$$

Proof. First of all, since $\mathfrak{m} \in \mathcal{P}_2(\mathcal{X})$, it follows that $W_2(\nu, \mathfrak{m})$ is finite for all $\nu \in \mathcal{P}_2(\mathcal{X})$.

(a) The HWI inequality is obtained by choosing $\mu_0 = \mathfrak{m}, \mu_1 = \nu$.

(b) The Talagrand inequality is obtained by choosing $\mu_0 = \nu, \mu_1 = \mathfrak{m}$.

(c) When $\kappa > 0$, the logarithmic Sobolev inequality with $\nu \in \mathcal{P}_2(\mathcal{X})$ follows by using the trivial inequality $xy - \kappa x^2/2 \le y^2/(2\kappa)$, for $x, y \ge 0$, in the right-hand side of the HWI inequality (a). To extend this result to the case where $\nu \in \mathcal{P}(\mathcal{X})$, a standard approximation argument (carried out for instance in Lemma 3.1 below) is sufficient.

□

Remarks 1.7.

(a) When $\kappa > 0$, any $\nu \in \mathcal{P}(\mathcal{X})$ such that $H(\nu \,|\, \mathfrak{m}) < \infty$ stands in $\mathcal{P}_2(\mathcal{X})$. This follows from the variational representation of the relative entropy as

$$H(\nu \,|\, \mathfrak{m}) = \sup_f \Big\{ \int_{\mathcal{X}} f \,\mathrm{d}\nu - \log \int_{\mathcal{X}} e^f \,\mathrm{d}\mathfrak{m} \,:\, \int_{\mathcal{X}} e^f \,\mathrm{d}\mathfrak{m} < \infty, \int_{\mathcal{X}} f^- \,\mathrm{d}\nu < \infty \Big\},$$

where $f^- := \max\{-f, 0\}$ (see for instance [17] for a proof). If we choose $f = \alpha \mathsf{d}^2(\cdot, x)$ for some $x \in \mathcal{X}$ and $0 < \alpha < \kappa$, then $\int_{\mathcal{X}} e^f \,\mathrm{d}\mathfrak{m} < \infty$ holds: in Setting 1-(a) this is due to (1.2) whereas in Setting 1-(b) to (1.4) and the Bishop-Gromov inequality. Therefore

$$\int_{\mathcal{X}} \mathsf{d}^2(\cdot, x) \,\mathrm{d}\nu \le \alpha^{-1} \Big(H(\nu \,|\, \mathfrak{m}) + \log \int_M e^{\alpha \mathsf{d}^2(\cdot, x)} \,\mathrm{d}\mathfrak{m} \Big),$$

whence the claim. In particular, $\mathfrak{m} \in \mathcal{P}_2(\mathcal{X})$.

(b) Talagrand inequality (b) is irrelevant when $\kappa \le 0$. When $\kappa > 0$, in view of previous remark it extends to all $\nu \in \mathcal{P}(\mathcal{X})$ provided that one sets $W_2(\nu, \mathfrak{m}) = \infty$ when ν does not belong to $\mathcal{P}_2(\mathcal{X})$.

(c) It follows from the logarithmic Sobolev inequality that when (1.2) or (1.4) holds with some $\kappa > 0$, any $\nu \in \mathcal{P}(\mathcal{X})$ such that $H(\nu \,|\, \mathfrak{m}) = \infty$ satisfies $I(\nu \,|\, \mathfrak{m}) = \infty$. Similarly, it follows from the HWI inequality that when (1.2) or (1.4) is only supposed to hold with κ real, as soon as $\nu \in \mathcal{P}_2(\mathcal{X})$, then $H(\nu \,|\, \mathfrak{m}) = \infty$ implies that $I(\nu \,|\, \mathfrak{m}) = \infty$.

2. Entropic interpolations

In this section we propose a short and self-contained presentation of entropic interpolations. The purpose is twofold: to provide the reader with those notions and results that will be frequently used later on and discuss their physical interpretation via Nelson's dynamical view of diffusion processes. For sake of simplicity, the latter will be carried out in the more familiar Euclidean setting.

Notations. Let $X = (X_t)_{0 \le t \le 1}$, $X_t : \Omega \to \mathcal{X}$ with $\Omega := C([0,1], \mathcal{X})$ be the canonical process, defined by

$$X_t(\omega) := \omega_t, \qquad \forall\, \omega \in \Omega, \forall 0 \le t \le 1.$$

For any path measure $\mathsf{Q} \in \mathcal{P}(\Omega)$ and each $0 \le t \le 1$, we denote by $\mathsf{Q}_t := (X_t)_\# \mathsf{Q} \in \mathcal{P}(\mathcal{X})$ the t-th marginal of Q, that is the law of the position X_t at time t of the random path X under Q. Moreover, for any $0 \le s, t \le 1$, we shall denote by Q_{st} the joint law of X_s and X_t under Q, namely $\mathsf{Q}_{st} := (X_s, X_t)_\# \mathsf{Q}$.

As reference path measure $\mathsf{R} \in \mathcal{P}(\Omega)$ we consider the law of the Markov diffusion process with generator

$$L := \frac{1}{2}(\Delta - \nabla V \cdot \nabla)$$

with initial law $\mathsf{R}_0 = \mathfrak{m}$, where the potential V appears at (1.1), (1.3), Δ is the Laplace-Beltrami operator on $\mathcal{X}$ and ∇ the Levi-Civita connection associated to the metric g (in Setting 1-(a) they are nothing but the standard Laplacian and gradient). It is well-known that R is a reversible Markov measure with reversing measure $\mathfrak{m}$. In particular it is stationary, that is $\mathsf{R}_t = \mathfrak{m}$ for all $0 \le t \le 1$. For any $\varepsilon > 0$ we denote by X^ε the time-rescaled process defined by $X^\varepsilon_t := X_{\varepsilon t}$, $0 \le t \le 1$, and by $\mathsf{R}^\varepsilon := (X^\varepsilon)_\# \mathsf{R}$ the

corresponding path measure. The parameter ε is meant to tend to zero so that R^ε is a slowed down version of R, whose generator is

$$L^\varepsilon = \varepsilon L = \frac{\varepsilon}{2}(\Delta - \nabla V \cdot \nabla).$$

For any $\varepsilon > 0$, as a time rescaling of R, R^ε is also $\mathfrak{m}$-reversible, so that in particular $\mathsf{R}^\varepsilon_t = \mathfrak{m}$ for all t.

The 1-parameter semigroup associated to L will be denoted by $(\mathcal{T}_t)$ and, in a completely analogous way, $(\mathcal{T}^\varepsilon_t)$ the one associated to L^ε; notice that $\mathcal{T}^\varepsilon_t = \mathcal{T}_{\varepsilon t}$ for all $t \geq 0$. Within Setting 1 it is well-known (see for instance [15]) that there exists a unique heat kernel $\mathsf{r}_t(x, y)$ associated to L which is a smooth function on $(0, \infty) \times \mathcal{X} \times \mathcal{X}$. Therefore, the semigroup $(\mathcal{T}_t)$ can be represented by

$$\mathcal{T}_t f(x) = \mathsf{E}_{\mathsf{R}}[f(X_0) \,|\, X_t = x] = \int_{\mathcal{X}} f(y)\mathsf{r}_t(x, y)\, \mathfrak{m}(\mathrm{d}y) \tag{2.1}$$

for all $f \in L^\infty(\mathfrak{m})$. Let us also recall that, in conjunction with (1.2) or (1.4), $(\mathcal{T}_t)$ enjoys the Bakry-Émery contraction estimate

$$|\nabla \mathcal{T}_t f|^2 \leq e^{-2\kappa t}\mathcal{T}_t(|\nabla f|^2) \qquad \forall f \in C^\infty_c(\mathcal{X}),\, t \geq 0. \tag{2.2}$$

For its proof as well as for all the regularizing properties of $\mathcal{T}$ that will be used throughout the paper, we address the reader to [4, Thm 3.2.3].

Finally, recall that the operators

$$\Gamma(f, g) := L(fg) - fLg - gLf, \quad \Gamma(f, f) = \Gamma(f), \quad \Gamma_2(f) := L\Gamma(f) - 2\Gamma(f, Lf), \tag{2.3}$$

defined for all $f, g \in C^\infty_c(\mathcal{X})$, are naturally associated to L. As it is not difficult to see, the drift V in $\mathfrak{m}$ does not affect Γ, since $\Gamma(f, g) = \langle \nabla f, \nabla g\rangle$. It is worth mentioning that, with respect to the standard definition provided in [4, Sec. 1.16], here Γ and Γ_2 are not divided by 2, as the factor $1/2$ already appears in L, which thus corresponds to an SDE driven by a standard Brownian motion.

The Schrödinger problem. Let $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{X})$ be two probability measures: the Schrödinger problem associated with $\mathsf{R}^\varepsilon, \mu_0, \mu_1$ is defined by

$$\varepsilon H(\mathsf{P} \,|\, \mathsf{R}^\varepsilon) \to \min; \qquad \mathsf{P} \in \mathcal{P}(\Omega) \,:\, \mathsf{P}_0 = \mu_0, \mathsf{P}_1 = \mu_1 \tag{S_ε}$$

and its value is called 'entropic cost'. As a strictly convex minimization problem, it admits at most one solution.

Definition 2.4 (R^ε-entropic interpolation)**.** *The solution P^ε of (S_ε), if it exists, is called the R^ε-entropic bridge between μ_0 and μ_1. The R^ε-entropic interpolation (μ^ε_t) between μ_0 and μ_1 is defined as the time marginal flow of the solution P^ε, namely*

$$\mu^\varepsilon_t := \mathsf{P}^\varepsilon_t, \qquad 0 \leq t \leq 1.$$

The name 'entropic interpolation' stems from the connection with displacement interpolation. Indeed, it is known from [20] that

$$\lim_{\varepsilon \downarrow 0} \inf (\mathrm{S}_\varepsilon) = \frac{1}{2} W_2^2(\mu_0, \mu_1). \tag{2.5}$$

This limit is a consequence of a more general result asserting that (S_ε) converges to the quadratic Monge-Kantorovich problem as $\varepsilon \downarrow 0$ in the sense of Γ-convergence, see [16].

As shown in [18], if P^ε is the solution of (S_ε), then there exist two non-negative measurable functions $f^\varepsilon, g^\varepsilon : \mathcal{X} \to [0,\infty)$ such that $\mathsf{P}^\varepsilon = f^\varepsilon(X_0) g^\varepsilon(X_1)\, \mathsf{R}^\varepsilon$ and defining

$$f_t^\varepsilon := \mathcal{T}_t^\varepsilon f^\varepsilon, \quad g_t^\varepsilon := \mathcal{T}_{1-t}^\varepsilon g^\varepsilon, \qquad 0 \le t \le 1 \tag{2.6}$$

we obtain $\mathsf{P}_t^\varepsilon(\mathrm{d}x) = \mu_t^\varepsilon(\mathrm{d}x) = \rho_t^\varepsilon(x)\, \mathfrak{m}(\mathrm{d}x)$ with $\rho_t^\varepsilon(x) := f_t^\varepsilon(x) g_t^\varepsilon(x)$. By computing the endpoint marginals of P^ε we obtain the following two conditions

$$\rho_0 := \frac{\mathrm{d}\mu_0}{\mathrm{d}\mathfrak{m}} = f^\varepsilon\, g_0^\varepsilon \qquad\qquad \rho_1 := \frac{\mathrm{d}\mu_1}{\mathrm{d}\mathfrak{m}} = f_1^\varepsilon\, g^\varepsilon \tag{2.7}$$

usually known as 'Schrödinger system': indeed, if we interprete them as a nonlinear system where the unknowns are f^ε and g^ε, then the (unique up to an obvious multiplicative rescaling) solution completely determines P^ε (see for instance [18]). As far as the convergence of entropic interpolations towards displacement ones is investigated, the following functions

$$\varphi_t^\varepsilon = \varepsilon \log f_t^\varepsilon \qquad\qquad \psi_t^\varepsilon := \varepsilon \log g_t^\varepsilon$$

are of special interest. We also set $\varphi^\varepsilon := \varepsilon \log f^\varepsilon$ in $\mathrm{supp}(\mu_0)$ and $\psi^\varepsilon := \varepsilon \log g^\varepsilon$ in $\mathrm{supp}(\mu_1)$. They are called Schrödinger potentials, in connection with Kantorovich ones.

Existence and regularity results. Let us now derive a criterion, in terms of the endpoint marginals μ_0 and μ_1, for the existence of regular functions $f^\varepsilon, g^\varepsilon$ with well-defined Fisher information solving the Schrödinger system (2.7). As noticed in [12], [13] the regularity (smoothness and integrability) of μ_0 (resp. μ_1) is inherited by f^ε (resp. g^ε). In the next result we extend this property: if $I(\mu_0 \,|\, \mathfrak{m})$ is finite, then so is $I(f^\varepsilon \mathfrak{m} \,|\, \mathfrak{m})$ and analogously with μ_1, g^ε.

Proposition 2.8. *Let $(\mathcal{X}, \mathsf{d}, \mathfrak{m})$ be as in Setting 1, $\varepsilon > 0$ and consider two probability measures $\mu_0 = \rho_0 \mathfrak{m}, \mu_1 = \rho_1 \mathfrak{m} \in \mathcal{P}(\mathcal{X})$ with compact supports. Then the following hold:*

(a) The Schrödinger problem (S_ε) admits a solution with finite entropy if and only if $H(\mu_0 \,|\, \mathfrak{m})$, $H(\mu_1 \,|\, \mathfrak{m}) < \infty$. This solution is unique and the R^ε-entropic interpolation between μ_0 and μ_1 exists.

(b) Suppose in addition that $\rho_0, \rho_1 \in L^\infty(\mathfrak{m})$.

(i) Then $f^\varepsilon, g^\varepsilon \in L^\infty(\mathfrak{m})$.

(ii) For any $0 < t < 1$ the functions $f_t^\varepsilon, g_t^\varepsilon, \rho_t^\varepsilon$ as well as f_1^ε and g_0^ε belong to $C^\infty(\mathcal{X}) \cap L^\infty(\mathfrak{m})$.

(iii) For any $k \in \mathbb{N} \cup \{\infty\}$, if $\rho_0 \in C^k(\mathcal{X})$ (resp. ρ_1), then the function f^ε (resp. g^ε) also belongs to $C^k(\mathcal{X})$.

(c) In addition to item (b), suppose that $I(\mu_0 \,|\, \mathfrak{m})$ is finite (resp. $I(\mu_1 \,|\, \mathfrak{m})$). Then so is $I(f^\varepsilon \mathfrak{m} \,|\, \mathfrak{m})$ (resp. $I(g^\varepsilon \mathfrak{m} \,|\, \mathfrak{m})$).

Proof. For (a) and (b-i) see [13, Proposition 2.1]. As regards (b-ii), the fact that $f^\varepsilon \in L^\infty(\mathfrak{m})$, $\mathsf{r}_t \in C^\infty(\mathcal{X} \times \mathcal{X})$ and (2.1) imply $f_t^\varepsilon \in C^\infty(\mathcal{X})$ for all $0 < t \le 1$, while the maximum principle ensures that $f_t^\varepsilon \in L^\infty(\mathfrak{m})$; the statements for g_t^ε and $\rho_t^\varepsilon = f_t^\varepsilon g_t^\varepsilon$ follow by the same reason.

(b-iii). Notice that the first equation in the Schrödinger system (2.7) can be rewritten as

$$f^\varepsilon = \frac{\rho_0}{\mathcal{T}_1^\varepsilon g^\varepsilon}.$$

Since $\mathcal{T}^\varepsilon$ is positivity improving, $\mathcal{T}_1^\varepsilon g^\varepsilon$ is smooth and $\rho_0 \in C^k(\mathcal{X})$ with compact support, the conclusion follows. A similar argument holds for g^ε.

(c). Observe that the Schrödinger system (2.7) and $g_0^\varepsilon > 0$ (as $\mathcal{T}^\varepsilon$ is positivity improving) force f^ε to have the same support as ρ_0. Thus f^ε has compact support and, as a consequence, $g_0^\varepsilon \geq c > 0$ in $\mathrm{supp}(f^\varepsilon)$ for some c. This remark and the chain rule allow us to say that

$$\begin{aligned} c\frac{|\nabla f^\varepsilon|^2}{f^\varepsilon} &\leq g_0^\varepsilon \frac{|\nabla f^\varepsilon|^2}{f^\varepsilon} \leq g_0^\varepsilon \frac{|\nabla f^\varepsilon|^2}{f^\varepsilon} + f^\varepsilon \frac{|\nabla g_0^\varepsilon|^2}{g_0^\varepsilon} = \frac{|\nabla \rho_0|^2}{\rho_0} - 2\langle \nabla f^\varepsilon, \nabla g_0^\varepsilon \rangle \\ &\leq \frac{|\nabla \rho_0|^2}{\rho_0} + |\nabla f^\varepsilon|^2 + |\nabla g_0^\varepsilon|^2, \end{aligned}$$

so that it remains to prove the integrability of the right-hand side. The first term is integrable by assumption, the third one by the regularization properties of $\mathcal{T}^\varepsilon$, while for the second one notice that $I(\mu_0 \,|\, \mathfrak{m}) < \infty$ and $\rho_0 \in L^\infty(\mathfrak{m})$ imply $|\nabla \rho_0| \in L^2(\mathfrak{m})$; plugging this information into

$$c|\nabla f^\varepsilon| \leq g_0^\varepsilon |\nabla f^\varepsilon| \leq |\nabla \rho_0| + f^\varepsilon |\nabla g_0^\varepsilon|$$

we get the desired conclusion. □

For this reason we formulate the following

Assumptions 2.9 (Hypotheses on μ_0, μ_1)**.** *The endpoint marginals $\mu_0, \mu_1 \in \mathcal{P}_2(\mathcal{X})$ are such that μ_0, μ_1 have compact supports, $H(\mu_0 \,|\, \mathfrak{m}), H(\mu_1 \,|\, \mathfrak{m}), I(\mu_0 \,|\, \mathfrak{m}), I(\mu_1 \,|\, \mathfrak{m}) < \infty$ and their densities ρ_0, ρ_1 belong to $C^\infty(\mathcal{X})$.*

A dynamical viewpoint. As concerns the evolution of entropic interpolations and Schrödinger potentials, let us first notice that under Assumptions 2.9, by the very definition (2.6), item (b-iii) of Proposition 2.8 and the fact that $\mathsf{r}_t(x,y) \in C^\infty((0,\infty) \times \mathcal{X}^2)$ we deduce that f_t^ε and g_t^ε are smooth on $[0,1] \times \mathcal{X}$ and solve

$$\partial_t f_t^\varepsilon = \varepsilon L f_t^\varepsilon \qquad\qquad -\partial_t g_t^\varepsilon = \varepsilon L g_t^\varepsilon \tag{2.10}$$

in the classical sense. Moreover, if we look at $(f_t^\varepsilon), (g_t^\varepsilon)$ as curves parametrized by t with values in $\mathcal{W}^{1,2}(\mathcal{X}, \mathfrak{m})$, they belong to the set $AC([0,1], \mathcal{W}^{1,2}(\mathcal{X}, \mathfrak{m}))$ of all absolutely continuous functions from $[0,1]$ to $\mathcal{W}^{1,2}(\mathcal{X}, \mathfrak{m})$. As a consequence the PDEs above hold also when $\partial_t f_t^\varepsilon, \partial_t g_t^\varepsilon$ are seen as strong $\mathcal{W}^{1,2}$-limits.

Relying on that, it follows that the Schrödinger potentials $\varphi_t^\varepsilon, \psi_t^\varepsilon$ are smooth on $(0,1] \times \mathcal{X}$ and $[0,1) \times \mathcal{X}$ and solve forward and backward Hamilton-Jacobi-Bellman equations respectively, i.e.

$$\partial_t \varphi_t^\varepsilon = \frac{1}{2}|\nabla \varphi_t^\varepsilon|^2 + \varepsilon L \varphi_t^\varepsilon \qquad\qquad \partial_t \psi_t^\varepsilon = -\frac{1}{2}|\nabla \psi_t^\varepsilon|^2 - \varepsilon L \psi_t^\varepsilon, \tag{2.11}$$

while for (ρ_t^ε) the continuity equation

$$\partial_t \rho_t^\varepsilon + \mathrm{div}_{\mathfrak{m}}(\nabla \vartheta_t^\varepsilon \, \rho_t^\varepsilon) = 0 \tag{2.12}$$

is satisfied in $(0,1) \times \mathcal{X}$, where $\vartheta_t^\varepsilon := (\psi_t^\varepsilon - \varphi_t^\varepsilon)/2$ and $\mathrm{div}_{\mathfrak{m}}$ denotes the divergence with respect to $\mathfrak{m}$, i.e. the opposite of the adjoint of the differential in $\mathcal{W}^{1,2}(\mathcal{X}, \mathfrak{m})$. This last PDE is strongly linked to the dynamical representation of the entropic cost $\inf(\mathrm{S}_\varepsilon)$, namely

$$\varepsilon \inf(\mathrm{S}_\varepsilon) = \frac{\varepsilon}{2}\Big(H(\mu_0 \,|\, \mathfrak{m}) + H(\mu_1 \,|\, \mathfrak{m})\Big) + \int_{[0,1]\times\mathcal{X}} \frac{|\nabla \vartheta_t^\varepsilon|^2}{2} \mathrm{d}t \mathrm{d}\mu_t^\varepsilon + \frac{\varepsilon^2}{8}\int_0^1 I(\mu_t^\varepsilon \,|\, \mathfrak{m}) \, \mathrm{d}t, \tag{2.13}$$

shown in [5, 9] for Setting 1-(a) and in [14] for rather general metric measure spaces including Setting 1-(b). By the very definition of ϑ_t^ε and since $\varepsilon \log \rho_t^\varepsilon = \varphi_t^\varepsilon + \psi_t^\varepsilon$, this implies

$$\int_{[0,1]\times\mathcal{X}} \frac{|\nabla\varphi_t^\varepsilon|^2}{2} \mathrm{d}t\mathrm{d}\mu_t^\varepsilon, \int_{[0,1]\times\mathcal{X}} \frac{|\nabla\psi_t^\varepsilon|^2}{2} \mathrm{d}t\mathrm{d}\mu_t^\varepsilon < \infty. \tag{2.14}$$

Remark 2.15. The regularity of Schrödinger potentials comes from the one of $f_t^\varepsilon, g_t^\varepsilon$, the fact that $f_t^\varepsilon, g_t^\varepsilon$ are everywhere positive and the logarithm is smooth on $(0,\infty)$. However, Schrödinger potentials are not integrable in general, as $f_t^\varepsilon, g_t^\varepsilon$ can be arbitrarily close to 0. Thus we cannot study their behaviour as curves with values into some $L^p(\mathfrak{m})$ space. This also explains why the regularity of φ_t^ε (resp. ψ_t^ε) does not extend up to $t=0$ (resp. $t=1$).

A physical interpretation. In the Euclidean framework of Setting 1-(a) it is possible to make a bridge between what is presented so far and Nelson's formalism [21], thus providing a physical motivation for some results stated above and a further perspective on some objects.

Following [21] we define the forward and backward velocities of the Markov measure P for any $x \in \mathbb{R}^n$ and $0 \le t < 1$, $0 < t \le 1$ respectively by

$$\overrightarrow{v}_t^{\mathsf{P}}(x) := \lim_{h\to 0^+} \mathsf{E}_{\mathsf{P}}\left(\frac{X_{t+h}-X_t}{h} \mid X_t = x\right) \qquad \overleftarrow{v}_t^{\mathsf{P}}(x) := \lim_{h\to 0^+} \mathsf{E}_{\mathsf{P}}\left(\frac{X_{t-h}-X_t}{h} \mid X_t = x\right)$$

when these limits are meaningful. The current velocity is defined, for any $0 < t < 1$ and $x \in \mathbb{R}^n$, by

$$v_t^{\mathrm{cu},\mathsf{P}}(x) := \lim_{h\to 0^+} \mathsf{E}_{\mathsf{P}}\left(\frac{X_{t+h}-X_{t-h}}{2h} \mid X_t = x\right)$$

and the osmotic velocity by

$$v_t^{\mathrm{os},\mathsf{P}}(x) := \lim_{h\to 0^+} \mathsf{E}_{\mathsf{P}}\left(\frac{X_{t+h}-2X_t+X_{t-h}}{2h} \mid X_t = x\right).$$

We immediately see that

$$\begin{cases} \overrightarrow{v}^P = v^{\mathrm{cu},P} + v^{\mathrm{os},P}, \\ \overleftarrow{v}^P = -v^{\mathrm{cu},P} + v^{\mathrm{os},P}, \end{cases} \qquad \text{and} \qquad \begin{cases} v^{\mathrm{cu},P} = (\overrightarrow{v}^P - \overleftarrow{v}^P)/2, \\ v^{\mathrm{os},P} = (\overrightarrow{v}^P + \overleftarrow{v}^P)/2. \end{cases}$$

For R^ε, it is easily seen that

$$v^{\mathrm{cu},\mathsf{R}^\varepsilon} = 0, \qquad \overrightarrow{v}^{\mathsf{R}^\varepsilon} = \overleftarrow{v}^{\mathsf{R}^\varepsilon} = v^{\mathrm{os},\mathsf{R}^\varepsilon} = -\frac{\varepsilon}{2}\nabla V$$

whereas for P^ε

$$\begin{cases} \overrightarrow{v}_t^{\mathsf{P}^\varepsilon} = \nabla\psi_t^\varepsilon + \frac{\varepsilon}{2}\nabla V, \\ \overleftarrow{v}_t^{\mathsf{P}^\varepsilon} = \nabla\varphi_t^\varepsilon - \frac{\varepsilon}{2}\nabla V, \end{cases} \qquad \text{and} \qquad \begin{cases} v_t^{\mathrm{cu},\mathsf{P}^\varepsilon} = \nabla\vartheta_t^\varepsilon - \frac{\varepsilon}{2}\nabla V, \\ v_t^{\mathrm{os},\mathsf{P}^\varepsilon} = \frac{\varepsilon}{2}\nabla\log\rho_t^\varepsilon. \end{cases}$$

This allows to rewrite the continuity equation (2.12) as

$$\partial_t\rho_t^\varepsilon + \mathrm{div}(v^{\mathrm{cu},\mathsf{P}^\varepsilon}\rho_t^\varepsilon) = 0$$

where now div denotes the divergence with respect to $\mathcal{L}^n$ if in Setting 1-(a) or Vol if in Setting 1-(b). This is perfectly coherent with (2.12) since $\mathrm{div}_{\mathfrak{m}}(w) = \mathrm{div}(w) - \frac{\varepsilon}{2}\nabla V \cdot w$ for any vector field w. Furthermore, (2.13) becomes

$$\varepsilon \inf(\mathrm{S}_\varepsilon) = \frac{\varepsilon}{2}\Big(H(\mu_0 \,|\, \mathfrak{m}) + H(\mu_1 \,|\, \mathfrak{m})\Big) + \frac{1}{2}\int_{[0,1]\times\mathcal{X}} \Big(|v_t^{\mathrm{cu},\mathsf{P}^\varepsilon}|^2 + |v_t^{\mathrm{os},\mathsf{P}^\varepsilon}|^2\Big)\mathrm{d}t\mathrm{d}\mu_t^\varepsilon.$$

3. Proof of Proposition 1.5

We need to state some preliminary lemmas before completing the proof of Proposition 1.5 at page 14. Throughout the whole section we shall assume to work within $(\mathcal{X}, \mathsf{d}, \mathfrak{m})$ as in Setting 1.

Auxiliary lemmas. Let us start with an approximation result.

Lemma 3.1. *Let $\mu \in \mathcal{P}_2(\mathcal{X})$ with $H(\mu \,|\, \mathfrak{m}) < \infty$. Then:*

(a) there exists a sequence $(\mu_n) \subset \mathcal{P}_2(\mathcal{X})$ with $\mu_n = \rho_n \mathfrak{m}$ and $\rho_n \in C_c^\infty(\mathcal{X})$ such that $W_2(\mu_n, \nu) \to W_2(\mu, \nu)$ for all $\nu \in \mathcal{P}_2(\mathcal{X})$ and $H(\mu_n \,|\, \mathfrak{m}) \to H(\mu \,|\, \mathfrak{m})$ as $n \to \infty$;

(b) if in addition $I(\mu \,|\, \mathfrak{m}) < \infty$, then there exists a sequence $(\mu_n') \subset \mathcal{P}_2(\mathcal{X})$ with $\mu_n' = \rho_n \mathfrak{m}$, $\rho_n \in C_c^\infty(\mathcal{X})$ such that $W_2(\mu_n', \nu) \to W_2(\mu, \nu)$ for all $\nu \in \mathcal{P}_2(\mathcal{X})$, $H(\mu_n' \,|\, \mathfrak{m}) \to H(\mu \,|\, \mathfrak{m})$ and $I(\mu_n' \,|\, \mathfrak{m}) \to I(\mu \,|\, \mathfrak{m})$ as $n \to \infty$.

Proof. Let us write $\mu = \rho\, \mathfrak{m}$ and, as a first step, let us prove that both in (a) and (b) it is possible to find a sequence of measures with smooth densities converging to μ in the desired sense. This can be proved by defining for $\varepsilon > 0$

$$\mu_\varepsilon := \rho_\varepsilon\, \mathfrak{m} \qquad \text{with } \rho_\varepsilon := \mathcal{T}_\varepsilon \rho,$$

which clearly have smooth densities by the regularizing properties of $(\mathcal{T}_\varepsilon)$. The convergence of $W_2(\mu_\varepsilon, \nu)$, $H(\mu_\varepsilon \,|\, \mathfrak{m})$ and $I(\mu_\varepsilon \,|\, \mathfrak{m})$ to $W_2(\mu, \nu)$, $H(\mu \,|\, \mathfrak{m})$ and $I(\mu \,|\, \mathfrak{m})$ respectively as $\varepsilon \downarrow 0$ is now a well-known fact in the theory of gradient flows (see for instance Theorem 2.4.15 and Remark 2.4.16 in [1] in conjunction with the fact that the squared slope of the entropy is the Fisher information, as proved in [2]).

Thus, it is not restrictive to suppose that μ has smooth density. Under this new assumption, let us prove that we can find a sequence of measures with compact supports and smooth densities converging to μ in the desired sense. To this aim define

$$\mu_n := \alpha_n \rho_n\, \mathfrak{m} \qquad \text{with } \rho_n := \chi_n^2 \rho,$$

where α_n is the renormalization constant and χ_n is a smooth cut-off function with support in $B_{n+1}(x)$ (the ball of radius $n+1$ centered in x), for some $x \in \mathcal{X}$, $\chi_n(x) = 1$ and Lipschitz constant controlled by C/n, where $C > 1$ does not depend on n (see e.g. [3] for a proof of the existence of such cut-off functions). By dominated convergence it is not difficult to see that $W_2(\mu, \mu_n) \to 0$ and thus $W_2(\mu_n, \nu) \to W_2(\mu, \nu)$ for all $\nu \in \mathcal{P}_2(\mathcal{X})$ as $n \to \infty$; for the same reason $H(\mu_n \,|\, \mathfrak{m}) \to H(\mu \,|\, \mathfrak{m})$. If we also assume that $I(\mu \,|\, \mathfrak{m}) < \infty$, then

$$\frac{|\nabla \rho_n|^2}{\rho_n} \le 2\chi_n^2 \frac{|\nabla \rho|^2}{\rho} + 8\rho |\nabla \chi_n|^2 \le 2\frac{|\nabla \rho|^2}{\rho} + 8\frac{C^2}{n^2}\rho.$$

Since the right-hand side is integrable and $\chi_n \to 1$ as $n \to \infty$, by dominated convergence we get $I(\mu_n \,|\, \mathfrak{m}) \to I(\mu \,|\, \mathfrak{m})$.

Combining the two steps and using a diagonal argument, the conclusion follows. □

The following conservation result was pointed out in [6] and [24] in the case $\mathcal{X}$ is compact with different approaches; see also [8]. Following [24], we extend the statement to the present framework.

Lemma 3.2. *Under Assumptions 2.9, for any $\varepsilon > 0$ the function*

$$(0,1) \ni t \mapsto \int_{\mathcal{X}} |\nabla \vartheta_t^\varepsilon|^2 \, \mathrm{d}\mu_t^\varepsilon - \frac{\varepsilon^2}{4} I(\mu_t^\varepsilon \,|\, \mathfrak{m}) =: Q_t^\varepsilon$$

is real-valued and constant. Thus we shall denote it by Q^ε.

Proof. As a first step, for all $0 < t < 1$ by algebraic manipulation we have

$$|\nabla \vartheta_t^\varepsilon|^2 \rho_t^\varepsilon - \frac{\varepsilon^2}{4} |\nabla \log \rho_t^\varepsilon|^2 \rho_t^\varepsilon = -\langle \nabla \varphi_t^\varepsilon, \nabla \psi_t^\varepsilon \rangle \rho_t^\varepsilon = -\varepsilon^2 \langle \nabla f_t^\varepsilon, \nabla g_t^\varepsilon \rangle$$

so that

$$Q_t^\varepsilon = -\varepsilon^2 \int_{\mathcal{X}} \langle \nabla f_t^\varepsilon, \nabla g_t^\varepsilon \rangle \, \mathrm{d}\mathfrak{m}.$$

By Proposition 2.8 (b-iii) we know that $f^\varepsilon, g^\varepsilon$ are smooth with compact support, hence $|\nabla f^\varepsilon|, |\nabla g^\varepsilon| \in L^\infty(\mathfrak{m})$ and by the regularization properties of $\mathcal{T}^\varepsilon$ the same holds for $|\nabla f_t^\varepsilon|, |\nabla g_t^\varepsilon|$: this implies that $Q_t^\varepsilon \in \mathbb{R}$ for all $0 < t < 1$. As concerns the constancy of Q_t^ε, it is sufficient to prove that the right-hand side above is constant in t. To this aim, $\Gamma(f_t^\varepsilon, g_t^\varepsilon)$ is smooth both in time and space, as so are $f_t^\varepsilon, g_t^\varepsilon$; moreover $|\nabla f_t^\varepsilon|, |\nabla g_t^\varepsilon| \in L^\infty(\mathfrak{m})$ by what just said. Therefore, by dominated convergence and (2.10) we obtain

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathcal{X}} \langle \nabla f_t^\varepsilon, \nabla g_t^\varepsilon \rangle \, \mathrm{d}\mathfrak{m} &= \int_{\mathcal{X}} \Big(\langle \nabla \partial_t f_t^\varepsilon, \nabla g_t^\varepsilon \rangle + \langle \nabla f_t^\varepsilon, \nabla \partial_t g_t^\varepsilon \rangle \Big) \, \mathrm{d}\mathfrak{m} \\ &= \varepsilon \int_{\mathcal{X}} \Big(\langle \nabla L f_t^\varepsilon, \nabla g_t^\varepsilon \rangle - \langle \nabla f_t^\varepsilon, \nabla L g_t^\varepsilon \rangle \Big) \, \mathrm{d}\mathfrak{m}. \end{aligned}$$

From integration by parts formula it is straightforward to see that the right-hand side vanishes, whence the conclusion. □

Motivated by (2.13), let us investigate separately the convergence of current and osmotic velocities as $\varepsilon \downarrow 0$.

Lemma 3.3. *Under Assumptions 2.9, for any $\varepsilon > 0$ we have*

$$\lim_{\varepsilon \downarrow 0} \int_{[0,1]\times\mathcal{X}} |\nabla \vartheta_t^\varepsilon|^2 \, \mathrm{d}t \mathrm{d}\mu_t^\varepsilon = W_2^2(\mu_0, \mu_1), \quad \lim_{\varepsilon \downarrow 0} \int_{[0,1]\times\mathcal{X}} t |\nabla \vartheta_t^\varepsilon|^2 \, \mathrm{d}t \mathrm{d}\mu_t^\varepsilon = \frac{1}{2} W_2^2(\mu_0, \mu_1) \tag{3.4}$$

and

$$\lim_{\varepsilon \downarrow 0} \varepsilon^2 \int_0^1 I(\mu_t^\varepsilon \,|\, \mathfrak{m}) \, \mathrm{d}t = 0, \qquad \lim_{\varepsilon \downarrow 0} \varepsilon^2 \int_0^1 t I(\mu_t^\varepsilon \,|\, \mathfrak{m}) \, \mathrm{d}t = 0. \tag{3.5}$$

Proof. Let us first notice that combining (2.13) and (2.5) we get

$$\lim_{\varepsilon \downarrow 0} \Big(\int_{[0,1]\times\mathcal{X}} |\nabla \vartheta_t^\varepsilon|^2 \, \mathrm{d}t \mathrm{d}\mu_t^\varepsilon + \frac{\varepsilon^2}{4} \int_0^1 I(\mu_t^\varepsilon \,|\, \mathfrak{m}) \, \mathrm{d}t \Big) = W_2^2(\mu_0, \mu_1).$$

Since the continuity equation (2.12) is satisfied by μ_t^ε, the Benamou-Brenier formula holds for any $\varepsilon > 0$, that is

$$\int_{[0,1]\times\mathcal{X}} |\nabla \vartheta_t^\varepsilon|^2 \, \mathrm{d}t \mathrm{d}\mu_t^\varepsilon \geq W_2^2(\mu_0, \mu_1),$$

and together with the above limit, this leads us to the first identities both in (3.4) and (3.5). From them we immediately deduce that

$$\lim_{\varepsilon \downarrow 0} Q^\varepsilon = W_2^2(\mu_0, \mu_1) \tag{3.6}$$

and from the first one in (3.5)

$$0 \leq \lim_{\varepsilon \downarrow 0} \varepsilon^2 \int_0^1 t I(\mu_t^\varepsilon \,|\, \mathfrak{m}) \, \mathrm{d}t \leq \lim_{\varepsilon \downarrow 0} \varepsilon^2 \int_0^1 I(\mu_t^\varepsilon \,|\, \mathfrak{m}) \, \mathrm{d}t = 0,$$

whence also the second identity in (3.5). Finally observe that

$$\lim_{\varepsilon\downarrow 0}\int_{[0,1]\times\mathcal{X}} t|\nabla\vartheta_t^\varepsilon|^2\,\mathrm{d}t\mathrm{d}\mu_t^\varepsilon = \lim_{\varepsilon\downarrow 0}\Big(\int_{[0,1]\times\mathcal{X}} t|\nabla\vartheta_t^\varepsilon|^2\,\mathrm{d}t\mathrm{d}\mu_t^\varepsilon - \frac{\varepsilon^2}{4}\int_0^1 tI(\mu_t^\varepsilon\,|\,\mathfrak{m})\,\mathrm{d}t\Big)$$
$$= \lim_{\varepsilon\downarrow 0}\int_0^1 tQ^\varepsilon\mathrm{d}t = W_2^2(\mu_0,\mu_1)\int_0^1 t\,\mathrm{d}t = \frac{1}{2}W_2^2(\mu_0,\mu_1).$$

□

As already noticed at Remark 2.15, although it is smooth on the open interval $(0,1)$, the density ρ_t^ε might be arbitrarily close to zero. Consequently, the Schrödinger potentials might not be integrable enough, and the Fisher information $I(\mu_t^\varepsilon\,|\,\mathfrak{m})$ might behave badly around $t=0$ and $t=1$. Next lemma provides related regularity and growth controls. Its proof is strongly inspired by [12, 13, 14, 19]; we thus address the reader to these articles for more details.

Lemma 3.7. *Under Assumptions 2.9, let $\delta>0$ and set, for all $0\le t\le 1$, $\rho_t^{\varepsilon,\delta} := (f_t^\varepsilon+\delta)(g_t^\varepsilon+\delta)$ as well as*

$$\varphi_t^{\varepsilon,\delta} := \varepsilon\log(f_t^\varepsilon+\delta) \qquad \psi_t^{\varepsilon,\delta} := \varepsilon\log(g_t^\varepsilon+\delta) \qquad \vartheta_t^{\varepsilon,\delta} := \frac{1}{2}\big(\psi_t^{\varepsilon,\delta}-\varphi_t^{\varepsilon,\delta}\big).$$

Then all the functions so defined belong to $C^\infty([0,1]\times\mathcal{X})\cap\mathcal{W}^{1,2}(\mathcal{X},\mathfrak{m})$ and, as curves parametrized by t with values in $\mathcal{W}^{1,2}(\mathcal{X},\mathfrak{m})$, to $AC([0,1],\mathcal{W}^{1,2}(\mathcal{X},\mathfrak{m}))$. The time derivatives of $\varphi_t^{\varepsilon,\delta},\psi_t^{\varepsilon,\delta},\rho_t^{\varepsilon,\delta}$ are given by

$$\begin{gathered}\partial_t\varphi_t^{\varepsilon,\delta} = \frac{1}{2}|\nabla\varphi_t^{\varepsilon,\delta}|^2+\varepsilon L\varphi_t^{\varepsilon,\delta} \qquad \partial_t\psi_t^{\varepsilon,\delta} = \frac{1}{2}|\nabla\psi_t^{\varepsilon,\delta}|^2+\varepsilon L\psi_t^{\varepsilon,\delta}\\ \partial\rho_t^{\varepsilon,\delta}+\mathrm{div}_{\mathfrak{m}}(\nabla\vartheta_t^{\varepsilon,\delta}\,\rho_t^{\varepsilon,\delta})=0\end{gathered} \tag{3.8}$$

where $\partial_t\varphi_t^{\varepsilon,\delta}$, $\partial_t\psi_t^{\varepsilon,\delta}$, $\partial_t\rho_t^{\varepsilon,\delta}$ have to be understood both in the classical sense and as strong $\mathcal{W}^{1,2}$-limits.

Furthermore, defining $u(z):=z\log z$, the function $t\mapsto\int_{\mathcal{X}}u(\rho_t^{\varepsilon,\delta})\,\mathrm{d}\mathfrak{m}$ belongs to $C^2([0,1])$ and for every $t\in[0,1]$ it holds

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{X}}u(\rho_t^{\varepsilon,\delta})\,\mathrm{d}\mathfrak{m} = \int_{\mathcal{X}}\langle\nabla\rho_t^{\varepsilon,\delta},\nabla\vartheta_t^{\varepsilon,\delta}\rangle\,\mathrm{d}\mathfrak{m} \tag{3.9a}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\int_{\mathcal{X}}u(\rho_t^{\varepsilon,\delta})\,\mathrm{d}\mathfrak{m} = \int_{\mathcal{X}}\Big(\Gamma_2(\vartheta_t^{\varepsilon,\delta})+\frac{\varepsilon^2}{4}\Gamma_2(\log\rho_t^{\varepsilon,\delta})\Big)\rho_t^{\varepsilon,\delta}\,\mathrm{d}\mathfrak{m}. \tag{3.9b}$$

Proof. As already explained in Section 2, under Assumptions 2.9 $f_t^\varepsilon,g_t^\varepsilon\in C^\infty([0,1],\mathcal{X})\cap\mathcal{W}^{1,2}(\mathcal{X},\mathfrak{m})$ and $(f_t^\varepsilon),(g_t^\varepsilon)\in AC([0,1],\mathcal{W}^{1,2}(\mathcal{X},\mathfrak{m}))$. Therefore the regularity and integrability properties of $\varphi_t^{\varepsilon,\delta},\psi_t^{\varepsilon,\delta},\vartheta_t^{\varepsilon,\delta},\rho_t^{\varepsilon,\delta}$ are a straightforward consequence of the chain rule and of the fact that the logarithm is smooth with bounded derivatives on $[\delta,\infty)$. Also the PDEs solved by $\varphi_t^{\varepsilon,\delta},\psi_t^{\varepsilon,\delta},\rho_t^{\varepsilon,\delta}$ are easily deduced, when interpreted in the classical sense, as they follow from (2.11) and (2.12). In order to deduce the same identities with $\partial_t\varphi_t^{\varepsilon,\delta}$, $\partial_t\psi_t^{\varepsilon,\delta}$, $\partial_t\rho_t^{\varepsilon,\delta}$ seen as strong $\mathcal{W}^{1,2}$-limits, notice that by the maximum principle

$$\varepsilon\log\delta\le\varphi_t^{\varepsilon,\delta}\le\varepsilon\log(\|f^\varepsilon\|_{L^\infty(\mathfrak{m})}+\delta),\qquad\forall t\ge 0$$

whence $(\varphi_t^{\varepsilon,\delta}) \in L^\infty((0,\infty), L^\infty(\mathfrak{m}))$. Moreover, the smoothness of the logarithm, the chain and Leibniz rules entail that

$$|\nabla\varphi_t^{\varepsilon,\delta}| \le \varepsilon\frac{|\nabla f_t^\varepsilon|}{\delta} \qquad |\Delta\varphi_t^{\varepsilon,\delta}| \le \varepsilon\frac{|\Delta f_t^\varepsilon}{\delta} + \varepsilon\frac{|\nabla f_t^\varepsilon|^2}{\delta^2}$$

whence $(|\nabla\varphi_t^{\varepsilon,\delta}|^2), (\Delta\varphi_t^{\varepsilon,\delta}) \in L^\infty((0,1), \mathcal{W}^{1,2}(\mathcal{X},\mathfrak{m}))$; analogous estimates hold for $\nabla|\nabla\varphi_t^{\varepsilon,\delta}|^2$ and $\nabla\Delta\varphi_t^{\varepsilon,\delta}$. These bounds together with the fact that the PDEs in (3.8) hold in the classical sense imply, by a dominated convergence argument, that (3.8) are satisfied also as strong $\mathcal{W}^{1,2}$-limits.

Relying on that, (3.9a) and (3.9b) follow by the computations carried out in [19]. Indeed, the validity of (3.8) as strong $\mathcal{W}^{1,2}$-limits on $[0,1]$ ensures that $t \mapsto u(\rho_t^{\varepsilon,\delta})$ and $t \mapsto \langle\nabla\rho_t^{\varepsilon,\delta}, \nabla\vartheta_t^{\varepsilon,\delta}\rangle$ belong to $AC([0,1], L^2(\mathfrak{m}))$ and thus we can pass the time derivatives under the integral sign, i.e.

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{X}} u(\rho_t^{\varepsilon,\delta})\,\mathrm{d}\mathfrak{m} = \int_{\mathcal{X}} \partial_t u(\rho_t^{\varepsilon,\delta})\,\mathrm{d}\mathfrak{m}, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{X}} \langle\nabla\rho_t^{\varepsilon,\delta}, \nabla\vartheta_t^{\varepsilon,\delta}\rangle\,\mathrm{d}\mathfrak{m} = \int_{\mathcal{X}} \partial_t\langle\nabla\rho_t^{\varepsilon,\delta}, \nabla\vartheta_t^{\varepsilon,\delta}\rangle\,\mathrm{d}\mathfrak{m}. \tag{3.10}$$

Then the Hamilton-Jacobi-Bellman equations for the Schrödinger potentials and the continuity equation for the entropic interpolation together with $\varepsilon\log\rho_t^\varepsilon = \varphi_t^\varepsilon + \psi_t^\varepsilon$, here replaced by $\varepsilon\log\rho_t^{\varepsilon,\delta} = \varphi_t^{\varepsilon,\delta} + \psi_t^{\varepsilon,\delta}$, are the only tools needed to deduce (3.9a) and (3.9b). Finally, the fact that:

- $(\rho_t^{\varepsilon,\delta}), (\vartheta_t^{\varepsilon,\delta}) \in AC([0,1], \mathcal{W}^{1,2}(\mathcal{X},\mathfrak{m}))$;
- $(\rho_t^{\varepsilon,\delta}) \in AC([0,1], \mathcal{W}^{1,2}(\mathcal{X},\mathfrak{m}))$, $(|\nabla\vartheta_t^{\varepsilon,\delta}|^2), (\Delta\vartheta_t^{\varepsilon,\delta}), (|\nabla\log\rho_t^{\varepsilon,\delta}|^2), (\Delta\log\rho_t^{\varepsilon,\delta})$ belong to $L^\infty((0,1), \mathcal{W}^{1,2}(\mathcal{X},\mathfrak{m}))$ and $\vartheta_t^{\varepsilon,\delta}, \log\rho_t^{\varepsilon,\delta} \in C^\infty([0,1]\times\mathcal{X})$;

imply the continuity on $[0,1]$ of the right-hand sides of (3.9a) and (3.9b) respectively. □

With this results at disposal we can prove our main lemma: a rigorous 'entropic' analogue of Otto-Villani's heuristic argument.

Lemma 3.11. *Under Assumptions 2.9, for any $\varepsilon > 0$ it holds*

$$\begin{aligned} H(\mu_1\,|\,\mathfrak{m}) - H(\mu_0\,|\,\mathfrak{m}) \le \int_{\mathcal{X}} \langle\nabla\vartheta_1^\varepsilon, \nabla\rho_1\rangle\,\mathrm{d}\mathfrak{m} - \kappa\int_{[0,1]\times\mathcal{X}} t|\nabla\vartheta_t^\varepsilon|^2\,\mathrm{d}\mu_t^\varepsilon\mathrm{d}t \\ - \kappa\frac{\varepsilon^2}{4}\int_{[0,1]\times\mathcal{X}} t|\nabla\log\rho_t^\varepsilon|^2\,\mathrm{d}\mu_t^\varepsilon\mathrm{d}t. \end{aligned} \tag{3.12}$$

Proof. The proof of the lemma is based on the standard calculus identity

$$h(1) = h(0) + h'(1) - \int_0^1 t h''(t)\,\mathrm{d}t, \tag{3.13}$$

valid for any C^2-regular function h, applied to $t \mapsto \int_{\mathcal{X}} u(\rho_t^{\varepsilon,\delta})\,\mathrm{d}\mathfrak{m}$ defined as in Lemma 3.7. Plugging (3.9a) and (3.9b) into (3.13) and using the well-known inequalities

$$\Gamma_2(\vartheta_t^{\varepsilon,\delta}) \ge \kappa|\nabla\vartheta_t^{\varepsilon,\delta}|^2 \qquad \Gamma_2(\log\rho_t^{\varepsilon,\delta}) \ge \kappa|\nabla\log\rho_t^{\varepsilon,\delta}|^2,$$

consequence of the lower Ricci bounds (1.2) and (1.4) rewritten in the form of the Bochner-Lichnerowicz-Weitzenböck formula, we obtain

$$\int_{\mathcal{X}} u(\rho_1^{\varepsilon,\delta})\,\mathrm{d}\mathfrak{m} - \int_{\mathcal{X}} u(\rho_0^{\varepsilon,\delta})\,\mathrm{d}\mathfrak{m} \le \int_{\mathcal{X}} \langle \nabla\vartheta_1^{\varepsilon,\delta}, \nabla\rho_1^{\varepsilon,\delta}\rangle\,\mathrm{d}\mathfrak{m} - \kappa\int_{[0,1]\times\mathcal{X}} t|\nabla\vartheta_t^{\varepsilon,\delta}|^2\,\mathrm{d}\mu_t^{\varepsilon,\delta}\mathrm{d}t - \kappa\frac{\varepsilon^2}{4}\int_{[0,1]\times\mathcal{X}} t|\nabla\log\rho_t^{\varepsilon,\delta}|^2\,\mathrm{d}\mu_t^{\varepsilon,\delta}\mathrm{d}t, \tag{3.14}$$

where $\mu_t^{\varepsilon,\delta} := \rho_t^{\varepsilon,\delta}\mathfrak{m}$. It is now sufficient to pass to the limit as $\delta \downarrow 0$. By dominated convergence (recall that $f^\varepsilon, g^\varepsilon \in L^\infty(\mathfrak{m})$) it is easy to see that the left-hand side converges to $H(\mu_1\,|\,\mathfrak{m}) - H(\mu_0\,|\,\mathfrak{m})$. In order to apply the dominated convergence theorem also to the first term on the right-hand side, notice that

$$\nabla\rho_1^{\varepsilon,\delta} = \nabla\rho_1 + \delta\nabla f_1^\varepsilon + \delta\nabla g^\varepsilon \qquad \text{and} \qquad \nabla\vartheta_1^{\varepsilon,\delta} = \frac{1}{2}\nabla\psi_1^{\varepsilon,\delta} - \frac{1}{2}\nabla\varphi_1^{\varepsilon,\delta}.$$

Thus, to control $\langle\nabla\vartheta_1^{\varepsilon,\delta}, \nabla\rho_1^{\varepsilon,\delta}\rangle$ let us first observe that

$$|\langle\nabla\psi_1^{\varepsilon,\delta}, \nabla\rho_1\rangle| \le \varepsilon\frac{|\nabla g^\varepsilon||\nabla\rho_1|}{g^\varepsilon+\delta} \le \varepsilon\frac{|\nabla g^\varepsilon||\nabla\rho_1|}{g^\varepsilon} \le \varepsilon f_1^\varepsilon\frac{|\nabla g^\varepsilon|^2}{g^\varepsilon} + \varepsilon|\nabla g^\varepsilon||\nabla f_1^\varepsilon|$$

and remark that the right-hand side is integrable by Proposition 2.8-(c); secondly

$$|\langle\nabla\varphi_1^{\varepsilon,\delta}, \nabla\rho_1\rangle| \le \varepsilon\frac{|\nabla f_1^\varepsilon||\nabla\rho_1|}{f_1^\varepsilon} \le \varepsilon g^\varepsilon\frac{|\nabla f_1^\varepsilon|^2}{f_1^\varepsilon} + \varepsilon|\nabla g^\varepsilon||\nabla f_1^\varepsilon|$$

and in this case the right-hand side is integrable as g^ε has compact support and f_1^ε is bounded away from 0 therein; finally,

$$\delta|\langle\nabla\psi_1^{\varepsilon,\delta}, \nabla f_1^\varepsilon\rangle| \le \varepsilon\delta\frac{|\nabla g^\varepsilon||\nabla f_1^\varepsilon|}{g^\varepsilon+\delta} \le \varepsilon|\nabla g^\varepsilon||\nabla f_1^\varepsilon|$$

as $g^\varepsilon+\delta \ge \delta$ and the same strategy applies to all the remaining terms that we obtain developing $\langle\nabla\vartheta_1^{\varepsilon,\delta}, \nabla\rho_1^{\varepsilon,\delta}\rangle$. Therefore, the first term on the right-hand side of (3.14) converges to the first term on the right-hand side of (3.12). As regards the other two summands, by the very definition of $\vartheta_t^{\varepsilon,\delta}$ and since

$$\varepsilon\log\rho_t^{\varepsilon,\delta} = \varphi_t^{\varepsilon,\delta} + \psi_t^{\varepsilon,\delta},$$

the conclusion will follow if we are able to prove that

$$\lim_{\delta\downarrow 0}\int_{[0,1]\times\mathcal{X}} |\nabla\varphi_t^{\varepsilon,\delta}|^2\,\mathrm{d}\mu_t^{\varepsilon,\delta}\mathrm{d}t = \int_{[0,1]\times\mathcal{X}} |\nabla\varphi_t^{\varepsilon}|^2\,\mathrm{d}\mu_t^{\varepsilon}\mathrm{d}t, \tag{3.15a}$$

$$\lim_{\delta\downarrow 0}\int_{[0,1]\times\mathcal{X}} |\nabla\psi_t^{\varepsilon,\delta}|^2\,\mathrm{d}\mu_t^{\varepsilon,\delta}\mathrm{d}t = \int_{[0,1]\times\mathcal{X}} |\nabla\psi_t^{\varepsilon}|^2\,\mathrm{d}\mu_t^{\varepsilon}\mathrm{d}t. \tag{3.15b}$$

To this aim, notice that $\rho_t^{\varepsilon,\delta} = \rho_t^\varepsilon + \delta f_t^\varepsilon + \delta g_t^\varepsilon + \delta^2$, whence using either $f_t^\varepsilon + \delta \geq f_t^\varepsilon$ or $f_t^\varepsilon + \delta \geq \delta$ it is easy to infer that

$$\begin{aligned}
|\nabla \varphi_t^{\varepsilon,\delta}|^2 \rho_t^\varepsilon &= \varepsilon^2 \frac{|\nabla f_t^\varepsilon|^2}{(f_t^\varepsilon + \delta)^2} \rho_t^\varepsilon \leq \varepsilon^2 \frac{|\nabla f_t^\varepsilon|^2}{(f_t^\varepsilon)^2} \rho_t^\varepsilon = |\nabla \varphi_t^\varepsilon|^2 \rho_t^\varepsilon, \\
\delta |\nabla \varphi_t^{\varepsilon,\delta}|^2 f_t^\varepsilon &= \delta \varepsilon^2 \frac{|\nabla f_t^\varepsilon|^2}{(f_t^\varepsilon + \delta)^2} f_t^\varepsilon \leq \varepsilon^2 |\nabla f_t^\varepsilon|^2, \\
\delta |\nabla \varphi_t^{\varepsilon,\delta}|^2 g_t^\varepsilon &= \delta \varepsilon^2 \frac{|\nabla f_t^\varepsilon|^2}{(f_t^\varepsilon + \delta)^2} g_t^\varepsilon \leq \varepsilon^2 \frac{|\nabla f_t^\varepsilon|^2}{f_t^\varepsilon} g_t^\varepsilon = |\nabla \varphi_t^\varepsilon|^2 \rho_t^\varepsilon, \\
\delta^2 |\nabla \varphi_t^\varepsilon|^2 &= \delta^2 \varepsilon^2 \frac{|\nabla f_t^\varepsilon|^2}{(f_t^\varepsilon + \delta)^2} \leq \varepsilon^2 |\nabla f_t^\varepsilon|^2.
\end{aligned}$$

All the right-hand sides above are integrable on $[0,1] \times \mathcal{X}$ (either by (2.14) or by the Bakry-Émery contraction estimate (2.2) together with $f^\varepsilon \in C_c^\infty(\mathcal{X})$), thus by dominated convergence (3.15a) follows. An analogous argument holds for (3.15b), whence the conclusion. □

Completion of the proof of Proposition 1.5. We are now ready to complete the proof of the HWI inequality.

Proof of Proposition 1.5. First of all, by Proposition 2.8 and Lemma 3.1, it is sufficient to prove the result for any μ_0, μ_1 satisfying Assumptions 2.9. This allows us to invoke Lemma 3.11. Secondly, observe that by the Cauchy-Schwarz inequality

$$\int_{\mathcal{X}} \langle \nabla \vartheta_1^\varepsilon, \nabla \rho_1 \rangle \, \mathrm{d}\mathfrak{m} = \int_{\mathcal{X}} \langle \nabla \vartheta_1^\varepsilon, \nabla \log \rho_1 \rangle \mathrm{d}\mu_1 \leq \sqrt{\int_{\mathcal{X}} |\nabla \vartheta_1^\varepsilon|^2 \mathrm{d}\mu_1} \sqrt{I(\mu_1 \,|\, \mathfrak{m})},$$

so that if we plug this information into (3.12) we obtain

$$\begin{aligned}
H(\mu_1 \,|\, \mathfrak{m}) - H(\mu_0 \,|\, \mathfrak{m}) \leq \sqrt{\int_{\mathcal{X}} |\nabla \vartheta_1^\varepsilon|^2 \mathrm{d}\mu_1} \sqrt{I(\mu_1 \,|\, \mathfrak{m})} &- \kappa \int_{[0,1]\times\mathcal{X}} t |\nabla \vartheta_t^\varepsilon|^2 \, \mathrm{d}\mu_t^\varepsilon \mathrm{d}t \\
&- \kappa \frac{\varepsilon^2}{4} \int_{[0,1]\times\mathcal{X}} t |\nabla \log \rho_t^\varepsilon|^2 \, \mathrm{d}\mu_t^\varepsilon \mathrm{d}t.
\end{aligned}$$

Let us now pass to the limit as $\varepsilon \downarrow 0$: by Lemma 3.2, at $t = 1$ we have

$$Q^\varepsilon = \int_{\mathcal{X}} |\nabla \vartheta_1^\varepsilon|^2 \, \mathrm{d}\mu_1 - \frac{\varepsilon^2}{4} I(\mu_1 \,|\, \mathfrak{m}),$$

so that (3.6) together with $I(\mu_1 \,|\, \mathfrak{m}) < \infty$ yields

$$\lim_{\varepsilon \downarrow 0} \int_{\mathcal{X}} |\nabla \vartheta_1^\varepsilon|^2 \mathrm{d}\mu_1 = W_2^2(\mu_0, \mu_1).$$

By Lemma 3.3 we can pass to the limit as $\varepsilon \downarrow 0$ also in the remaining terms on the right-hand side, thus concluding. □

Let us mention that another approach to prove the HWI inequality via the Schrödinger problem is pointed out in a recent work [7] where an entropic counterpart of the HWI inequality is formally obtained by differentiating the convexity estimate of the entropy along the entropic interpolations introduced in [6, Thm. 1.4].

4. Final remarks and comments

Lagrangian and Hamiltonian interpretation. From a heuristic point of view, the expression of the constant quantity Q^ε can be deduced by standard arguments in Lagrangian and Hamiltonian formalism. Indeed, motivated by (2.13) let us consider the action functional

$$\mathscr{A}(\nu, v) = \int_{[0,1]\times\mathcal{X}} \Big(\frac{|v_t|^2}{2} + \frac{\varepsilon^2}{8}|\nabla \log \nu_t|^2\Big)\nu_t \,\mathrm{d}t\mathrm{d}\mathfrak{m} \tag{4.1}$$

associated to the Lagrangian

$$\mathscr{L}(\nu, v) = \int_{\mathcal{X}} \Big(\frac{|v|^2}{2}\nu + \frac{\varepsilon^2}{8}\frac{|\nabla \nu|^2}{\nu}\Big)\,\mathrm{d}\mathfrak{m}.$$

By means of Legendre's transform, the corresponding Hamiltonian is given by

$$\mathscr{H}(\nu, p) = \int_{\mathcal{X}} \Big(\frac{|p|^2}{2\nu} - \frac{\varepsilon^2}{8}\frac{|\nabla \nu|^2}{\nu}\Big)\,\mathrm{d}\mathfrak{m}$$

and, at least formally, $\mathscr{H}$ is constant along the critical points of $\mathscr{A}$. Since μ_0 and μ_1 are prescribed, the Euler-Lagrange equation for (4.1) reads as

$$\begin{aligned}
&\partial_t \nu_t + \mathrm{div}_{\mathfrak{m}}(\nu_t v_t) = 0\\
&\partial_t v_t + \frac{|v_t|^2}{2} = -\frac{\varepsilon^2}{8}\big(2\Delta \log \nu_t + |\nabla \log \nu_t|^2\big)
\end{aligned}$$

and, as it is not difficult to see (e.g. following the computations carried out in [13] which fit to Setting 1), these PDEs are satisfied along $(\rho_t^\varepsilon, \vartheta_t^\varepsilon)$. Finally, as in the Hamiltonian p represents a momentum density, it is natural to set $p_t := \nu_t v_t$. From these considerations, the guess on the existence of the conserved quantity of Lemma 3.2 and on its expression follows.
This point of view is particularly investigated in the recent paper [8], to which we refer for more details.

The compact case. As already mentioned in Remark 2.15, in general Schrödinger potentials fail to be smooth in $t = 0$ or $t = 1$ and are not even integrable for any $0 < t < 1$, because $f_t^\varepsilon, g_t^\varepsilon$ can be arbitrarily close to 0. This problem can be overcome if $f^\varepsilon, g^\varepsilon \geq c > 0$ for some constant c: then $f_t^\varepsilon, g_t^\varepsilon \geq c$ as well by the maximum principle and $\varphi_t^\varepsilon, \psi_t^\varepsilon \geq \varepsilon \log c$ from their very definition. However, even if we assume μ_0, μ_1 to have densities bounded away from 0 (thus forcing $\mathfrak{m}$ to belong to $\mathcal{P}_2(\mathcal{X})$, a condition which does not follow from our Setting 1, unless $\kappa > 0$, cf. Remark 1.7-(a)), it is not known yet whether this implies an analogous lower bound on f^ε and g^ε. This is the motivation behind the definitions provided in Lemma 3.7.

On the contrary, if $\mathcal{X}$ is assumed to be compact, then it is not restrictive to assume $\mu_0, \mu_1 \geq c\mathfrak{m}$ for some $c > 0$: indeed any $\mu \in \mathcal{P}(\mathcal{X})$ can be approximated by

$$\mu_n := \alpha_n \rho_n \mathfrak{m} \qquad \text{with } \rho_n := \max\{\rho, 1/n\}$$

(where α_n is the renormalization constant) in W_2, entropy and Fisher information in the sense of Lemma 3.1. Secondly, r_ε is bounded from above since $\mathsf{r}_\varepsilon \in C^\infty(\mathcal{X})$. These two facts imply that $f^\varepsilon, g^\varepsilon \geq c' > 0$, as shown in [12]. As a consequence the proof of the HWI inequality is easier and the parallelism with the heuristic proof of Otto-Villani [23]

is stronger, since no 'δ-argument' in Lemma 3.7 and Lemma 3.11 is needed; in particular, Lemma 3.7 is already known to hold by [12] and thus no proof is required.

RCD spaces. All the results presented in this paper, with particular mention of the key Lemma 3.11 and Proposition 1.5, are also true in a more general framework than the one of Setting 1-(b), namely in $\mathsf{RCD}^*(K,N)$ spaces (introduced in [11]). If $\mathcal{X}$ is assumed to be compact, then this has been shown by the last author in his PhD thesis [24]. When $\mathcal{X}$ is not compact, the most important steps in our entropic approach still hold. Namely:

- all the regularity and integrability results concerning Schrödinger potentials and entropic interpolations mentioned in Section 2 as well as the dynamic representation of the entropic cost;
- the regularizing and contraction properties of $(\mathcal{T}_t)$;
- the existence of 'good' cut-off functions;
- the Benamou-Brenier formula and the Bochner-Lichnerowicz-Weitzenböck inequality.

The reader is addressed to [13, 14] for the first point and to [10] for all the others.

Acknowledgements. This research was supported by the French ANR-17-CE40-0030 EFI project and the LABEX MILYON ANR-10-LABX-0070.

References

[1] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, second ed., 2008.

[2] ——, *Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below*, Invent. Math., 195 (2014), pp. 289–391.

[3] D. Azagra, J. Ferrera, F. López-Mesas, and Y. Rangel, *Smooth approximation of Lipschitz functions on Riemannian manifolds*, J. Math. Anal. Appl., 326 (2007), pp. 1370–1378.

[4] D. Bakry, I. Gentil, and M. Ledoux, *Analysis and geometry of Markov diffusion operators*, vol. 348 of Grundlehren der Mathematischen Wissenschaften, Springer, Cham, 2014.

[5] Y. Chen, T. Georgiou, and M. Pavon, *On the relation between optimal transport and Schrödinger bridges: A stochastic control viewpoint*, Journal of Optimization Theory and Applications, 169 (2016), pp. 671–691.

[6] G. Conforti, *A second order equation for Schrödinger bridges with applications to the hot gas experiment and entropic transportation cost*. To appear in Probab. Theory Relat. Fields, (2018).

[7] G. Conforti and L. Ripani, *Around the entropic Talagrand inequality*. Preprint, arXiv:1809.02062 (2018).

[8] I. Gentil, C. Léonard, and L. Ripani, *Dynamical aspects of generalized Schrödinger problem via Otto calculus - A heuristic point of view*. Preprint, arXiv:1806.01553 (2018).

[9] ——, *About the analogy between optimal transport and minimal entropy*, Ann. Fac. Toulouse, 26 (2017), pp. 569–600.

[10] N. Gigli, *Nonsmooth differential geometry - an approach tailored for spaces with Ricci curvature bounded from below*. Mem. Am. Math. Soc. 1196, iii-vi, 166 p. 2018.

[11] ——, *On the differential structure of metric measure spaces and applications*, Mem. Amer. Math. Soc., 236 (2015), pp. vi+91.

[12] N. Gigli and L. Tamanini, *Second order differentiation formula on compact $RCD^*(K,N)$ spaces*. Preprint, arXiv:1701.03932, (2017).

[13] ——, *Second order differentiation formula on $RCD^*(K,N)$ spaces*. To appear in J. Eur. Math. Soc. (2018).

[14] ——, *Benamou-Brenier and duality formulas for the entropic cost on $RCD^*(K,N)$ spaces*. To appear in Probab. Theory Relat. Fields (2018).

[15] A. Grigoryan, *Heat kernel and analysis on manifolds*, vol. 47, American Mathematical Soc., 2009.

[16] C. Léonard, *From the Schrödinger problem to the Monge-Kantorovich problem*, J. Funct. Anal., 262 (2012), pp. 1879–1920.
[17] ——, *Some properties of path measures*, Séminaire de Probabilités XLVI, (2014), pp. 207–230.
[18] ——, *A survey of the Schrödinger problem and some of its connections with optimal transport*, Discrete Contin. Dyn. Syst., 34 (2014), pp. 1533–1574.
[19] ——, *On the convexity of the entropy along entropic interpolations*, in Measure Theory in Non-Smooth Spaces, N. Gigli, ed., Partial Differential Equations and Measure Theory, De Gruyter Open, 2017.
[20] T. Mikami, *Monge's problem with a quadratic cost by the zero-noise limit of h-path processes*, Probab. Theory Relat. Fields 129 (2004), pp. 245–260.
[21] E. Nelson, *Dynamical theories of Brownian motion*, Princeton University Press, Princeton, N.J., 1967.
[22] F. Otto, *The geometry of dissipative evolution equations: the porous medium equation*, Comm. Partial Differential Equations, 26 (2001), pp. 101–174.
[23] F. Otto and C. Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal., 173 (2000), pp. 361–400.
[24] L. Tamanini, *Analysis and geometry of RCD spaces via the Schrödinger problem*, PhD thesis, Université Paris Nanterre and SISSA, 2017.
[25] C. Villani, *Optimal transport. Old and new*, vol. 338 of Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin, 2009.

Ivan Gentil. Univ Lyon, Université Claude Bernard Lyon 1, CNRS UMR 5208, Institut Camille Jordan, F-69622 Villeurbanne, France
E-mail address: gentil@math.univ-lyon1.fr

Christian Léonard. Modal-X. Université Paris Nanterre. F-92001 Nanterre, France
E-mail address: leonard@parisnanterre.fr

Luigia Ripani. Univ Lyon, Université Claude Bernard Lyon 1, CNRS UMR 5208, Institut Camille Jordan, F-69622 Villeurbanne, France
E-mail address: ripani@math.univ-lyon1.fr

Luca Tamanini. Institut für Angewandte Mathematik. Universität Bonn. Bonn, Germany
E-mail address: tamanini@iam.uni-bonn.de