Journal of Functional Analysis 285 (2023) 110153

Contents lists available at ScienceDirect

Journal of Functional Analysis

journal homepage: www.elsevier.com/locate/jfa

Full Length Article

Density of subalgebras of Lipschitz functions in metric Sobolev spaces and applications to Wasserstein Sobolev spaces

Massimo Fornasier [a], Giuseppe Savaré [b,*], Giacomo Enrico Sodini [c]

[a] *TUM Fakultät für Mathematik, Boltzmannstrasse 3, 85748 Garching bei München, Germany*
[b] *Bocconi University, Department of Decision Sciences and BIDSA, Via Roentgen 1, 20136 Milano, Italy*
[c] *TUM Fakultät für Mathematik, Boltzmannstrasse 3, 85748 Garching bei München, Germany*

ARTICLE INFO

Article history:
Received 2 September 2022
Accepted 21 August 2023
Available online 30 August 2023
Communicated by K.-T. Sturm

MSC:
primary 46E36, 31C25
secondary 49Q20, 28A33

Keywords:
Metric Sobolev spaces
Dirichlet forms
Cheeger energy
Kantorovich-Wasserstein distance
Optimal transport
Moreau-Yosida regularization

ABSTRACT

We prove a general criterion for the density in energy of suitable subalgebras of Lipschitz functions in the metric Sobolev space $H^{1,p}(X,\mathsf{d},\mathfrak{m})$ associated with a positive and finite Borel measure $\mathfrak{m}$ in a separable and complete metric space (X,d).
We then provide a relevant application to the case of the algebra of cylinder functions in the Wasserstein Sobolev space $H^{1,2}(\mathcal{P}_2(\mathbb{M}),W_2,\mathfrak{m})$ arising from a positive and finite Borel measure $\mathfrak{m}$ on the Kantorovich-Rubinstein-Wasserstein space $(\mathcal{P}_2(\mathbb{M}),W_2)$ of probability measures in a finite dimensional Euclidean space, a complete Riemannian manifold, or a separable Hilbert space $\mathbb{M}$. We will show that such a Sobolev space is always Hilbertian, independently of the choice of the reference measure $\mathfrak{m}$ so that the resulting Cheeger energy is a Dirichlet form.
We will eventually provide an explicit characterization for the corresponding notion of $\mathfrak{m}$-Wasserstein gradient, showing

* Corresponding author.
E-mail addresses: massimo.fornasier@ma.tum.de (M. Fornasier), giuseppe.savare@unibocconi.it (G. Savaré), sodini@ma.tum.de (G.E. Sodini).

https://doi.org/10.1016/j.jfa.2023.110153

useful calculus rules and its consistency with the tangent bundle and the Γ-calculus inherited from the Dirichlet form.

Contents

1. Introduction

The theory of Sobolev spaces associated to a metric measure space $(X, \mathsf{d}, \mathfrak{m})$ has been much developed in recent years. One of the most important approaches (we refer to the monographs [12,28] and to the lecture notes [26,44]) is based on the notion of upper gradient [27,29] of a map $f : X \to \mathbb{R}$: it is a Borel map $g : X \to [0, +\infty]$ satisfying

$$|f(\gamma(b)) - f(\gamma(a))| \leq \int_\gamma g \tag{1.1}$$

along every d-Lipschitz (or even rectifiable) curve $\gamma : [a, b] \to X$. The Dirichlet space $D^{1,p}(X, \mathsf{d}, \mathfrak{m})$, $p \in (1, +\infty)$ can then be defined as the class of measurable functions $f : X \to \mathbb{R}$ that possess a p-integrable upper gradient, thus resulting in a finite Newtonian energy

$$\mathsf{NE}_p(f) := \inf\Big\{ \int_X g^p \, \mathrm{d}\mathfrak{m} : g \text{ is an upper gradient of } f \Big\}. \tag{1.2}$$

A crucial and nontrivial fact is that NE_p admits a local representation

$$\mathsf{NE}_p(f) = \int_X |\mathrm{D}f|_N^p \,\mathrm{d}\mathfrak{m} \tag{1.3}$$

in terms of the *minimal p-weak upper gradient* $|\mathrm{D}f|_N^p$ of f, which can be characterized in terms of the upper gradient property (1.1) along Mod_p-a.e. curve and enjoys many nice metric-nonsmooth calculus rules. When f is Lipschitz, then the pointwise and asymptotic Lipschitz constants

$$|\mathrm{D}f|(x) := \limsup_{y\to x} \frac{|f(y)-f(x)|}{\mathsf{d}(x,y)}, \quad \operatorname{lip} f(x) := \limsup_{y,z\to x,\ y\neq z} \frac{|f(y)-f(z)|}{\mathsf{d}(y,z)} \tag{1.4}$$

are upper gradients, so that

$$\text{if } f \in \mathrm{Lip}_b(X) \text{ then} \quad |\mathrm{D}f|_N \le |\mathrm{D}f| \le \operatorname{lip} f \quad \mathfrak{m}\text{-a.e. in } X. \tag{1.5}$$

Functions in $L^p(X,\mathfrak{m})$ which admit a good representative (in the usual Lebesgue class defined up to $\mathfrak{m}$-negligible sets) in $D^{1,p}(X,\mathsf{d},\mathfrak{m})$ give raise to the Newtonian spaces $\hat{N}^{1,p}(X,\mathsf{d},\mathfrak{m})$ [45] [12, Def. 1.19], which is a Banach space with the norm $\|f\|_{\hat{N}^{1,p}} := \big(\|f\|_{L^p}^p + \mathsf{NE}_p(f)\big)^{1/p}$. $\hat{N}^{1,p}(X,\mathsf{d},\mathfrak{m})$ can also be identified with the domain of the L^p-relaxation of NE_p [15].

Density of Lipschitz functions: the case of doubling spaces supporting a Poincaré inequality It is a natural question if NE_p can be recovered starting from the distinguished class of upper gradients given by the pointwise or asymptotic Lipschitz constants (1.4) of Lipschitz functions. When $(X,\mathsf{d},\mathfrak{m})$ satisfies a doubling condition and supports a p-Poincaré inequality, then Lipschitz functions are dense in $\hat{N}^{1,p}(X,\mathsf{d},\mathfrak{m})$ [45, Theorem 4.1] and for Lipschitz functions the minimal p-weak upper gradient $|\mathrm{D}f|_N$ coincides with the pointwise Lipschitz constant $|\mathrm{D}f|$ [15, Theorem 6.1]. In particular for every $f \in \hat{N}^{1,p}(X,\mathsf{d},\mathfrak{m})$ there exists a sequence $f_n \in \mathrm{Lip}_b(X)$ such that

$$f_n \to f, \quad |\mathrm{D}f_n| \to |\mathrm{D}f|_N \quad \text{strongly in } L^p(X,\mathfrak{m}). \tag{1.6}$$

It is worth noticing that in this case $\hat{N}^{1,p}(X,\mathsf{d},\mathfrak{m})$ is a reflexive space [15, Theorem 4.48].

Density in energy of subalgebras of Lipschitz functions The strong approximation property (1.6) holds in fact for arbitrary complete and separable metric spaces, a result obtained in [7] (when $p = 2$) and [5], where also the approximation by the asymptotic Lipschitz constant is considered.

The first aim of the present paper is to discuss the extension of this result when the sequence f_n in (1.6) is chosen in a suitable unital subalgebra $\mathscr{A} \subset \mathrm{Lip}_b(X)$ separating the points of X, i.e.

$$1 \in \mathscr{A}, \quad \text{for every } x_0, x_1 \in X \text{ there exists } f \in \mathscr{A}\colon \quad f(x_0) \neq f(x_1). \tag{1.7}$$

The use of a subalgebra is not a formal exercise with no implications. In fact, in relevant examples such as Wasserstein Sobolev spaces, which we shall introduce and discuss below, and the related subalgebra of cylinder functions, it is possible to recover the Cheeger energy as suitable relaxation of an *explicitly computable Dirichlet form.* Namely, from the so-called pre-Cheeger energy

$$\mathsf{pCE}_p(f) := \int_X (\operatorname{lip} f)^p \,\mathrm{d}\mathfrak{m}, \quad f \in \operatorname{Lip}_b(X), \tag{1.8}$$

we recover the Cheeger energy as its relaxation starting from $\mathscr{A}$:

$$\mathsf{CE}_{p,\mathscr{A}}(f) = \inf \left\{ \liminf_{n\to+\infty} \mathsf{pCE}_p(f_n) : f_n \in \mathscr{A},\ f_n \to f \text{ in } L^0(X,\mathfrak{m}) \right\}. \tag{1.9}$$

Thanks to the algebraic properties of $\mathscr{A}$ and (1.7) it is possible to prove [44, Sec. 3] that $\mathsf{CE}_{p,\mathscr{A}}$ admits a local representation of the form

$$\mathsf{CE}_{p,\mathscr{A}}(f) = \int_X |\mathrm{D}f|^p_{\star,\mathscr{A}}(x)\,\mathrm{d}\mathfrak{m}(x) \quad \text{whenever } \mathsf{CE}_{p,\mathscr{A}}(f) < +\infty, \tag{1.10}$$

in terms of a minimal $(p,\mathscr{A})$-relaxed gradient $|\mathrm{D}f|_{\star,\mathscr{A}}$, enjoying the same calculus rules as $|\mathrm{D}f|_N$, see Theorem 2.3 below. We denote by $H^{1,p}(X,\mathsf{d},\mathfrak{m};\mathscr{A})$ the class of functions in $L^p(X,\mathfrak{m})$ with finite $(p,\mathscr{A})$-Cheeger energy.

It is easy to check that $H^{1,p}(X,\mathsf{d},\mathfrak{m};\mathscr{A}) \subset \hat{N}^{1,p}(X,\mathsf{d},\mathfrak{m})$ with $|\mathrm{D}f|_N \le |\mathrm{D}f|_{\star,\mathscr{A}}$ $\mathfrak{m}$-a.e. It turns out that the strong approximation property

$$\text{for every } f \in \hat{N}^{1,p}(X,\mathsf{d},\mathfrak{m}) \text{ there exist } f_n \in \mathscr{A}: \quad f_n \to f,\ \operatorname{lip} f_n \to |\mathrm{D}f|_N \text{ in } L^p(X,\mathfrak{m}) \tag{1.11}$$

is equivalent to the identification

$$H^{1,p}(X,\mathsf{d},\mathfrak{m};\mathscr{A}) = \hat{N}^{1,p}(X,\mathsf{d},\mathfrak{m}), \quad |\mathrm{D}f|_N = |\mathrm{D}f|_{\star,\mathscr{A}} \quad \text{for every } f \in \hat{N}^{1,p}(X,\mathsf{d},\mathfrak{m}). \tag{1.12}$$

The density results of [7,5] show that (1.12) always hold if $\mathscr{A} = \operatorname{Lip}_b(X)$. When $\mathscr{A}$ is a proper subalgebra of $\operatorname{Lip}_b(X)$, a first sufficient condition for the validity of (1.12), in the more general framework of extended topological metric measure spaces, is provided by the compatibility condition between d and $\mathscr{A}$ [44, Theorems 3.2.7, 5.3.1]

$$\mathsf{d}(x,y) = \sup \Big\{ f(x) - f(y) : f \in \mathscr{A},\ \operatorname{Lip}(f,X) \le 1 \Big\}. \tag{1.13}$$

We are able to improve (1.13) and to show (Theorem 2.13) that a necessary and sufficient condition for (1.12) is that for every $y \in X$ (or in a dense subset of X) the distance function $\mathsf{d}_y : x \mapsto \mathsf{d}(x,y)$ satisfies

$$|\mathrm{D}\mathsf{d}_y|_{\star,\mathscr{A}}(x) \le 1 \quad \text{for } \mathfrak{m}\text{-a.e. } x \in X. \tag{1.14}$$

As mentioned above, the density of distinguished subalgebras of Lipschitz functions can provide valuable information on the structure of the metric Sobolev space $\hat{N}^{1,p}(X, \mathsf{d}, \mathfrak{m})$, in particular when the asymptotic Lipschitz constant $\operatorname{lip} f$ exhibits a more regular behavior when restricted to $\mathscr{A}$. A relevant example arises when an algebra $\mathscr{A}$ exists for which the pre-Cheeger energy (1.8) is induced by a bilinear form and one wants to study its closure as it is typical in the theory of Dirichlet forms. In this case our result shows that this construction is intrinsically linked to the metric structure, so that it is independent of the particular choice of the algebra $\mathscr{A}$ satisfying (1.14) and it is invariant with respect to measure-preserving isometries. As a byproduct, we will recover in a simple way previous Hilbertianity results of [21,31,44].

The Wasserstein Sobolev space An important application, which has been one of the inspiring motivations of our investigation, concerns functional analysis over spaces of probability measures. In fact, smooth functions do appear recently as solutions of new types of partial differential equations over spaces of probability measures defined by diverse forms of differentiation, namely nonlinear transport equations [2,3] for describing population evolutionary games, Kolmogorov equations [35,34] in nonlinear filtering, and Hamilton-Jacobi-Bellman equations [11,23,39,14] as appearing, e.g., in the theory of mean-field games and mean-field optimal control. In some of these instances, the solutions are considered in classical sense, because of the lack of weak formulations and variational descriptions. Moreover, while the expression of these equations is in most cases of foundational interest, their relevance in terms of providing insights about solutions and their explicit computation remained so far rather unclear. Hence, a proper definition of function spaces of regular functions, rules of calculus, and density properties are fundamental for developing a more systematic framework for the analysis of such novel forms of infinite dimensional PDEs and explaining their practical use and impact.

Moreover, thanks to significant advances in computational optimal transport that made its numerical realization feasible also for problems of relatively high dimension, in the past decade there has been an increasing and more accepted adoption of probability measures to model data points in image and shape processing and other machine learning applications. While the first applications were about discriminating data encoded as distributions available in the form of bags-of-features or descriptors, more recent developments explored geometric interpolation of data provided by optimal transport, for instance in the form of Wasserstein barycenters. In the meanwhile the literature on the subject has grown significantly to be really able to offer a complete account and we may more simply refer to the recent survey [38] for insights and references.

Building upon these advances, approximating or interpolating efficiently functions over data points modeled as (probability) measures can also provide a novel framework for machine learning tasks, such as classification and regression. Also for such developments a proper foundation of functional analysis is necessary.

These are relevant motivations for us to focus on the study of Sobolev spaces generated by a finite measure $\mathfrak{m}$ on the Wasserstein space $\mathcal{P}_2(\mathbb{M})$ of Borel probability measures in a complete Riemannian manifold $(\mathbb{M}, \mathsf{d}_{\mathbb{M}})$ with finite quadratic moment

$$\int_{\mathbb{M}} \mathsf{d}_{\mathbb{M}}^2(x, x_o)\,\mathrm{d}\mu(x) < +\infty \quad \text{for some, and thus any, } x_o \in \mathbb{M}, \tag{1.15}$$

endowed with the L^2-Kantorovich-Rubinstein-Wasserstein distance W_2

$$W_{2,\mathsf{d}_{\mathbb{M}}}^2(\mu,\nu) := \min\left\{ \int_{\mathbb{M}\times\mathbb{M}} \mathsf{d}_{\mathbb{M}}^2(x,y)\,\mathrm{d}\boldsymbol{\mu}(x,y) \mid \boldsymbol{\mu} \in \Gamma(\mu,\nu) \right\}; \tag{1.16}$$

here $\Gamma(\mu,\nu)$ is the set of couplings between μ and ν, i.e. probability measures $\boldsymbol{\mu}$ in $\mathbb{M}\times\mathbb{M}$ whose marginals are μ and ν.

The space of probability measures $(\mathcal{P}_2(\mathbb{M}), W_{2,\mathsf{d}_{\mathbb{M}}})$ may be considered a model class for the above mentioned applications and it is an example of complete and separable metric space, which exhibits a non-smooth, infinite dimensional pseudo-Riemannian character [37,4,49]. In particular, it is not isometric to a finite dimensional Riemannian manifold or a $\mathrm{Cat}(\kappa)$ space [20]; when $\mathbb{M}$ has nonnegative sectional curvature as in the case of the Euclidean space $\mathbb{R}^d$, then $(\mathcal{P}_2(\mathbb{M}), W_{2,\mathsf{d}_{\mathbb{M}}})$ has nonnegative curvature in the sense of Aleksandrov [49]; for a general Riemannian manifold $\mathbb{M}$, $(\mathcal{P}_2(\mathbb{M}), W_{2,\mathsf{d}_{\mathbb{M}}})$ is not an Aleksandrov space and lacks of any lower or upper curvature bound.

When $\mathbb{M}$ is compact, Sobolev spaces on $(\mathcal{P}_2(\mathbb{M}), W_2)$ have been constructed in [17] starting from measures $\mathfrak{m}$ which have full support and satisfy an integration-by-parts formula (see Section 5.2 below) on the unital algebra of cylinder maps $\mathrm{FC}_c^\infty(\mathcal{P}_2(\mathbb{M}))$, generated by linear functionals of the form

$$\mathsf{L}_\phi : \mu \mapsto \int_{\mathbb{M}} \phi\,\mathrm{d}\mu, \quad \phi \in \mathrm{C}_c^\infty(\mathbb{M}). \tag{1.17}$$

It turns out that the restriction of the pre-Cheeger energy pCE_2 (1.8) to $\mathrm{FC}_c^\infty(\mathcal{P}_2(\mathbb{M}))$ is induced by a bilinear form so that one can study the Dirichlet form arising by its closure.

Our main result is that for every separable and complete Riemannian manifold $\mathbb{M}$ and for every positive and finite Borel measure $\mathfrak{m}$ on $\mathcal{P}_2(\mathbb{M})$ (so, full support and integration-by-parts properties are not required) the algebra $\mathrm{FC}_c^\infty(\mathcal{P}_2(\mathbb{M}))$ satisfies property (1.14) and therefore it is dense in the metric Sobolev space $H^{1,2}(\mathcal{P}_2(\mathbb{M}), W_{2,\mathsf{d}_{\mathbb{M}}}, \mathfrak{m})$, which is therefore a Hilbert space. Due to the non-smooth character of W_2 such a Hilbertianity property was far from obvious even in the flat case $\mathbb{M} = \mathbb{R}^d$. We will also show that this result holds when $\mathbb{M}$ is an infinite-dimensional, separable, Hilbert space.

Our metric analysis is also supplemented with a detailed discussion of the structure of the Cheeger energy and of the minimal relaxed gradient, in the case when $\mathbb{M}$ is the

Euclidean space $\mathbb{R}^d$. Introducing the measure $\boldsymbol{\mathfrak{m}} = \int (\delta_\mu \otimes \mu)\,\mathrm{d}\mathfrak{m}(\mu)$ in $\mathcal{P}\big(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d\big)$, we will show that there is a linear continuous Wasserstein-gradient operator $\mathrm{D}_{\mathfrak{m}} : H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m}) \to L^2(\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$ representing the bilinear form associated to the Cheeger energy as

$$\mathsf{CE}_2(F, G) = \int \mathrm{D}_{\mathfrak{m}} F(\mu, x) \cdot \mathrm{D}_{\mathfrak{m}} G(\mu, x)\,\mathrm{d}\boldsymbol{\mathfrak{m}}(\mu, x), \tag{1.18}$$

and satisfying useful calculus rules which are typical of Γ-calculus for Dirichlet form. $\mathrm{D}_{\mathfrak{m}}$ also allows for an explicit characterization of the tangent bundle $L^2\big(\mathrm{T}\mathcal{P}_2(\mathbb{R}^d)\big)$ in the sense of Gigli [25,26].

We are also able to study the relaxation effect occurring in the construction of the Cheeger energy starting from (1.8). We claim that our results are sufficiently strong and provide useful tools to pave the way for further studies on the structure and the promising applications of Wasserstein Sobolev spaces. In particular, the techniques developed in the present paper can also be applied to study the general class of Wasserstein Sobolev spaces $H^{1,q}(\mathcal{P}_p(\mathbb{M}), W_p, \mathfrak{m})$, with $p, q \in (1, +\infty)$, a topic that has been addressed in [46].

Moreover, as a direct consequence of our results, the recovery of the Cheeger energy in terms of relaxation of the explicitly computable pre-Cheeger energy pCE_2 (1.8) on $\mathrm{FC}_c^\infty(\mathcal{P}_2(\mathbb{M}))$ does allow the equally explicit formulation of Euler-Lagrange equations of properly formulated variational problems defined on $H^{1,2}(\mathcal{P}_2(\mathbb{M}), W_{2,\mathsf{d}_{\mathbb{M}}}, \mathfrak{m})$, which can be solved numerically over finite dimensional suitably graduated approximations of $\mathrm{FC}_c^\infty(\mathcal{P}_2(\mathbb{M}))$, as a sort of (nonlinear) Galerkin approximation. Hence, as a concluding remark, perhaps surprisingly, the use of the subalgebra of cylindric functions $\mathrm{FC}_c^\infty(\mathcal{P}_2(\mathbb{M}))$ instead of $\mathrm{Lip}_b\big(\mathcal{P}_2(\mathbb{M})\big)$ as a fundamental nucleus to define Wasserstein Sobolev spaces allows to bring the theory from its foundational level to rather concrete applicability. In particular, we have in mind the above mentioned applications to the solutions of PDEs over $\mathcal{P}_2(\mathbb{M})$ and machine learning.

Plan of the paper After a quick review of the construction of the Cheeger energy starting from a subalgebra $\mathscr{A}$, **Section 2** is devoted to prove our main density result under condition (1.14) (Section 2.2). The last part 2.4 extends the applicability of the results to a larger class of distances: one of its quite useful applications will concern the extension of the results for the Wasserstein Sobolev spaces modeled on $\mathcal{P}_2(\mathbb{R}^d)$ to the general case of $\mathcal{P}_2(\mathbb{M})$ for a complete Riemannian manifold $\mathbb{M}$, which will be carried out in Sections 6.1 and 6.2.

We will recap a few properties of the Wasserstein distance in **Section 3**. **Section 4** contains a collection of some properties of cylinder functions, of their asymptotic Lipschitz constants (Section 4.1), and our main density and Hilbertianity result for the Wasserstein Sobolev space $H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m})$ (Theorem 4.10).

Calculus rules for the $\mathfrak{m}$-differential are presented in **Section 5**; the structure of the tangent bundle, the properties of the residual differentials, and the study of the relaxation effect are discussed in Section 5.1, together with a few examples in Section 5.2.

The last **Section 6** shows how to extend the result of Section 4 from $\mathbb{R}^d$ to an arbitrary complete Riemannian manifold $\mathbb{M}$ and to a separable Hilbert space $\mathbb{H}$.

Acknowledgments The authors warmly thank N. Gigli, E. Pasqualetto, G. Peyré for their comments and L. Dello Schiavo for the careful reading and his valuable remarks on a first draft of the present paper.

The authors gratefully acknowledge the support of the Institute for Advanced Study of the Technical University of Munich, funded by the German Excellence Initiative. G.S. has also been supported by IMATI-CNR, Pavia and by the MIUR-PRIN 2017 project *Gradient flows, Optimal Transport and Metric Measure Structures.* The authors are grateful to the anonymous reviewers for their valuable comments.

2. Metric Sobolev spaces and density of unital algebras

In this section we will briefly recap the construction of metric Sobolev spaces adapting the relaxation viewpoint of the Cheeger energy to the presence of a distinguished algebra of Lipschitz functions [7,5,44].

2.1. Sobolev functions and minimal relaxed gradients

Let (X, d) be a complete and separable metric space. We will denote by $\mathrm{Lip}_b(X, \mathsf{d})$ the space of bounded and Lipschitz real functions $f : X \to \mathbb{R}$. The asymptotic Lipschitz constant of $f \in \mathrm{Lip}_b(X, \mathsf{d})$ is defined as

$$\mathrm{lip}_{\mathsf{d}} f(x) := \lim_{r\downarrow 0} \mathrm{Lip}(f, \mathrm{B}(x, r), \mathsf{d}) = \limsup_{y,z\to x,\ y\neq z} \frac{|f(y) - f(z)|}{\mathsf{d}(y, z)}, \tag{2.1}$$

where $\mathrm{B}(x, r)$ denotes the open ball centered at x with radius r and, for $A \subset X$, the quantity $\mathrm{Lip}(f, A, \mathsf{d})$ is defined as

$$\mathrm{Lip}(f, A, \mathsf{d}) := \sup_{x,y\in A,\, x\neq y} \frac{|f(x) - f(y)|}{\mathsf{d}(x, y)}.$$

We will simply write $\mathrm{Lip}_b(X), \mathrm{lip}\, f, \mathrm{Lip}(f, A)$, omitting to explicitly mention d, when the choice of the metric d is clear from the context.

We will also deal with a unital algebra $\mathscr{A} \subset \mathrm{Lip}_b(X)$ separating the points of X, i.e.

$$1 \in \mathscr{A}, \quad \text{for every } x_0, x_1 \in X \text{ there exists } f \in \mathscr{A}\colon \quad f(x_0) \neq f(x_1). \tag{2.2}$$

The initial Hausdorff topology $\tau_{\mathscr{A}}$ induced on X by $\mathscr{A}$ is clearly coarser than the metric topology of X.

Let $\mathfrak{m}$ be a finite and positive Borel measure on X (being X a Polish space, $\mathfrak{m}$ is also a Radon measure). We will denote by $\mathcal{L}^0(X,\mathfrak{m})$ the set of $\mathfrak{m}$-measurable real functions defined in X; $L^0(X,\mathfrak{m})$ is the usual quotient of $\mathcal{L}^0(X,\mathfrak{m})$ obtained by identifying two functions which coincide $\mathfrak{m}$-a.e. in X. In a similar way, $\mathcal{L}^p(X,\mathfrak{m})$ and $L^p(X,\mathfrak{m})$ are the usual Lebesgue spaces of p-summable $\mathfrak{m}$-measurable (equivalence classes of) real functions, $p \in [1,+\infty]$. It is worth noticing that by [44, Lemma 2.1.27] we have that

$$\begin{gathered}\text{for every } p \in [1,\infty) \text{ and every } f \in \mathcal{L}^p(X,\mathfrak{m}) \text{ taking values in an interval } I \subset \mathbb{R} \\ \text{there exists a sequence } (f_n)_n \subset \mathscr{A} \text{ with values in } I \text{ converging to } f \text{ in } L^p(X,\mathfrak{m}).\end{gathered} \tag{2.3}$$

We will endow $L^0(X,\mathfrak{m})$ with the topology of the convergence in measure, which is induced by the metric

$$\mathrm{d}_{L^0}(f_1,f_2) := \int_X \vartheta(|f_1-f_2|)\,\mathrm{d}\mathfrak{m}, \quad f_1,f_2 \in L^0(X,\mathfrak{m}), \tag{2.4}$$

where $\vartheta : [0,+\infty) \to [0,+\infty)$ is any increasing, concave, bounded function with $\vartheta(0) = \lim_{r\downarrow 0}\vartheta(r) = 0$. In the following we fix an exponent $p \in (1,+\infty)$.

Definition 2.1 *($(p,\mathscr{A})$-relaxed gradient).* We say that $G \in L^p(X,\mathfrak{m})$ is a $(p,\mathscr{A})$-relaxed gradient of a $\mathfrak{m}$-measurable function $f \in L^0(X,\mathfrak{m})$ if there exists a sequence $(f_n)_{n\in\mathbb{N}} \in \mathscr{A}$ such that:

(1) $f_n \to f$ in $\mathfrak{m}$-measure and $\operatorname{lip} f_n \to \tilde{G}$ weakly in $L^p(X,\mathfrak{m})$;
(2) $\tilde{G} \le G$ $\mathfrak{m}$-a.e. in X.

The minimal $(p,\mathscr{A})$-relaxed gradient of f (denoted by $|\mathrm{D}f|_{\star,\mathscr{A}}$) is the element of minimal L^p-norm among all the $(p,\mathscr{A})$-relaxed gradient of f. We will just write $|\mathrm{D}f|_\star$ if $\mathscr{A} = \mathrm{Lip}_b(X)$.

Remark 2.2. Notice that the minimal relaxed gradient $|\mathrm{D}f|_{\star,\mathscr{A}}$ depends also on $p \in [1,+\infty)$, see e.g. [6,12,28]. Since it will be always clear from the context which value of p we are considering (a general one or, in the second part of the paper, $p = 2$), we omit to write explicitly this dependence.

We collect in the following Theorem the main properties of $|\mathrm{D}f|_{\star,\mathscr{A}}$ we will extensively use.

Theorem 2.3.

(1) *The set*

$$S := \Big\{(f, G) \in L^0(X, \mathfrak{m}) \times L^p(X, \mathfrak{m}) : G \text{ is a } (p, \mathscr{A})\text{-relaxed gradient of } f\Big\}$$

is convex and it is closed with respect to the product topology of the convergence in $\mathfrak{m}$*-measure and the weak convergence in* $L^p(X, \mathfrak{m})$*. In particular, the restriction* $S_q := S \cap L^q(X, \mathfrak{m}) \times L^p(X, \mathfrak{m})$ *is weakly closed in* $L^q(X, \mathfrak{m}) \times L^p(X, \mathfrak{m})$ *for every* $q \in (1, +\infty)$*.*

(2) *(Strong approximation) If* $f \in L^0(X, \mathfrak{m})$ *has a* $(p, \mathscr{A})$*-relaxed gradient then* $|\mathrm{D}f|_{\star,\mathscr{A}}$ *is well defined. If* f *takes values in a closed (possibly unbounded) interval* $I \subset \mathbb{R}$ *then there exists a sequence* $f_n \in \mathscr{A}$ *with values in* I *such that*

$$f_n \to f \ \mathfrak{m}\text{-a.e. in } X, \quad \operatorname{lip} f_n \to |\mathrm{D}f|_{\star,\mathscr{A}} \text{ strongly in } L^p(X, \mathfrak{m}). \tag{2.5}$$

If moreover $f \in L^q(X, \mathfrak{m})$ *for some* $q \in [1, +\infty)$ *then we can also find a sequence as in* (2.5) *converging strongly to* f *in* $L^q(X, \mathfrak{m})$*.*

(3) *(Pointwise minimality) If* G *is a* $(p, \mathscr{A})$*-relaxed gradient of* $f \in L^0(X, \mathfrak{m})$ *then* $|\mathrm{D}f|_{\star,\mathscr{A}} \le G$ $\mathfrak{m}$*-a.e. in* X*.*

(4) *(Leibniz rule) If* $f, g \in L^\infty(X, \mathfrak{m})$ *have* $(p, \mathscr{A})$*-relaxed gradient, then* $h := fg$ *has* $(p, \mathscr{A})$*-relaxed gradient and*

$$|\mathrm{D}(fg)|_{\star,\mathscr{A}} \le |f|\, |\mathrm{D}g|_{\star,\mathscr{A}} + |g|\, |\mathrm{D}f|_{\star,\mathscr{A}} \quad \mathfrak{m}\text{-a.e. in } X. \tag{2.6}$$

(5) *(Sub-linearity) If* $f, g \in L^0(X, \mathfrak{m})$ *have* $(p, \mathscr{A})$*-relaxed gradient and* $\alpha, \beta \in \mathbb{R}$*, then*

$$|\mathrm{D}(\alpha f + \beta g)|_{\star,\mathscr{A}} \le |\alpha|\, |\mathrm{D}f|_{\star,\mathscr{A}} + |\beta|\, |\mathrm{D}g|_{\star,\mathscr{A}} \quad \mathfrak{m}\text{-a.e. in } X. \tag{2.7}$$

(6) *(Locality) If* $f \in L^0(X, \mathfrak{m})$ *has a* $(p, \mathscr{A})$*-relaxed gradient, then for any* $\mathscr{L}^1$*-negligible Borel subset* $N \subset \mathbb{R}$ *we have*

$$|\mathrm{D}f|_{\star,\mathscr{A}} = 0 \quad \mathfrak{m}\text{-a.e. on } f^{-1}(N). \tag{2.8}$$

(7) *(Chain rule) If* $f \in L^0(X, \mathfrak{m})$ *has a* $(p, \mathscr{A})$*-relaxed gradient and* $\phi \in \mathrm{Lip}(\mathbb{R})$ *then* $\phi \circ f$ *has* $(p, \mathscr{A})$*-relaxed gradient and*

$$|\mathrm{D}(\phi \circ f)|_{\star,\mathscr{A}} \le |\phi'(f)|\, |\mathrm{D}f|_{\star,\mathscr{A}} \quad \mathfrak{m}\text{-a.e. in } X, \tag{2.9}$$

and equality holds in (2.9) *if* ϕ *is monotone or* C^1*.*

(8) *(Truncations) If* $f_j \in L^0(X, \mathfrak{m})$ *has* $(p, \mathscr{A})$*-relaxed gradient,* $j = 1, \cdots, J$*, then also the functions* $f_+ := \max(f_1, \cdots, f_J)$ *and* $f_- := \min(f_1, \cdots, f_J)$ *have* $(p, \mathscr{A})$*-relaxed gradient and*

$$|\mathrm{D}f_+|_{\star,\mathscr{A}} = |\mathrm{D}f_j|_{\star,\mathscr{A}} \quad \mathfrak{m}\text{-a.e. on } \{x \in X : f_+ = f_j\}, \tag{2.10}$$

$$|\mathrm{D}f_-|_{\star,\mathscr{A}} = |\mathrm{D}f_j|_{\star,\mathscr{A}} \quad \mathfrak{m}\text{-a.e. on } \{x \in X : f_- = f_j\}. \tag{2.11}$$

Remark 2.4. Notice that the product in (2.9) is well defined since there exists a $\mathscr{L}^1$-negligible Borel set $N \subset \mathbb{R}$ such that ϕ is differentiable in $\mathbb{R} \setminus N$ and $|\mathrm{D}f|_{\star,\mathscr{A}}$ vanishes $\mathfrak{m}$-a.e. in $f^{-1}(N)$ thanks to the locality property (2.8).

Proof. We give a few references for the proofs. The case when $p = 2$, $\mathscr{A} = \mathrm{Lip}_b(X)$ and the local slope of f is used to define relaxed gradients have been considered in [7, Sec. 4], whose proof generalizes easily to the case $p \in (1, \infty)$ and the asymptotic Lipschitz constant (2.1), see also [5].

The definition and the properties involving a general unital subalgebra $\mathscr{A}$ have been discussed in [44, Sec. 3]: points (1,2) correspond to Lemma 3.1.6 and Corollary 3.1.9, (3) has been stated in Lemma 3.1.11, (4) refers to Corollary 3.1.10, (5,6,7,8) are proved in Theorem 3.1.12 and its Corollary 3.1.13.

Let us make three further technical comments:

- both [7,44] involve an auxiliary topology τ: in the present case, being X complete and separable and d a canonical metric (thus d only take finite values), we can select τ as the (Polish) topology induced by d.
- In order to deal with *extended* distances, in [44] has also been assumed that the unital algebra $\mathscr{A}$ satisfies the *stronger* compatibility condition

$$\mathsf{d}(x, y) = \sup \Big\{ f(x) - f(y) : f \in \mathscr{A},\ \mathrm{Lip}(f, X) \le 1 \Big\}, \tag{2.12}$$

  which clearly implies that $\mathscr{A}$ separates the points of X as in (2.2). However, such a property is not needed in the construction and the proofs of Section 3.1.1 of [44]. The only point where (2.12) explicitly occurs is in the proof of Locality [44, Lemma 3.1.11], to ensure that the restriction of $\mathscr{A}$ to each compact set $K \subset X$ is uniformly dense in $\mathrm{C}(K)$, a property which is guaranteed in the present setting by (2.2) thanks to Stone-Weierstrass Theorem.
- The standard approach of [7,44] considers first functions f belonging to $L^p(X, \mathfrak{m})$ instead of general $\mathfrak{m}$-measurable functions. However, the compatibility with truncations showing that for every $k > 0$

$$|\mathrm{D}\, T_k(f)|_{\star,\mathscr{A}}(x) = \begin{cases} |\mathrm{D}f|_{\star,\mathscr{A}}(x) & \text{if } |f(x)| < k, \\ 0 & \text{if } |f(x)| \ge k, \end{cases} \qquad T_k(f) := -k \vee f \wedge k, \tag{2.13}$$

  and the possibility to find strong approximations of $T_k(f) \in L^p(X, \mathfrak{m})$ (recall that $\mathfrak{m}$ is finite) satisfying (2.5) and taking values in $[-k, k]$ (see [44, Cor 2.1.24, Cor. 3.1.9] where an approximation argument involving odd polynomials is implemented) allow for a standard extension of the theory from $L^p(X, \mathfrak{m})$ to $L^0(X, \mathfrak{m})$, see also the discussion related to (4.16) of [7]. Notice also that, from a metric point of view, there is no reason to couple the integrability of a function f and the one of its minimal relaxed gradient $|\mathrm{D}f|_{\star,\mathscr{A}}$. Also the choice of working in $L^0(X, \mathfrak{m})$ gives more flexibility and in particular allows to treat the distance function d_y from one point $y \in X$ without imposing any integrability condition. This will be crucial in the rest of this paper (see e.g. Theorem 2.13). $\square$

Starting from Definition 2.1 and using the properties of Theorem 2.3 it is natural to introduce the following notions.

Definition 2.5 *(Cheeger energy and Sobolev space).* We call $D^{1,p}(X, \mathsf{d}, \mathfrak{m}; \mathscr{A})$ the set of functions in $L^0(X, \mathfrak{m})$ with a $(p, \mathscr{A})$-relaxed gradient and we set

$$\mathsf{CE}_{p,\mathscr{A}}(f) := \int_X |\mathrm{D}f|^p_{\star,\mathscr{A}}(x)\,\mathrm{d}\mathfrak{m}(x) \quad \text{for every } f \in D^{1,p}(X, \mathsf{d}, \mathfrak{m}; \mathscr{A}), \tag{2.14}$$

with $\mathsf{CE}_{p,\mathscr{A}}(f) := +\infty$ if $f \notin D^{1,p}(X, \mathsf{d}, \mathfrak{m}; \mathscr{A})$. The Sobolev space $H^{1,p}(X, \mathsf{d}, \mathfrak{m}; \mathscr{A})$ is defined as $L^p(X, \mathfrak{m}) \cap D^{1,p}(X, \mathsf{d}, \mathfrak{m}; \mathscr{A})$ and it is a Banach space with the norm $\|f\|^p_{H^{1,p}(X,d,\mathfrak{m};\mathscr{A})} := \|f\|^p_{L^p} + \mathsf{CE}_{p,\mathscr{A}}(f)$. As usual, we will write $D^{1,p}(X, \mathsf{d}, \mathfrak{m}), \mathsf{CE}_p(f)$, $H^{1,p}(X, \mathsf{d}, \mathfrak{m})$ and $\|f\|_{H^{1,p}}$ when $\mathscr{A} = \mathrm{Lip}_b(X)$.

Remark 2.6 *(Cheeger energy as relaxation of the pre-Cheeger energy).* We can equivalently define the Cheeger energy $\mathsf{CE}_{p,\mathscr{A}}$ as a sort of L^0-lower semicontinuous relaxation of the restriction to $\mathscr{A}$ of the *pre-Cheeger energy* pCE_p, the latter being defined as

$$\mathsf{pCE}_p(f) := \int_X (\mathrm{lip}\, f)^p\,\mathrm{d}\mathfrak{m}, \quad f \in \mathrm{Lip}_b(X). \tag{2.15}$$

In other words, for every $f \in L^0(X, \mathfrak{m})$ it holds ([44, Corollary 3.1.7])

$$\mathsf{CE}_{p,\mathscr{A}}(f) = \inf\left\{\liminf_{n\to+\infty} \mathsf{pCE}_p(f_n) : f_n \in \mathscr{A},\, f_n \to f \text{ in } L^0(X, \mathfrak{m})\right\}. \tag{2.16}$$

In particular the functional $\mathsf{CE}_{p,\mathscr{A}}$ is lower semicontinuous in $L^0(X, \mathfrak{m})$. Here the choice of the L^0-topology does not play a crucial role, since, by Theorem 2.3(2), the restriction of $\mathsf{CE}_{p,\mathscr{A}}$ to $L^q(X, \mathfrak{m})$, $q \in [1, \infty)$, can be equivalently obtained as L^q-relaxation:

$$\mathsf{CE}_{p,\mathscr{A}}(f) = \inf\left\{\liminf_{n\to+\infty} \mathsf{pCE}_p(f_n) : f_n \in \mathscr{A},\, f_n \to f \text{ in } L^q(X, \mathfrak{m})\right\}, \quad f \in L^q(X, \mathfrak{m}). \tag{2.17}$$

Notice also that, when $\mathfrak{m}$ has not full support, two different elements $f_1, f_2 \in \mathscr{A}$ may give rise to the same equivalence class in $L^0(X, \mathfrak{m})$. In this case, $\mathsf{CE}_{p,\mathscr{A}}$ can be equivalently defined as the L^0-lower semicontinuous relaxation of the functional

$$\widetilde{\mathsf{pCE}}_p(f) := \inf\left\{\mathsf{pCE}_p(g) : g \in \mathscr{A},\, g = f \ \mathfrak{m}\text{-a.e.}\right\}, \quad f \in \mathscr{A}_{\mathfrak{m}},$$

where $\mathscr{A}_{\mathfrak{m}}$ is the quotient of $\mathscr{A}$ with respect to equality $\mathfrak{m}$-a.e.

It is clear that we have the obvious implication for $f \in L^0(X, \mathfrak{m})$:

$$f \text{ has a } (p, \mathscr{A})\text{-relaxed gradient} \quad \Rightarrow \quad \begin{cases} f \text{ has a } (p, \mathrm{Lip}_b(X))\text{-relaxed gradient and} \\ |\mathrm{D}f|_\star \le |\mathrm{D}f|_{\star,\mathscr{A}} \ \mathfrak{m}\text{-a.e. in } X. \end{cases} \tag{2.18}$$

The converse implication together with the identity $|\mathrm{D}f|_\star = |\mathrm{D}f|_{\star,\mathscr{A}}$ is an important density property for an algebra $\mathscr{A}$: by Theorem 2.3(2), it is equivalent to the following property.

Definition 2.7 *(Density in energy of a subalgebra of Lipschitz functions).* We say that a subalgebra $\mathscr{A} \subset \mathrm{Lip}_b(X)$ is *dense in p-energy* if for every $f \in L^0(X, \mathfrak{m})$ with a p-relaxed gradient there exists a sequence $(f_n)_{n\in\mathbb{N}}$ satisfying

$$f_n \in \mathscr{A}, \quad f_n \to f \ \mathfrak{m}\text{-a.e. in } X, \quad \mathrm{lip}\, f_n \to |\mathrm{D}f|_\star \text{ strongly in } L^p(X, \mathfrak{m}). \tag{2.19}$$

When $\mathscr{A}$ is unital and separating, this is equivalent to the fact that f has a $(p, \mathscr{A})$-relaxed gradient and

$$|\mathrm{D}f|_{\star,\mathscr{A}} = |\mathrm{D}f|_\star \quad \mathfrak{m}\text{-a.e. in } X. \tag{2.20}$$

In particular $D^{1,p}(X, \mathsf{d}, \mathfrak{m}; \mathscr{A}) = D^{1,p}(X, \mathsf{d}, \mathfrak{m})$.

Remark 2.8 *(Comparison with the Newtonian approach).* By the identification (1.12) when $\mathscr{A} = \mathrm{Lip}_b(X)$ we always have

$$H^{1,p}(X, \mathsf{d}, \mathfrak{m}) = \hat{N}^{1,p}(X, \mathsf{d}, \mathfrak{m}), \quad |\mathrm{D}f|_N = |\mathrm{D}f|_\star \quad \text{for every } f \in \hat{N}^{1,p}(X, \mathsf{d}, \mathfrak{m}). \tag{2.21}$$

If $\mathscr{A}$ is dense in p-energy and $f \in \hat{N}^{1,p}(X, \mathsf{d}, \mathfrak{m})$ we thus obtain

$$|\mathrm{D}f|_{\star,\mathscr{A}} = |\mathrm{D}f|_\star = |\mathrm{D}f|_N \quad \mathfrak{m}\text{-a.e. in } X. \tag{2.22}$$

Notice that (2.22) and (1.5) immediately yield the uniform upper bound in terms of the pointwise Lipschitz constant

$$\text{if } f \text{ is Lipschitz then} \quad |\mathrm{D}f|_{\star,\mathscr{A}} \le |\mathrm{D}f| \quad \mathfrak{m}\text{-a.e. in } X. \tag{2.23}$$

Remark 2.9. As we already mentioned in Remark 2.6, the choice of arbitrary measurable maps $f \in L^0(X, \mathfrak{m})$ in Definition 2.7 and of the pointwise $\mathfrak{m}$-a.e. convergence in (2.19) is not restrictive: a simple truncation argument (which can be implemented by using odd polynomials, see [44, Corollary 2.1.24]) shows that $\mathscr{A}$ is dense in p-energy if and only if for every $f \in L^p(X, \mathfrak{m})$ with a p-relaxed gradient there exists a sequence $(f_n)_{n\in\mathbb{N}}$ satisfying

$$f_n \in \mathscr{A}, \quad f_n \to f \text{ in } L^p(X,\mathfrak{m}), \quad \operatorname{lip} f_n \to |\mathrm{D}f|_\star \text{ strongly in } L^p(X,\mathfrak{m}). \tag{2.24}$$

If $\mathscr{A}$ is unital and separating this is equivalent to $H^{1,p}(X,\mathsf{d},\mathfrak{m};\mathscr{A}) = H^{1,p}(X,\mathsf{d},\mathfrak{m})$ with equal norms.

A first sufficient condition to obtain the density in energy of a subalgebra $\mathscr{A}$, in the more general framework of extended topological metric measure spaces, is provided by the compatibility condition (2.12) [44, Theorems 3.2.7, 5.3.1] (see also [9] for the algebra generated by truncated distance functions).

In the present Polish setting, we notice that (2.19) (and, a fortiori, (2.12)) implies the weaker condition

$$\text{for every } y \in X \text{ the function } \mathsf{d}_y : x \mapsto \mathsf{d}(x,y) \text{ has } (p,\mathscr{A})\text{-relaxed gradient } 1, \tag{2.25}$$

which is equivalent, thanks to Theorem 2.3(3), to

$$|\mathrm{D}\mathsf{d}_y|_{\star,\mathscr{A}} \le 1 \quad \mathfrak{m}\text{-a.e. in } X. \tag{2.26}$$

In fact, using the truncations (2.13), each function d_y can be approximated by the increasing sequence $f_k := T_k\mathsf{d}_y$ of bounded 1-Lipschitz maps, so that

$$|\mathrm{D}\mathsf{d}_y|_\star \le 1 \quad \mathfrak{m}\text{-a.e. in } X \text{ for every } y \in X, \tag{2.27}$$

and therefore (2.19) yields (2.26).

Remark 2.10 *(The effect of truncations).* The $(p,\mathscr{A})$-relaxed gradient is not affected by truncations of the distance functions, in particular it is not restrictive to assume d bounded above by a constant, e.g. 1. In fact, if we introduce a parameter $a > 0$ and the truncated distance

$$\mathsf{d}_a(x_1,x_2) := \mathsf{d}(x_1,x_2) \wedge a \quad \text{for every } x_1,x_2 \in X, \tag{2.28}$$

(X,d_a) is still a complete and separable metric space, the sets $\mathrm{Lip}_b(X,\mathsf{d})$ and $\mathrm{Lip}_b(X,\mathsf{d}_a)$ coincide, and it is easy to check that

$$\operatorname{lip}_{\mathsf{d}} f = \operatorname{lip}_{\mathsf{d}_a} f \quad \text{for every bounded and Lipschitz function } f. \tag{2.29}$$

We deduce that d and d_a induce the same $(p,\mathscr{A})$-relaxed gradient. Notice moreover that using (2.28) we can also easily cover the case of extended distances (i.e. possibly assuming the value $+\infty$), *provided* (X,d_a) *is a separable metric space.* The case when (X,d_a) is not separable requires a more refined setting involving an auxiliary topology τ [44].

It is possible to express (2.26) in a more flexible way, by using suitable nonlinear functions of d_y. We state a general result.

Lemma 2.11. *Let $I = (a, b)$ be an interval (possibly unbounded) of $\mathbb{R}$ and let $\zeta : \mathbb{R} \to \mathbb{R}$ be a Lipschitz and nondecreasing map satisfying*

$$\text{the restriction of } \zeta \text{ to } I \text{ is of class } \mathrm{C}^1 \text{ with } \zeta'(s) > 0 \text{ if } s \in I. \tag{2.30}$$

If $f : X \to \overline{I}$ is a Borel function, then the condition

$$f \in D^{1,p}(X, \mathsf{d}, \mathfrak{m}; \mathscr{A}), \quad |\mathrm{D}f|_{\star,\mathscr{A}} \le 1 \tag{2.31}$$

is equivalent to

$$\zeta \circ f \in D^{1,p}(X, \mathsf{d}, \mathfrak{m}; \mathscr{A}), \quad \big|\mathrm{D}(\zeta \circ f)\big|_{\star,\mathscr{A}}(x) \le \zeta'(f(x)) \quad \text{for } \mathfrak{m}\text{-a.e. } x \in X. \tag{2.32}$$

Proof. It is clear that if $|\mathrm{D}f|_{\star,\mathscr{A}} \le 1$ then (2.32) holds, thanks to (2.9). In order to prove the converse implication, we consider a strictly decreasing sequence $a_n \downarrow a$, a strictly increasing sequence $b_n \uparrow b$ and nondecreasing and bounded Lipschitz functions $\psi_n : \mathbb{R} \to \mathbb{R}$ such that

$$\psi_n(z) = a_n \text{ if } z < \zeta(a_n), \quad \psi_n(\zeta(s)) = s \text{ for every } s \in [a_n, b_n], \quad \psi_n(z) = b_n \text{ if } z > \zeta(b_n).$$

The restriction of ψ_n to the interval $[\zeta(a_n), \zeta(b_n)]$ is of class C^1.

Setting $h(x) := \zeta(f(x))$, the Chain rule (2.9) yields

$$|\mathrm{D}(\psi_n \circ h)|_{\star,\mathscr{A}}(x) \le (\psi_n' \circ h)\, |\mathrm{D}h|_{\star,\mathscr{A}}(x) \le (\psi_n' \circ \zeta(f(x)))\zeta'(f(x)).$$

Since $\psi_n(h(x)) = a_n \vee f(x) \wedge b_n$, the locality property (2.8), the truncation Property 2.3(8), and the fact that ${\psi_n}'(\zeta(s))\zeta'(s) = 1$ if $s \in [a_n, b_n]$ yield

$$|\mathrm{D}(\psi_n \circ h)|_{\star,\mathscr{A}} \le 1 \quad \mathfrak{m}\text{-a.e.} \tag{2.33}$$

Since $\psi_n \circ h \to f$ pointwise in X as $n \to \infty$, passing to the limit in (2.33) we get $|\mathrm{D}f|_{\star,\mathscr{A}} \le 1$. $\square$

Remark 2.12. Thanks to Lemma 2.11, if d is a bounded metric and $q > 1$, (2.26) is equivalent to

$$|\mathrm{D}\mathsf{d}_y^q|_{\star,\mathscr{A}}(x) \le q\, \mathsf{d}_y^{q-1}(x) \quad \text{for } \mathfrak{m}\text{-a.e. } x \in X. \tag{2.34}$$

In particular, if (2.34) holds for some $q \ge 1$, it holds for any $q \ge 1$.

2.2. A density result

We have seen that in the present setting of Polish spaces, condition (2.26) (or, equivalently, (2.32) for some admissible truncation satisfying (2.30)) is a necessary condition for the validity of the approximation property (2.19) and of the identification $|\mathrm{D}f|_\star = |\mathrm{D}f|_{\star,\mathscr{A}}$. We want to show that (2.26) or (2.32) are also *sufficient* conditions.

Theorem 2.13. *Let $(X, \mathsf{d}, \mathfrak{m})$ be a Polish metric measure space, let $Y \subset X$ be a dense subset, and let $\mathscr{A}$ be a unital separating subalgebra of $\mathrm{Lip}_b(X)$ as in (2.2). If*

$$\text{for every } y \in Y \text{ it holds} \quad \mathsf{d}_y \in D^{1,p}(X, \mathsf{d}, \mathfrak{m}; \mathscr{A}), \quad \left|\mathrm{D}\mathsf{d}_y\right|_{\star,\mathscr{A}} \le 1 \tag{2.35}$$

then $\mathscr{A}$ is dense in p-energy according to Definition 2.7.

Proof. We split the proof in various steps. Notice that by (2.18) it is sufficient to prove that

$$|\mathrm{D}f|_{\star,\mathscr{A}} \le |\mathrm{D}f|_\star \quad \mathfrak{m}\text{-a.e. in } X. \tag{2.36}$$

(1) *It is not restrictive to assume* d *bounded above by* 1 : see Remark 2.10.

By Lemma 2.11 and Remark 2.12 we know that (2.34) holds for every $y \in Y$ and every $q \ge 1$.

(2) *It is sufficient to prove that*

$$\mathsf{CE}_{p,\mathscr{A}}(f) \le \int_X (\mathrm{lip}\, f)^p \,\mathrm{d}\mathfrak{m} = \mathsf{pCE}_p(f) \quad \textit{for every } f \in \mathrm{Lip}_b(X). \tag{2.37}$$

In fact, if f has $(p, \mathrm{Lip}_b(X))$-relaxed gradient, by (2.5) we can find a sequence $f_n \in \mathrm{Lip}_b(X)$ such that $f_n \to f$ $\mathfrak{m}$-a.e. and $\mathrm{lip}\, f_n \to |\mathrm{D}f|_\star$ strongly in $L^p(X, \mathfrak{m})$ as $n \to \infty$. By the L^0-lower semicontinuity of the $\mathsf{CE}_{p,\mathscr{A}}$-energy, passing to the limit in (2.37) written for f_n we get

$$\mathsf{CE}_{p,\mathscr{A}}(f) = \int_X |\mathrm{D}f|^p_{\star,\mathscr{A}} \,\mathrm{d}\mathfrak{m} \le \int_X |\mathrm{D}f|^p_\star \,\mathrm{d}\mathfrak{m} = \mathsf{CE}_p(f) < \infty.$$

We deduce that f has a $(p, \mathscr{A})$-relaxed gradient and that (2.20) holds, since $|\mathrm{D}f|_\star \le |\mathrm{D}f|_{\star,\mathscr{A}}$ $\mathfrak{m}$-a.e.

(3) For every $f \in \mathrm{Lip}_b(X)$ and $t > 0$ we introduce the Hopf-Lax regularization $\mathsf{Q}_t f : X \to \mathbb{R}$ defined by

$$\mathsf{Q}_t f(x) := \inf_{y \in X} \frac{1}{q t^{q-1}} \mathsf{d}^q(x, y) + f(y), \quad x \in X, \tag{2.38}$$

where $q \in (1, +\infty)$ is the conjugate exponent of p i.e. $1/q + 1/p = 1$. It is clear that $\mathsf{Q}_t f$ is bounded (it takes values in the interval $[\inf_X f, \sup_X f]$) and Lipschitz, being the infimum of a family of uniformly Lipschitz functions. We consider the upper semicontinuous function [7, (3.4) and Prop. 3.2]

$$\mathsf{D}_t^+ f(x) := \sup_{(y_n)} \limsup_{n\to\infty} \mathsf{d}(x, y_n), \tag{2.39}$$

where the $(y_n)_n$'s vary among all the minimizing sequences of (2.38). $\mathsf{D}_t^+ f$ is also uniformly bounded and satisfies (see e.g. [44, Lemma 3.2.1])

$$\left(\frac{\mathsf{D}_t^+ f(x)}{t}\right)^q \le (q \operatorname{Lip}(f, X))^p \,. \tag{2.40}$$

In fact, if y_n is a minimizing sequence of (2.38), for every $\varepsilon > 0$ we eventually have

$$\frac{1}{qt^{q-1}} \mathsf{d}^q(x, y_n) + f(y_n) \le \mathsf{Q}_t f(x) + \varepsilon \le f(x) + \varepsilon$$

i.e., setting $L := \operatorname{Lip}(f, X)$,

$$\frac{1}{t^q} \mathsf{d}^q(x, y_n) \le \frac{\varepsilon q}{t} + \frac{q}{t}(f(x) - f(y_n)) \le \frac{\varepsilon q}{t} + qL\frac{\mathsf{d}(x, y_n)}{t} \le \frac{\varepsilon q}{t} + (qL)^p + \frac{\mathsf{d}^q(x, y_n)}{qt^q p^{1/(p-1)}}.$$

We thus get

$$\limsup_{n\to\infty} \frac{1}{t^q} \mathsf{d}^q(x, y_n) \le \frac{\varepsilon q}{t} + (qL)^p$$

which yields (2.40) since $\varepsilon > 0$ is arbitrary.

(4) *For every $f \in \operatorname{Lip}_b(X)$ and for every $t > 0$*

$$|\mathrm{D}\mathsf{Q}_t f|_{\star,\mathscr{A}}(x) \le \left(t^{-1} \mathsf{D}_t^+ f(x)\right)^{q-1} \quad \text{for } \mathfrak{m}\text{-a.e. } x \in X. \tag{2.41}$$

Let $Y' = \{y_n\}_{n\in\mathbb{N}}$ be a countable set dense in Y; since $f \in \operatorname{Lip}_b(X)$ it is easy to check that

$$\begin{aligned} \mathsf{Q}_t f(x) &= \inf_{y\in Y} \frac{1}{qt^{q-1}} \mathsf{d}^q(x, y) + f(y) = \lim_{n\to\infty} \mathsf{Q}_t^n f(x), \\ \mathsf{Q}_t^n f(x) &:= \min_{1\le k\le n} \frac{1}{qt^{q-1}} \mathsf{d}^q(x, y_k) + f(y_k). \end{aligned} \tag{2.42}$$

We consider now the upper semicontinuous function

$$\mathsf{D}_t^n(x) := \max\Big\{\mathsf{d}(x, y_k) : 1 \le k \le n,\ \mathsf{Q}_t^n(x) = \frac{1}{qt^{q-1}} \mathsf{d}^q(x, y_k) + f(y_k)\Big\}. \tag{2.43}$$

By (2.34) and Theorem 2.3(8), we have that $(t^{-1}\mathsf{D}_t^n)^{q-1}$ is a $(p,\mathscr{A})$-relaxed gradient of $\mathsf{Q}_t^n f$. It is then clear that for every x there exists a sequence $n \mapsto y'(n,x)$ with $y'(n,x) \in \{y_1, \cdots, y_n\}$ such that $\mathsf{D}_t^n(x) = \mathsf{d}(x, y'(n,x))$ and $\mathsf{Q}_t^n f(x) = \frac{1}{qt^{q-1}}\mathsf{d}^q(x, y'(n,x)) + f(y'(n,x)) \to \mathsf{Q}_t f(x)$ as $n \to \infty$, i.e. $y'(n,x)$ is a minimizing sequence of (2.38). We deduce that

$$\limsup_{n\to\infty} \mathsf{D}_t^n(x) = \limsup_{n\to\infty} \mathsf{d}(x, y'(n,x)) \le \mathsf{D}_t^+ f(x) \text{ for every } x \in X. \tag{2.44}$$

Since $\mathsf{D}_t^n f$ are uniformly bounded, up to extracting a suitable subsequence we can suppose that $(t^{-1}\mathsf{D}_t^n)^{q-1} \rightharpoonup^* G$ weakly* $L^\infty(X,\mathfrak{m})$ so that, by Theorem 2.3(1), G is a $(p,\mathscr{A})$-relaxed gradient of $\mathsf{Q}_t f$, hence $|\mathrm{D}\mathsf{Q}_t f|_{\star,\mathscr{A}} \le G$ $\mathfrak{m}$-a.e. by Theorem 2.3(3). Also notice that by Fatou's lemma and weak* $L^\infty(X,\mathfrak{m})$ convergence, we have

$$\begin{aligned}\int_B G \,\mathrm{d}\mathfrak{m} &= \lim_{n\to+\infty} \int_B (t^{-1}\mathsf{D}_t^n)^{q-1} \,\mathrm{d}\mathfrak{m} \le \int_B \limsup_{n\to+\infty} (t^{-1}\mathsf{D}_t^n(x))^{q-1} \,\mathrm{d}\mathfrak{m}(x) \\ &\le \int_B (t^{-1}\mathrm{D}_t^+ f(x))^{q-1} \,\mathrm{d}\mathfrak{m}(x),\end{aligned}$$

for every Borel set $B \subset X$. We conclude that $|\mathrm{D}\mathsf{Q}_t f|_{\star,\mathscr{A}} \le (t^{-1}\mathrm{D}_t^+ f(x))^{q-1}$ for $\mathfrak{m}$-a.e. $x \in X$.

(5) *For every* $x \in X,\ t > 0,$ *and* $f \in \mathrm{Lip}_b(X)$ *we have*

$$\frac{f(x) - \mathsf{Q}_t f(x)}{t} = \frac{1}{p}\int_0^1 \Big(\frac{\mathrm{D}_{rt}^+ f(x)}{rt}\Big)^q \,\mathrm{d}r, \tag{2.45}$$

$$\limsup_{t\downarrow 0} \frac{f(x) - \mathsf{Q}_t f(x)}{t} \le \frac{1}{p}\big(\operatorname{lip} f(x)\big)^p. \tag{2.46}$$

This follows by [44, Thm. 3.2.4] (see also [4, Thm. 3.1.4, Lemma 3.1.5]).

(6) *Conclusion.* We argue as in [44, Theorem 3.2.7]: (2.45) and (2.40) yield the uniform bound

$$\frac{f(x) - \mathsf{Q}_t f(x)}{t} \le \frac{1}{p}\big(q \,\mathrm{Lip}(f,X)\big)^p \quad \text{for every } x \in X,\ t > 0. \tag{2.47}$$

Integrating (2.46) in X and applying Fatou's Lemma we get

$$\limsup_{t\downarrow 0} \int_X \frac{f(x) - \mathsf{Q}_t f(x)}{t} \,\mathrm{d}\mathfrak{m}(x) \le \frac{1}{p}\int_X \big(\operatorname{lip} f(x)\big)^p \,\mathrm{d}\mathfrak{m}(x). \tag{2.48}$$

On the other hand, (2.45) and Fubini's Theorem yield

$$\int_X \frac{f(x) - \mathsf{Q}_t f(x)}{t} \, \mathrm{d}\mathfrak{m}(x) = \frac{1}{p} \int_0^1 \int_X \Big(\frac{\mathsf{D}^+_{rt} f(x)}{rt} \Big)^q \, \mathrm{d}\mathfrak{m}(x) \, \mathrm{d}r. \tag{2.49}$$

A further application of Fatou's Lemma yields

$$\liminf_{t \downarrow 0} \int_X \frac{f(x) - \mathsf{Q}_t f(x)}{t} \, \mathrm{d}\mathfrak{m}(x) \geq \frac{1}{p} \liminf_{t \downarrow 0} \int_X \Big(\frac{\mathsf{D}^+_t f(x)}{t} \Big)^q \, \mathrm{d}\mathfrak{m}(x). \tag{2.50}$$

Using the fact that $t^{-1}\mathsf{D}^+_t f$ is uniformly bounded by (2.40), we can find a decreasing and vanishing sequence $n \mapsto t(n)$ and a limit function $G \in L^\infty(X, \mathfrak{m})$ such that

$$\begin{gathered} \Big(t(n)^{-1} \mathsf{D}^+_{t(n)} f \Big)^{q-1} \rightharpoonup^* G \quad \text{weakly}^* \text{ in } L^\infty(X, \mathfrak{m}) \quad \text{as } n \to \infty, \\ \lim_{n \to \infty} \int_X \Big(\frac{\mathsf{D}^+_{t(n)} f(x)}{t(n)} \Big)^q \, \mathrm{d}\mathfrak{m}(x) = \liminf_{t \downarrow 0} \int_X \Big(\frac{\mathsf{D}^+_t f(x)}{t} \Big)^q \, \mathrm{d}\mathfrak{m}(x). \end{gathered} \tag{2.51}$$

Since $\big(t^{-1}\mathsf{D}^+_t f\big)^{q-1}$ is a $(p, \mathscr{A})$-relaxed gradient of $\mathsf{Q}_t f$ by claim (4) and $\mathsf{Q}_t f \to f$ pointwise everywhere, using Theorem 2.3(1) we get that G is a $(p, \mathscr{A})$-relaxed gradient of f. Using the lower semicontinuity of the L^p-norm w.r.t. the weak* $L^\infty(X, \mathfrak{m})$ convergence, we get that

$$\lim_{n \to \infty} \int_X \left(\frac{\mathsf{D}^+_{t(n)} f(x)}{t(n)} \right)^q \mathrm{d}\mathfrak{m}(x) = \lim_{n \to \infty} \int_X \left(\frac{\mathsf{D}^+_{t(n)} f(x)}{t(n)} \right)^{p(q-1)} \mathrm{d}\mathfrak{m}(x) \tag{2.52}$$

$$\geq \int_X G^p \, \mathrm{d}\mathfrak{m}(x) \tag{2.53}$$

$$\geq \int_X |\mathrm{D} f|^p_{\star, \mathscr{A}}(x) \, \mathrm{d}\mathfrak{m}(x), \tag{2.54}$$

where we also used the pointwise minimality of $|\mathrm{D} f|_{\star, \mathscr{A}}$ given by Theorem 2.3(3). Combining (2.52), (2.51), (2.50) and (2.48) we deduce that

$$\int_X |\mathrm{D} f|^p_{\star, \mathscr{A}}(x) \, \mathrm{d}\mathfrak{m}(x) \leq \int_X \big(\mathrm{lip}\, f(x) \big)^p \, \mathrm{d}\mathfrak{m}(x)$$

so that (2.37) holds. □

Corollary 2.14 *(Density in energy of $\mathscr{A}$ in $H^{1,p}$). If $\mathscr{A}$ is a separating unital subalgebra of* $\mathrm{Lip}_b(X)$ *satisfying* (2.35) *then*

$$\mathsf{CE}_{p, \mathscr{A}}(f) = \mathsf{CE}_p(f) = \mathsf{NE}_p(f) \quad \textit{for every } \mathfrak{m}\textit{-measurable function } f : X \to \mathbb{R}. \tag{2.55}$$

In particular, $H^{1,p}(X, \mathsf{d}, \mathfrak{m}) = H^{1,p}(X, \mathsf{d}, \mathfrak{m}; \mathscr{A})$.

As we have already said, (2.55) can be interpreted as a density result in $H^{1,p}(X, \mathsf{d}, \mathfrak{m})$: for every $f \in H^{1,p}(X, \mathsf{d}, \mathfrak{m})$ there exists a sequence $f_n \in \mathscr{A}$, $n \in \mathbb{N}$, such that

$$f_n \to f,\ \operatorname{lip} f_n \to |\mathrm{D}f|_* \quad \text{strongly in } L^p(X, \mathfrak{m}), \quad \int_X |\operatorname{lip} f_n|^p \,\mathrm{d}\mathfrak{m} \to \mathsf{CE}_p(f) \quad \text{as } n \to \infty. \tag{2.56}$$

2.3. Applications

We first recall a useful result showing that it is possible to remove the assumption that $\mathscr{A}$ is unital, if $\mathscr{A}$ satisfies a suitable tightness condition. We will denote by $\mathbb{1}$ the unit constant function.

Proposition 2.15. *Let $\mathscr{A} \subset \mathrm{Lip}_b(X)$ be a separating subalgebra of Lipschitz functions and let*

$$\mathscr{A}_1 := \mathscr{A} \oplus \{c\mathbb{1}\} = \Big\{ f + c\mathbb{1} : f \in \mathscr{A},\, c \in \mathbb{R} \Big\} \tag{2.57}$$

be the minimal unital subalgebra containing $\mathscr{A}$. If $\mathscr{A}_1$ is dense in p-energy and there exist sequences of compact sets $K_n \subset X$ and functions $f_n \in \mathscr{A}$ such that

$$f_n(x) \geq 1 \text{ for every } x \in K_n, \qquad \lim_{n\to\infty} \int_{X\setminus K_n} \Big(1 + |\operatorname{lip} f_n(x)|^p\Big) \,\mathrm{d}\mathfrak{m}(x) = 0, \tag{2.58}$$

then $\mathscr{A}$ is dense in p-energy as well.

The *proof* is a simple adaptation of [44, Proposition 5.3.2]. The next result shows that the algebra generated by (suitable compositions/truncations of) distance functions is always sufficient to generate the Sobolev space $H^{1,2}(X, \mathsf{d}, \mathfrak{m})$.

Theorem 2.16. *Let Y be a dense subset of X and let $\zeta : [0, +\infty) \to [0, +\infty)$ be a Lipschitz nondecreasing function such that $\zeta' > 0$ in an interval $I = (0, r) \subset (0, +\infty)$ and $\zeta \in \mathrm{C}^1(I)$. Then the unital algebra $\mathscr{A}$ generated by the functions $x \mapsto \zeta(\mathsf{d}(x, y))$ is dense in p-energy.*

Proof. Thanks to Remark 2.10, we can assume that d is bounded above by r. It is not difficult to check that $\mathscr{A}$ separates the points of X, so that in order to apply Theorem 2.13, it is enough to check that (2.32) with $f := \mathsf{d}_y$ (recall the notation (2.25)) holds.

Such a property follows immediately from the corresponding estimate on the asymptotic Lipschitz constant: for every $y \in Y$ and $g(x) := \zeta(\mathsf{d}_y(x))$, a simple direct computation shows that

$$\operatorname{lip} g(x) \le \zeta'(\mathsf{d}_y(x)) \quad \text{for every } x \in \mathrm{B}(y, r).$$

Since $g \in \mathscr{A}$ we have $|\mathrm{D}g|_{\star,\mathscr{A}} \le \operatorname{lip} g \le \zeta'(f)$, so that (2.32) holds and we conclude by applying Lemma 2.11. □

We now consider a simple application of Theorem 2.13 to the case when $p = 2$ and $\operatorname{lip} f$ has good properties for functions of $\mathscr{A}$.

Theorem 2.17 *(A Hilbertianity condition). Let $p = 2$ and let $\mathscr{A}$ be a separating unital subalgebra of $\mathrm{Lip}_b(X)$ satisfying (2.35). If for every $f, g \in \mathscr{A}$*

$$\int_X \Big(|\operatorname{lip}(f+g)|^2 + |\operatorname{lip}(f-g)|^2 \Big) \,\mathrm{d}\mathfrak{m} = 2 \int_X \Big(|\operatorname{lip} f|^2 + |\operatorname{lip} g|^2 \Big) \,\mathrm{d}\mathfrak{m}, \tag{2.59}$$

then $H^{1,2}(X, \mathsf{d}, \mathfrak{m})$ is a Hilbert space, CE_2 is a Dirichlet (thus quadratic) form, and $\mathscr{A}$ is strongly dense.

Proof. It is sufficient to prove that the Cheeger energy is a quadratic form in its domain. Thanks to [16, Prop. 11.9] and the 2-homogeneity of CE_2, this property is equivalent to

$$\mathsf{CE}_2(f+g) + \mathsf{CE}_2(f-g) \le 2\mathsf{CE}_2(f) + 2\mathsf{CE}_2(g) \quad \text{for every } f, g \in H^{1,2}(X, \mathsf{d}, \mathfrak{m}). \tag{2.60}$$

We can find two sequences $f_n, g_n \in \mathscr{A}$ such that $f_n \to f$, $g_n \to g$ in $\mathfrak{m}$-measure as $n \to \infty$ and $\operatorname{lip} f_n \to |\mathrm{D}f|_\star$, $\operatorname{lip} g_n \to |\mathrm{D}g|_\star$ in $L^2(X, \mathfrak{m})$. Clearly we have $f_n + g_n \to f + g$, $f_n - g_n \to f - g$ in $\mathfrak{m}$-measure and (2.59) shows that $\operatorname{lip}(f_n + g_n)$ and $\operatorname{lip}(f_n - g_n)$ are uniformly bounded in $L^2(X, \mathfrak{m})$. Up to extracting a suitable sequence, it is not restrictive to assume that $\operatorname{lip}(f_n + g_n) \rightharpoonup G_+ \ge |\mathrm{D}(f+g)|_\star$ and $\operatorname{lip}(f_n - g_n) \rightharpoonup G_- \ge |\mathrm{D}(f-g)|_\star$ $\mathfrak{m}$-a.e. in X. (2.59) then yields

$$\begin{aligned}
\mathsf{CE}_2(f+g) + \mathsf{CE}_2(f-g) &= \int_X |\mathrm{D}(f+g)|_\star^2 \,\mathrm{d}\mathfrak{m} + \int_X |\mathrm{D}(f-g)|_\star^2 \,\mathrm{d}\mathfrak{m} \\
&\le \liminf_{n\to\infty} \int_X |\operatorname{lip}(f_n+g_n)|^2 \,\mathrm{d}\mathfrak{m} + \int_X |\operatorname{lip}(f_n-g_n)|^2 \,\mathrm{d}\mathfrak{m} \\
&= \liminf_{n\to\infty} 2 \int_X |\operatorname{lip} f_n|^2 \,\mathrm{d}\mathfrak{m} + 2 \int_X |\operatorname{lip} g_n|^2 \,\mathrm{d}\mathfrak{m} \\
&= 2\mathsf{CE}_2(f) + \mathsf{CE}_2(g).
\end{aligned}$$

Since $H^{1,2}(X, \mathsf{d}, \mathfrak{m})$ is Banach space, we deduce that $H^{1,2}(X, \mathsf{d}, \mathfrak{m})$ is a Hilbert space, so it is reflexive. This also shows that $\mathscr{A}$ is strongly dense. $\square$

Remark 2.18. In the framework of Theorem 2.17, there exists a scalar product $\langle \cdot, \cdot \rangle_{H^{1,2}}$ on $H^{1,2}(X, \mathsf{d}, \mathfrak{m})$ inducing the norm $\| \cdot \|_{H^{1,2}}$ and satisfying

$$\langle f, g \rangle_{H^{1,2}} = \int_X fg \, \mathrm{d}\mathfrak{m} + \mathsf{CE}_2(f, g) \text{ for every } f, g \in H^{1,2}(X, \mathsf{d}, \mathfrak{m}), \tag{2.61}$$

where $\mathsf{CE}_2(\cdot, \cdot)$ denotes the bilinear form associated to $\mathsf{CE}_2(\cdot)$.

Remark 2.19. If (2.59) holds then the restriction $(\mathsf{pCE}_2, \mathscr{A})$ of pCE_2 to $\mathscr{A}$ is a quadratic form which is induced by a corresponding bilinear form $\mathsf{pCE}_2(\cdot, \cdot)$ defined by the parallelogram rule. We recall that such a form is *closable* (see e.g. [13, §1.3], [32, Chapter I, §.3]) if for any sequence $(f_n)_{n \in \mathbb{N}}$ in $\mathscr{A}$

$$f_n \to 0 \quad \text{in } L^2(X, \mathfrak{m}), \quad \limsup_{m,n \to \infty} \mathsf{pCE}_2(f_n - f_m) = 0 \quad \Rightarrow \quad \lim_{n \to \infty} \mathsf{pCE}_2(f_n) = 0. \tag{2.62}$$

Theorem 2.17 shows in particular that if $(\mathsf{pCE}_2, \mathscr{A})$ is quadratic and closable, then the Cheeger energy $(\mathsf{CE}_2, H^{1,2}(X, \mathsf{d}, \mathfrak{m}))$ coincides with the smallest closed extension of $(\mathsf{pCE}_2, \mathscr{A})$. In this case, trivially, the restriction of CE_2 to $\mathscr{A}$ coincides with pCE_2. Since the Cheeger energy CE_2 is quasi-regular (see [8, Lemma 6.7], [43, Thm. 4.1], [19, Prop. 3.21], [42]), as a by-product we obtain the quasi-regularity of the closure of $(\mathsf{pCE}_2, \mathscr{A})$.

An immediate consequence is the Hilbertianity of $H^{1,2}(\mathbb{H}, \mathsf{d}_{\mathbb{H}}, \mathfrak{m})$ in the case when $(\mathbb{H}, \mathsf{d}_{\mathbb{H}})$ is a separable Hilbert space (in particular $\mathbb{R}^d$) endowed with the distance induced by its Hilbertian norm [20,21,44].

Corollary 2.20. *Let $(\mathbb{H}, \mathsf{d}_{\mathbb{H}})$ be a separable Hilbert space and let $\mathfrak{m}$ be a finite and positive Borel measure on $\mathbb{H}$. Then $H^{1,2}(\mathbb{H}, \mathsf{d}_{\mathbb{H}}, \mathfrak{m})$ is a Hilbert space.*

Proof. Let $\mathscr{A}$ be the algebra $\mathrm{C}^1_b(\mathbb{H})$ of bounded C^1 functions with bounded continuous gradient. It is immediate to check that for every $\phi \in \mathrm{C}^1_b(\mathbb{H})$ we have $\operatorname{lip} \phi(x) = \|\nabla \phi(x)\|_{\mathbb{H}}$ so that pCE_2 is a quadratic form on $\mathscr{A}$, thus satisfying (2.59).

On the other hand $\mathscr{A}$ contains the functions $x \mapsto \tanh(\mathsf{d}^2(x, y))$, $y \in \mathbb{H}$, so that we can apply Theorem 2.16. $\square$

Remark 2.21 *(Density of $\mathrm{C}^\infty_c(\mathbb{R}^d)$ in $H^{1,2}(\mathbb{R}^d, \mathsf{d}, \mathfrak{m})$).* When $\mathbb{H} = \mathbb{R}^d$ is finite dimensional, we can also prove that the algebra $\mathscr{A} = \mathrm{C}^\infty_c(\mathbb{R}^d)$ is strongly dense in $H^{1,2}(\mathbb{R}^d, \mathsf{d}, \mathfrak{m})$. In fact, if ζ is the restriction to $[0, \infty)$ of a smooth nondecreasing transition function $\tilde{\zeta} \in \mathrm{C}^\infty(\mathbb{R})$ satisfying $\tilde{\zeta}(s) = 0$ if $s \le 0$, $\tilde{\zeta}(s) = 1$ is $s > 1$ and $\tilde{\zeta}'(s) > 0$ if $s \in (0, 1)$, it is immediate to check that for every $y \in \mathbb{R}^d$ the functions $\tilde{\zeta}(\mathsf{d}_y)$ belong to $\mathscr{A}_1$, so that $\mathscr{A}_1$ is dense in 2-energy by Theorem 2.16.

On the other hand, being $\mathfrak{m}$ tight, it is easy to check that $\mathscr{A}$ satisfies (2.58), so that we can apply Proposition 2.15.

2.4. Intrinsic distances

By using the general properties of metric Sobolev spaces and the equivalence with the Newtonian viewpoint based on the notion of upper gradient [12,28] it is possible to improve considerably the density result of Corollary 2.14. Let us first recall the notion of *metric velocity*

$$|\dot\gamma|_{\mathsf{d}}(t) := \limsup_{h\to 0} \frac{\mathsf{d}(\gamma(t+h),\gamma(t))}{|h|} \tag{2.63}$$

and *length*

$$\begin{aligned}\ell_{\mathsf{d}}(\gamma,[\alpha,\beta]) := \sup\Big\{ \sum_{n=1}^{N} \mathsf{d}(\gamma(t_{n-1}),\gamma(t_n)) : t_0=\alpha<t_1<\cdots<t_{N-1}<t_N=\beta\Big\} \\ = \int_\alpha^\beta |\dot\gamma|_{\mathsf{d}}(t)\,\mathrm{d}t\end{aligned} \tag{2.64}$$

of a d-Lipschitz curve $\gamma : [a,b] \to X$; here $[\alpha,\beta]\subset[a,b]$ and we just write $\ell_{\mathsf{d}}(\gamma)$ for $\ell_{\mathsf{d}}(\gamma,[a,b])$.

If $Y \subset X$ is a given set, we can introduce the *length (or intrinsic) extended distance* $\mathsf{d}_{Y,\ell}$ induced by d on Y, as the infimum of the length of Y-valued Lipschitz curves connecting two given points $y_0, y_1 \in Y$:

$$\mathsf{d}_{Y,\ell}(y_0,y_1) := \inf\Big\{\ell_{\mathsf{d}}(\gamma) : \gamma\in \mathrm{Lip}([0,1];(Y,\mathsf{d})),\ \gamma(0)=y_0,\ \gamma(1)=y_1\Big\} \tag{2.65}$$

$$= \inf\Big\{\ell>0 : \gamma\in \mathrm{Lip}([0,\ell];(Y,\mathsf{d})),\ \gamma(0)=y_0,\ \gamma(\ell)=y_1,\ |\dot\gamma|_{\mathsf{d}}\le 1 \text{ a.e.}\Big\}. \tag{2.66}$$

Clearly we have

$$\mathsf{d}(y_0,y_1)\le \mathsf{d}_{X,\ell}(y_0,y_1)\le \mathsf{d}_{Y,\ell}(y_0,y_1) \quad \text{for every } y_0,y_1\in Y. \tag{2.67}$$

If $g : X \to [0,+\infty]$ is a Borel function, the integral of g along γ is defined by

$$\int_\gamma g := \int_a^b g(\gamma(t))|\dot\gamma|_{\mathsf{d}}(t)\,\mathrm{d}t. \tag{2.68}$$

It is well known that length and integral are invariant with respect to arc-length reparametrization of γ and it is always possible to find a 1-Lipschitz curve R_γ : $[0, \ell_{\mathsf{d}}(\gamma)] \to X$ such that

$$R_\gamma(\ell_{\mathsf{d}}(\gamma, [a,t])) = \gamma(t) \text{ for every } t \in [a,b], \quad |\dot{R}_\gamma|(s) = 1 \text{ a.e. in } [0, \ell_{\mathsf{d}}(\gamma)], \quad \int_{R_\gamma} g = \int_\gamma g \tag{2.69}$$

for every nonnegative Borel function g (see e.g. [44, Section 3.3]). A Borel function $g : X \to [0, +\infty]$ is an upper gradient of $f : X \to \mathbb{R}$ if

$$|f(\gamma(b)) - f(\gamma(a))| \le \int_\gamma g \quad \text{for every } \gamma \in \mathrm{Lip}([a,b]; (X, \mathsf{d})). \tag{2.70}$$

Functions in $\mathcal{L}^p(X, \mathfrak{m})$ which admit an upper gradient in $\mathcal{L}^p(X, \mathfrak{m})$ characterize the Newtonian Sobolev space $N^{1,p}(X, \mathsf{d}, \mathfrak{m})$ [12,28]. We state here a useful consequence of the main equivalence results [7, Theorem 6.2] [5, Theorem 7.4].

Theorem 2.22. *Let Y be a Borel subset of X of full $\mathfrak{m}$-measure (i.e. $\mathfrak{m}(X \setminus Y) = 0$) satisfying*

$$\gamma \in \mathrm{Lip}([a,b]; (X, \mathsf{d})), \quad R_\gamma(s) \in Y \textit{ for } \mathscr{L}^1\textit{-a.e. } s \in [0, \ell_{\mathsf{d}}(\gamma)] \quad \Rightarrow \quad \gamma([a,b]) \subset Y, \tag{2.71}$$

let $f : X \to \mathbb{R}$ be a $\mathfrak{m}$-measurable function and let $g : Y \to [0, +\infty]$ be a Borel function satisfying

$$|f(\gamma(b)) - f(\gamma(a))| \le \int_\gamma g \quad \textit{for every } \gamma \in \mathrm{Lip}([a,b]; (Y, \mathsf{d})). \tag{2.72}$$

If $\int_Y |g|^p \, \mathrm{d}\mathfrak{m} < \infty$ then f has a p-relaxed gradient and

$$|\mathrm{D}f|_\star \le g \quad \mathfrak{m}\textit{-a.e. in } Y. \tag{2.73}$$

Notice that condition (2.72) is weaker than (2.70), since the upper gradient condition is imposed only along curves taking values in Y; however, starting from any function $g \in \mathcal{L}^p(Y, \mathfrak{m})$ satisfying (2.72) we can define a new Borel function $\tilde{g} : X \to [0, +\infty]$ whose restriction to Y coincides with g such that $\tilde{g}|_{X \setminus Y} \equiv +\infty$. Clearly

$$\int_X \tilde{g}^p \, \mathrm{d}\mathfrak{m} = \int_Y g^p \, \mathrm{d}\mathfrak{m} < +\infty \quad \text{since } \mathfrak{m}(X \setminus Y) = 0.$$

Moreover $\tilde{g}$ is an upper gradient for f according to (2.70): in fact it is sufficient to check (2.70) for those curves γ with $\gamma = R_\gamma$ and $\int_\gamma \tilde{g} < +\infty$; since $\tilde{g}(\gamma(s)) = +\infty$ if $\gamma(s) \notin Y$, we deduce that $\gamma(s) \in Y$ for $\mathscr{L}^1$ -a.e. $s \in [0, \ell_{\mathsf{d}}(\gamma)]$ so that $\gamma \in \mathrm{Lip}([0, \ell_{\mathsf{d}}(\gamma)]; (Y, \mathsf{d}))$ by (2.71), and (2.70) then follows by (2.72).

It is also immediate to check that (2.71) holds if Y is closed.

We consider the situation where

(A) $Y \subset X$ is a Borel set with full $\mathfrak{m}$-measure satisfying (2.71);
(B) a metric $\delta : Y \times Y \to [0, +\infty)$ is given on Y such that (Y, δ) is complete and separable and (recall Remark 2.10)

$$\mathsf{d}_1(y_1, y_2) \le \delta(y_1, y_2) \le \mathsf{d}_{Y,\ell}(y_1, y_2) \quad \text{for every } y_1, y_2 \in Y. \tag{2.74}$$

Remark 2.23 *(Y-intrinsic distance).* δ is intrinsically equivalent to d on Y, i.e. every d-Lipschitz curve $\gamma : [0, 1] \to Y$ is also δ-Lipschitz, its δ-length coincides with the corresponding d-length, and integration along γ does not depend on the choice of the distance. In particular condition (2.72) can be equivalently stated in terms of δ.
To see that these conditions are implied by (2.74), let us fix a d-Lipschitz curve $\gamma : [0, 1] \to Y$ with Lipschitz constant bounded by $L \ge 0$; then

$$\mathsf{d}_{Y,\ell}(\gamma(s), \gamma(t)) \le \ell_{\mathsf{d}}\left(\gamma|_{[s,t]}\right) = \int_s^t |\dot{\gamma}|_{\mathsf{d}}(r)\,\mathrm{d}r \le L|t - s| \quad 0 \le s \le t \le 1,$$

so that γ is $\mathsf{d}_{Y,\ell}$-Lipschitz continuous and thus, by (2.74), also δ-Lipschitz continuous. To see that the δ and the d-lengths of γ coincide, it is enough to show that $\ell_\delta(\gamma) \le \ell_{\mathsf{d}}(\gamma)$, since (2.74) and the trivial equality $\ell_{\mathsf{d}_1}(\gamma) = \ell_{\mathsf{d}}(\gamma)$ already give the other inequality; by (2.74) we immediately have $\ell_\delta(\gamma) \le \ell_{\mathsf{d}_{Y,\ell}}(\gamma)$ and by the very definition of $\mathsf{d}_{Y,\ell}$ we see that $\ell_{\mathsf{d}_{Y,\ell}}(\gamma) \le \ell_{\mathsf{d}}(\gamma)$. Finally, to see that the integral along γ does not depend on the choice of the distance, it is enough to see that $|\dot{\gamma}|_{\mathsf{d}} = |\dot{\gamma}|_\delta$ a.e. in $[0, 1]$. The $\le$ inequality is an immediate consequence of (2.74) and (2.63), while the $\ge$ follows by

$$\frac{\delta(\gamma(s), \gamma(t))}{t - s} \le \frac{\ell_\delta(\gamma|_{[s,t]})}{t - s} = \frac{\ell_{\mathsf{d}}(\gamma|_{[s,t]})}{t - s} = \frac{1}{t - s}\int_s^t |\dot{\gamma}|_{\mathsf{d}}(r)\,\mathrm{d}r \quad 0 \le s < t \le 1,$$

and passing to the limit as $s \to t$ for every Lebesgue point t of $|\dot{\gamma}|_{\mathsf{d}}$.

Since $\mathfrak{m}(X \setminus Y) = 0$ we can identify $L^p(Y, \mathfrak{m})$ with $L^p(X, \mathfrak{m})$. In general, the topology induced by δ is finer than the d topology on Y, and they coincide if δ is continuous w.r.t. d. It is also clear from property (B) that the restriction to Y of every bounded d-Lipschitz function $f : X \to \mathbb{R}$ is also δ-Lipschitz. Thanks to (2.74) (which in particular implies that δ-balls of radius $r < 1$ centered at some point $y \in Y$ are included in d-balls of the same radius and with the same center) it is also clear that

$$\mathrm{lip}_\delta f(y) \le \mathrm{lip}_{\mathsf{d}} f(y) \quad \text{for every } y \in Y,\ f \in \mathrm{Lip}_b(X, \mathsf{d}). \tag{2.75}$$

Since $\mathrm{lip}_\delta f$ is bounded and δ-u.s.c. in Y, it is $\mathfrak{m}$-measurable and we can define the δ pre-Cheeger energy

$$\mathsf{pCE}_{p,\delta}(f) := \int_Y |\,\mathrm{lip}_\delta f(y)|^p \,\mathrm{d}\mathfrak{m}(y) \tag{2.76}$$

and we can still consider its l.s.c. envelope in $L^0(Y, \mathfrak{m})$

$$\mathsf{CE}_{p,\delta,\mathscr{A}}(f) := \inf\Big\{ \liminf_{n\to\infty} \mathsf{pCE}_{p,\delta}(f_n) : f_n \in \mathscr{A},\ f_n \to f \text{ in } L^0(X, \mathfrak{m})\Big\}. \tag{2.77}$$

Theorem 2.24. *Let* $\mathscr{A}(X, \mathsf{d}) := \mathrm{Lip}_b(X, \mathsf{d})$*, let* $\mathscr{A}$ *be a separating unital subalgebra of* $\mathrm{Lip}_b(X, \mathsf{d})$ *satisfying* (2.35) *and assume that* (Y, δ) *satisfies the conditions* (A), (B) *above. Then we have*

$$\mathsf{CE}_{p,\delta,\mathscr{A}(X,\mathsf{d})}(f) = \mathsf{CE}_{p,\delta,\mathscr{A}}(f) = \mathsf{CE}_{p,\mathscr{A}}(f) = \mathsf{CE}_p(f) \quad \textit{for every } f \in L^0(X, \mathfrak{m}). \tag{2.78}$$

In particular, the minimal p-relaxed gradients of $f \in L^0(X, \mathfrak{m})$ *computed w.r.t.* $(\delta, \mathscr{A})$, $(\delta, \mathrm{Lip}_b(Y))$, $(\mathsf{d}, \mathscr{A})$ *or* $(\mathsf{d}, \mathrm{Lip}_b(X))$ *coincide and we have* $D^{1,p}(Y, \delta, \mathfrak{m}) = D^{1,p}(Y, \delta, \mathfrak{m}; \mathscr{A}) = D^{1,p}(X, \mathsf{d}, \mathfrak{m}) = D^{1,p}(X, \mathsf{d}, \mathfrak{m}; \mathscr{A})$.

Proof. Since $\mathsf{pCE}_{p,\delta}(f) \le \int_X \big(\mathrm{lip}_{\mathsf{d}} f(x)\big)^p \,\mathrm{d}\mathfrak{m}$ for every $f \in \mathrm{Lip}_b(X, \mathsf{d})$, we clearly have

$$\mathsf{CE}_{p,\delta,\mathscr{A}(X,\mathsf{d})}(f) \le \mathsf{CE}_{p,\delta,\mathscr{A}}(f) \le \mathsf{CE}_{p,\mathscr{A}}(f) = \mathsf{CE}_p(f) \quad \text{for every } f \in L^0(X, \mathfrak{m}),$$

where the last equality follows from Corollary 2.14. It is then sufficient to prove that $\mathsf{CE}_{p,\delta,\mathscr{A}(X,\mathsf{d})}(f) \ge \mathsf{CE}_p(f)$ in order to get (2.78). Using (2.77) and the $L^0(X, \mathfrak{m})$-lower semicontinuity of CE_p (see Remark 2.6), the latter inequality will be a consequence of

$$\int_Y |\,\mathrm{lip}_\delta f(y)|^p \,\mathrm{d}\mathfrak{m}(y) \ge \mathsf{CE}_p(f) \quad \text{for every } f \in \mathrm{Lip}_b(X, \mathsf{d}). \tag{2.79}$$

In order to prove (2.79) it is sufficient to apply Theorem 2.22 and prove that the Borel function $g := \mathrm{lip}_\delta f$ satisfies (2.72). Now we use the fact that the restriction to Y of a function $f \in \mathrm{Lip}_b(X, \mathsf{d})$ belongs to $\mathrm{Lip}_b(Y, \delta)$ and every d-Lipschitz curve γ with values in Y is also δ-Lipschitz, the respective lengths coincide and therefore also the arc-length reparametrizations are the same. Since lip_δ is an upper gradient we thus obtain

$$|f(\gamma(b)) - f(\gamma(a))| \leq \int_\gamma \operatorname{lip}_\delta f \quad \text{for every } \gamma \in \operatorname{Lip}([a,b];(Y,\delta)). \quad \square$$

By Theorem 2.22 and also using that $\mathfrak{m}(X \setminus Y) = 0$, we conclude.

Combining Theorem 2.24 with Corollary 2.20 we recover the following result of [31].

Corollary 2.25. *Let $(\mathbb{M}, \mathsf{d}_{\mathbb{M}})$ be a complete Riemannian manifold endowed with the canonical Riemannian distance and let $\mathfrak{m}$ be a finite and positive Borel measure on $\mathbb{M}$. Then $H^{1,2}(\mathbb{M}, \mathsf{d}_{\mathbb{M}}, \mathfrak{m})$ is a Hilbert space and $\mathrm{C}_c^\infty(\mathbb{M})$ is dense in $H^{1,2}(\mathbb{M}, \mathsf{d}_{\mathbb{M}}, \mathfrak{m})$.*

Proof. By Nash isometric embedding Theorem [36] we can find a dimension d, and an isometric embedding $\jmath : \mathbb{M} \to \jmath(\mathbb{M}) \subset \mathbb{R}^d$.

Since $\mathbb{M}$ is complete and $\jmath$ is an imbedding, $M := \jmath(\mathbb{M})$ is a closed subset of $\mathbb{R}^d$ and the (Riemannian) metric d_M inherited by $\mathsf{d}_{\mathbb{M}}$ given by $\mathsf{d}_M(\jmath(x), \jmath(y)) := \mathsf{d}_{\mathbb{M}}(x,y)$ is an isometry. In particular d_M induces on M the relative topology of $\mathbb{R}^d$ and (M, d_M) is a complete and separable metric space. Setting $\tilde{\mathfrak{m}} := \jmath_\sharp \mathfrak{m}$, it is clear that the map $\jmath^* : f \to f \circ \jmath$ is a linear isometric isomorphism between $H^{1,2}(M, \mathsf{d}_M, \tilde{\mathfrak{m}})$ and $H^{1,2}(\mathbb{M}, \mathsf{d}_{\mathbb{M}}, \mathfrak{m})$. It is then sufficient to prove the statement for $H^{1,2}(M, \mathsf{d}_M, \tilde{\mathfrak{m}})$.

We can now apply Theorem 2.24 with the choices $(Y, \delta) := (M, \mathsf{d}_M)$ and $X = \mathbb{R}^d$ endowed with the Euclidean distance d. Condition (A) clearly holds since M is closed in $\mathbb{R}^d$ and $\tilde{\mathfrak{m}}$ is supported on M. Similarly, also (B) holds since $\jmath$ is an isometric immersion.

Remark 2.21 shows that $\mathrm{C}_c^\infty(\mathbb{R}^d)$ is dense in $H^{1,2}(\mathbb{R}^d, \mathsf{d}_{\mathbb{R}^d}, \tilde{\mathfrak{m}})$ so that $\jmath^*\big(\mathrm{C}_c^\infty(\mathbb{R}^d)\big) \subset \mathrm{C}_c^\infty(\mathbb{M})$ is dense in $H^{1,2}(\mathbb{M}, \mathsf{d}_{\mathbb{M}}, \mathfrak{m})$. $\square$

3. Wasserstein spaces

In this section we list some properties of Wasserstein spaces we will use in the sequel. A complete account of this matter can be found e.g. in [49,4].
If (X, d) is a complete and separable metric space, we denote by $\mathcal{P}(X)$ the space of Borel probability measures on X and by $\mathcal{P}_2(X)$, the set

$$\mathcal{P}_2(X) := \left\{ \mu \in \mathcal{P}(X) \mid \int_X \mathsf{d}^2(x, x_0) \mathrm{d}\mu(x) < +\infty \text{ for some } x_0 \in X \right\}.$$

Given $\mu, \nu \in \mathcal{P}(X)$ the set of transport plans between μ and ν is denoted by $\Gamma(\mu,\nu)$ and defined as

$$\Gamma(\mu,\nu) := \left\{ \boldsymbol{\mu} \in \mathcal{P}(X \times X) \mid \pi^1_\sharp \boldsymbol{\mu} = \mu,\ \pi^2_\sharp \boldsymbol{\mu} = \nu \right\},$$

where $\pi^i(x_1, x_2) = x_i$ for every $(x_1, x_2) \in X \times X$ and $\sharp$ denotes the push forward operator. The L^2-Wasserstein distance W_2 between $\mu, \nu \in \mathcal{P}_2(X)$ is defined as

$$W_2^2(\mu,\nu) := \inf\left\{ \int_{X\times X} \mathsf{d}^2 \,\mathrm{d}\boldsymbol{\mu} \mid \boldsymbol{\mu} \in \Gamma(\mu,\nu) \right\}.$$

It is well known that the infimum above is attained in a non-empty and convex set $\Gamma_o(\mu,\nu) \subset \Gamma(\mu,\nu)$; elements of $\Gamma_o(\mu,\nu)$ are called optimal transport plans.

The space $(\mathcal{P}_2(X), W_2)$ is complete and separable and its topology is stronger than the narrow topology, the latter being defined as the coarsest topology on $\mathcal{P}(X)$ making the maps

$$\mu \mapsto \int_X \varphi \,\mathrm{d}\mu$$

continuous for every $\varphi \in \mathrm{C}_b(X)$, the space of continuous and bounded functions on X. In particular, for a sequence $(\mu_n)_n \subset \mathcal{P}_2(X)$ and a point $\mu \in \mathcal{P}_2(X)$, we have

$$W_2(\mu_n,\mu) \to 0 \Leftrightarrow \begin{cases} \displaystyle\int_X \mathsf{d}^2(x,x_0)\mathrm{d}\mu_n(x) \to \int_{\mathbb{R}^d} \mathsf{d}^2(x,x_0)\mathrm{d}\mu(x) & \text{for some } x_0 \in X, \\ \mu_n \to \mu \quad \text{narrowly in } \mathcal{P}(X). \end{cases} \tag{3.1}$$

Moreover, the Wasserstein distance is narrowly lower semicontinuous, meaning that, if $(\mu_n)_n$ and $(\mu'_n)_n$ are two sequences in $\mathcal{P}_2(X)$, $\mu,\mu' \in \mathcal{P}_2(X)$ and $\mu_n \to \mu$, $\mu'_n \to \mu'$ narrowly in $\mathcal{P}(X)$, then we have

$$\liminf_{n\to\infty} W_2(\mu_n,\mu'_n) \geq W_2(\mu,\mu').$$

The following Theorem is [4, Theorem 8.3.1, Proposition 8.4.5 and Proposition 8.4.6] in case $X = \mathbb{R}^d$. Recall that for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$

$$\operatorname{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d) := \overline{\{\nabla\varphi \mid \varphi \in \mathrm{C}_c^\infty(\mathbb{R}^d)\}}^{L^2(\mathbb{R}^d,\mu;\mathbb{R}^d)}. \tag{3.2}$$

Theorem 3.1 *(Wasserstein velocity field). Let $(\mu_t)_{t\in\mathcal{J}} \subset \mathcal{P}_2(\mathbb{R}^d)$ be a locally absolutely continuous curve defined in an open interval $\mathcal{J} \subset \mathbb{R}$. There exists a Borel vector field $v : \mathcal{J}\times\mathbb{R}^d \to \mathbb{R}^d$ and a set $A((\mu_t)_{t\in\mathcal{J}}) \subset \mathcal{J}$ with $\mathscr{L}^1(\mathcal{J} \setminus A((\mu_t)_{t\in\mathcal{J}})) = 0$ such that for every $t \in A((\mu_t)_{t\in\mathcal{J}})$*

$$v_t \in \operatorname{Tan}_{\mu_t} \mathcal{P}_2(\mathbb{R}^d), \quad \int_{\mathbb{R}^d} |v_t|^2 \,\mathrm{d}\mu_t = |\dot\mu_t|^2 = \lim_{h\to 0} \frac{W_2^2(\mu_{t+h},\mu_t)}{h^2},$$

and the continuity equation

$$\partial_t \mu_t + \nabla\cdot(v_t\mu_t) = 0$$

holds in the sense of distributions in $\mathcal{J} \times \mathbb{R}^d$*. Moreover,* v_t *is uniquely determined in* $L^2(\mathbb{R}^d, \mu_t; \mathbb{R}^d)$ *for* $t \in A((\mu_t)_{t\in\mathcal{J}})$ *and*

$$\lim_{h\to 0} \frac{W_2((\boldsymbol{i}_{\mathbb{R}^d} + hv_t)_\sharp \mu_t, \mu_{t+h})}{|h|} = 0 \quad \textit{for every } t \in A((\mu_t)_{t\in\mathcal{J}}), \tag{3.3}$$

where $\boldsymbol{i}_{\mathbb{R}^d}$ *is the identity map on* $\mathbb{R}^d$*.*

3.1. Kantorovich duality and estimates for Kantorovich potentials

The Kantorovich duality for the Wasserstein distance states that

$$W_2^2(\mu,\nu) = \sup\left\{ \int_X u \,\mathrm{d}\mu + \int_X v \,\mathrm{d}\nu \mid (u,v) \in \mathrm{Adm}_2(X) \right\} \quad \text{for every } \mu,\nu \in \mathcal{P}_2(X), \tag{3.4}$$

where $\mathrm{Adm}_2(X)$ is the set of pairs $(u,v) \in \mathrm{C}_b(X) \times \mathrm{C}_b(X)$ such that

$$u(x) + v(y) \le \mathsf{d}^2(x,y) \quad \text{for every } x,y \in X.$$

It is easy to check that for every $f \in \mathrm{Lip}(X,\mathsf{d})$

$$\int_X f \,\mathrm{d}(\mu-\nu) \le \mathrm{Lip}(f,X) W_2(\mu,\nu), \tag{3.5}$$

since choosing $\boldsymbol{\mu} \in \Gamma_o(\mu,\nu)$ and setting $L := \mathrm{Lip}(f,X)$,

$$\int_X f \,\mathrm{d}(\mu-\nu) = \int (f(x) - f(y)) \,\mathrm{d}\boldsymbol{\mu}(x,y) \le L \int \mathsf{d} \,\mathrm{d}\boldsymbol{\mu} \le L\Big(\int \mathsf{d}^2 \,\mathrm{d}\boldsymbol{\mu}\Big)^{1/2} = LW_2(\mu,\nu).$$

When $X = \mathbb{R}^d$, we denote by $\mathcal{P}_2^r(\mathbb{R}^d)$ the subset of $\mathcal{P}_2(\mathbb{R}^d)$ of probability measures that are absolutely continuous w.r.t. the d-dimensional Lebesgue measure. We also set

$$\mathsf{m}_2^2(\mu) := \int_{\mathbb{R}^d} |x|^2 \,\mathrm{d}\mu(x) = W_2^2(\mu,\delta_0). \tag{3.6}$$

The next result uses the celebrated Brenier-Knott-Smith Theorem [48, Section 3] to collect various useful properties of the optimal potentials realizing the supremum in (3.4) in a particular geometric situation. We will use the elementary property that

$$\text{if } u : \overline{B(0,R)} \to [-\infty,+\infty) \text{ is concave with } u(0) > -\infty \text{ then } \quad \sup_{\overline{\mathrm{B}(0,R)}} u = \sup_{\mathrm{B}(0,R)} u, \tag{3.7}$$

which follows by the fact that for every $y_0 \in \partial\mathrm{B}(0,R)$ the concavity of $t \mapsto u(ty_0)$ in $[0,1]$ yields $u(y_0) \le \sup_{0\le t<1} u(ty_0)$.

Theorem 3.2. *Let $\mu, \nu \in \mathcal{P}_2^r(\mathbb{R}^d)$ with $\operatorname{supp}\nu = \overline{\mathrm{B}(0,R)}$ for some $R > 0$. Then there exists a unique pair of continuous and convex functions*

$$\varphi = \Phi(\nu,\mu) : \mathrm{B}(0,R) \to \mathbb{R}, \quad \varphi^* = \Phi^*(\nu,\mu) : \mathbb{R}^d \to \mathbb{R} \tag{3.8}$$

such that

(i) *φ^* is R-Lipschitz and*

$$\varphi^*(y) = \sup_{x \in \mathrm{B}(0,R)} \langle x, y\rangle - \varphi(x) \qquad \text{for every } y \in \mathbb{R}^d, \tag{3.9}$$

$$\varphi(x) = \sup_{y \in \mathbb{R}^d} \langle y, x\rangle - \varphi^*(y) \qquad \text{for every } x \in \mathrm{B}(0,R), \tag{3.10}$$

(ii) $\varphi^*(0) = \inf_{\mathrm{B}(0,R)} \varphi = 0,$

(iii) $\displaystyle\int_{\mathrm{B}(0,R)} \varphi \,\mathrm{d}\nu + \int_{\mathbb{R}^d} \varphi^* \,\mathrm{d}\mu = \frac{1}{2}\mathsf{m}_2^2(\nu) + \frac{1}{2}\mathsf{m}_2^2(\mu) - \frac{1}{2}W_2^2(\nu,\mu).$

Moreover the pair (φ, φ^*) satisfies

$$W_2^2(\mu,\nu) = \int_{\mathrm{B}(0,R)} |x - \nabla\varphi(x)|^2 \,\mathrm{d}\nu(x) = \int_{\mathbb{R}^d} |y - \nabla\varphi^*(y)|^2 \,\mathrm{d}\mu(y). \tag{3.11}$$

Proof. Let us set $D := \mathrm{B}(0,R)$. We know (see e.g. [48, Theorem 2.9, Lemma 2.10]) that there exists a pair (ϕ, ϕ^*) of lower semicontinuous proper conjugate functions such that $\phi \in L^1(\overline{D}, \nu; (-\infty, +\infty])$, $\phi^* \in L^1(\mathbb{R}^d, \mu; (-\infty, +\infty])$ and it holds

$$\int_{\overline{D}} \phi \,\mathrm{d}\nu + \int_{\mathbb{R}^d} \phi^* \,\mathrm{d}\mu = \frac{1}{2}\mathsf{m}_2^2(\nu) + \frac{1}{2}\mathsf{m}_2^2(\mu) - \frac{1}{2}W_2^2(\nu,\mu), \tag{3.12}$$

where

$$\phi^*(y) := \sup_{x \in \overline{D}} \langle x, y\rangle - \phi(x) = \max_{x \in \overline{D}} \langle x, y\rangle - \phi(x). \tag{3.13}$$

Recalling that ϕ is bounded from below by an affine mapping (and thus it is uniformly bounded from below in $\overline{D}$) we immediately see that ϕ^* takes values in $\mathbb{R}$ and it is R-Lipschitz. Up to adding a suitable constant to ϕ we can also suppose that $\phi^*(0) = 0$.

We want to show that the restriction φ of ϕ to $\mathrm{B}(0,R)$ combined with ϕ^* satisfies conditions (i), (ii), and (iii).

(i): Since $\int_{\overline{D}} \phi \,\mathrm{d}\nu < +\infty$ and ν has full support, we deduce that the proper domain of ϕ $\{x \in \overline{D} : \phi(x) < +\infty\}$ is dense in $\overline{D}$; since the proper domain of a l.s.c. and convex function is convex and contains the interior of its closure, we deduce that $\phi(x) < +\infty$ for every $x \in D$ and $\varphi := \phi_{|D}$ is continuous in D.

(3.7) shows that the supremum defining ϕ^* in (3.13) can be restricted to D

$$\phi^*(y) = \sup_{x\in D} \langle x, y\rangle - \varphi(x), \tag{3.14}$$

so that (3.9) holds. (3.10) is just an application of Fenchel-Moreau Theorem $\phi = \phi^{**}$.

(ii): simply follows by (3.9) and the fact that $\phi^*(0) = 0$.

(iii) It is sufficient to notice that the first integral in (3.12) can be restricted to D since $\nu(\partial D) = 0$. The equality (3.11) follows by [48, Theorem 2.12].

Let us show that points (i)-(iii) are also sufficient to get uniqueness. If (φ_0, φ_0^*) is another pair as in the statement satisfying points (i)-(iii), then [48, Theorem 2.12] yields that both $\nabla\varphi$ and $\nabla\varphi_0$ are optimal transport maps from ν to μ, implying that $\nabla\varphi_0 = \nabla\varphi$ $\mathscr{L}^d$-a.e. in $\mathrm{B}(0, R)$ by the a.e. uniqueness of the optimal transport map. Since $\inf_{\mathrm{B}(0,R)} \varphi = \inf_{\mathrm{B}(0,R)} \varphi_0 = 0$ by (ii), we get that $\varphi = \varphi_0$ in $\mathrm{B}(0, R)$ and therefore $\varphi^* = \varphi_0^*$ in $\mathbb{R}^d$ by (3.9). $\square$

The next Lemma collects useful estimates on convex functions; we set $\omega_d := \mathscr{L}^d\big(\mathrm{B}(0, 1)\big)$.

Lemma 3.3. *Let $R, I > 0$ and let $\varphi : \mathrm{B}(0, R) \to \mathbb{R}$, $\psi : \mathbb{R}^d \to \mathbb{R}$, be two (continuous and) convex functions satisfying*

$$\begin{gathered} |\psi(y)| \le R|y| \quad \textit{for every } y \in \mathbb{R}^d, \\ \varphi(x) = \sup_{y\in\mathbb{R}^d} \langle x, y\rangle - \psi(y) \textit{ for every } x \in \mathrm{B}(0, R), \quad \int\limits_{\mathrm{B}(0,R)} \varphi(x)\,\mathrm{d}x \le I. \end{gathered} \tag{3.15}$$

Then φ is nonnegative and satisfy the uniform bounds

$$\sup_{|x|\le r} \varphi(x) \le \frac{I}{\omega_d (R - r)^d}, \quad \mathrm{Lip}\left(\varphi, \overline{\mathrm{B}(0, r)}\right) \le \frac{2^{d+1} I}{\omega_d (R - r)^{d+1}} \quad 0 < r < R, \tag{3.16}$$

and

$$\psi \textit{ is } R\textit{-Lipschitz}, \quad \psi(y) = \sup_{x\in\mathrm{B}(0,R)} \langle y, x\rangle - \varphi(x) \quad \textit{for every } y \in \mathbb{R}^d. \tag{3.17}$$

Proof. Notice that $\psi(0) = 0$ yields $\varphi \ge 0$; the integral estimate of (3.15) and Jensen inequality yield for every $x \in \mathrm{B}(0, R)$ with $\varrho := R - |x|$

$$\varphi(x) \le \frac{1}{\omega_d \varrho^d} \int\limits_{\mathrm{B}(x,\varrho)} \varphi(z)\,\mathrm{d}z \le \frac{I}{\omega_d \varrho^d}, \quad \text{so that} \quad \max_{|x|\le r_0} \varphi(x) \le \frac{I}{\omega_d (R - r_0)^d} \tag{3.18}$$

for every $0 < r_0 < R$. Still using the fact that φ is nonnegative, (3.18) with $r_0 := \frac{1}{2}R + \frac{1}{2}r$ and the estimate of [22, Corollary 2.4] yield the Lipschitz bound of (3.16).

The Legendre transform of ψ defined by $\psi^*(x) := \sup_{y \in \mathbb{R}^d} \langle x, y\rangle - \psi(y)$ coincides with φ in $\mathrm{B}(0,R)$ (in particular it is finite in $\mathrm{B}(0,R)$) and takes the value $+\infty$ for every $x \in \mathbb{R}^d$ with $|x| > R$, since

$$\psi^*(x) \geq \sup_{y \in \mathbb{R}^d} \langle x, y\rangle - R|y| = +\infty \quad \text{if } |x| > R.$$

Fenchel-Moreau Theorem and (3.7) then yield

$$\begin{aligned}\psi(y) &= \sup_{x \in \mathbb{R}^d} \langle x, y\rangle - \psi^*(x) = \sup_{x \in \overline{\mathrm{B}(0,R)}} \langle x, y\rangle - \psi^*(x) = \sup_{x \in \mathrm{B}(0,R)} \langle x, y\rangle - \psi^*(x) \\ &= \sup_{x \in \mathrm{B}(0,R)} \langle x, y\rangle - \varphi(x)\end{aligned}$$

thus showing (3.17); in particular we get that that ψ is R-Lipschitz. $\square$

We conclude this part with the study of the stability properties of pairs of potentials.

Lemma 3.4. *Let $R, I > 0$ and let $\varphi_n : \mathrm{B}(0,R) \to \mathbb{R}$, $\psi_n : \mathbb{R}^d \to \mathbb{R}$, $n \in \mathbb{N}$, be two sequences of (continuous and) convex functions satisfying for every $n \in \mathbb{N}$:*

$$\begin{gathered}|\psi_n(y)| \leq R|y| \quad \text{for every } y \in \mathbb{R}^d, \\ \varphi_n(x) = \sup_{y \in \mathbb{R}^d} \langle x, y\rangle - \psi_n(y) \text{ for every } x \in \mathrm{B}(0,R), \qquad \int_{\mathrm{B}(0,R)} \varphi_n(x)\,\mathrm{d}x \leq I.\end{gathered} \tag{3.19}$$

Then there exist a subsequence $j \mapsto n(j)$ and two convex and continuous functions $\varphi : \mathrm{B}(0,R) \to \mathbb{R}$ and $\psi : \mathbb{R}^d \to \mathbb{R}$ such that

(i) *$\varphi_{n(j)} \to \varphi$ locally uniformly on $\mathrm{B}(0,R)$;*
(ii) *$\psi_{n(j)} \to \psi$ locally uniformly in $\mathbb{R}^d$;*
(iii) *ψ is R-Lipschitz and $\nabla\psi_{n(j)} \to \nabla\psi$ $\mathscr{L}^d$-a.e. on $\mathbb{R}^d$.*

Moreover the pair (φ, ψ) satisfies (3.15), (3.16), and (3.17).

Proof. Thanks to (3.19) and Lemma 3.3, the sequence of pairs (φ_n, ψ_n) satisfies the equicontinuity estimates (3.16) and (3.17) with constants R, I independent of n.

By Arzelà-Ascoli Theorem, we can find a subsequence $j \mapsto n(j)$ and convex and continuous functions $\varphi : \mathrm{B}(0,R) \to \mathbb{R}$ and $\psi : \mathbb{R}^d \to \mathbb{R}$ such that $\varphi_{n(j)} \to \varphi$ and $\psi_{n(j)} \to \psi$ locally uniformly in their respective domains.

In particular, $\psi_{n(j)}$ Mosco converges (see e.g. [10, Definition 3.17, Proposition 3.19]) to ψ and therefore the sequence of its Legendre transforms $\psi^*_{n(j)}$ Mosco converges to ψ^*

([10, Theorem 3.18]). Since ψ_n^* coincides with φ_n in $\mathrm{B}(0,R)$ and $\varphi_{n(j)}$ converge locally uniformly to φ, we deduce that φ coincides with ψ^* in $\mathrm{B}(0,R)$:

$$\varphi(x) = \sup_{y\in\mathbb{R}^d} \langle x,y\rangle - \psi(y) \quad \text{for every } x \in \mathrm{B}(0,R).$$

By Fatou's Lemma φ also satisfies the integral bound of (3.15). A further application of Lemma 3.3 yields (3.16) and (3.17).

Finally, the local uniform convergence of $\psi_{n(j)}$ to ψ gives [40, Theorem 24.5] the pointwise convergence of $\nabla\psi_{n(j)}(x)$ to $\nabla\psi(x)$ at every point $x \in \mathbb{R}^d$ where all the $\psi_{n(j)}$ and ψ are differentiable. This proves (iii) and concludes the proof of the Lemma. □

4. The Wasserstein Sobolev space $H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m})$

In this section we consider the metric space $\mathcal{P}_2(\mathbb{R}^d)$, endowed with the L^2-Wasserstein distance $\mathsf{d} = W_2$ and a finite positive Borel measure $\mathfrak{m}$. We will denote by $\mathbb{W}_2 = \mathbb{W}_2(\mathbb{R}^d, \mathfrak{m})$ the metric-measure space $(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m})$ and we want to study the Wasserstein Sobolev space $H^{1,2}(\mathbb{W}_2)$.

We will show that $H^{1,2}(\mathbb{W}_2)$ is Hilbertian (and therefore the metric space $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ is infinitesimally Hilbertian) and its functions admit a nice approximation in terms of the distinguished algebra of cylinder functions.

4.1. The algebra of C^1-cylinder functions

We denote by $\mathrm{C}^1_b(\mathbb{R}^d)$ the space of bounded and Lipschitz C^1 functions $\phi : \mathbb{R}^d \to \mathbb{R}$. This in particular implies that $\sup_{x\in\mathbb{R}^d} |\phi(x)| + |\nabla\phi(x)| < +\infty$ if $\phi \in \mathrm{C}^1_b(\mathbb{R}^d)$. Every $\phi \in \mathrm{C}^1_b(\mathbb{R}^d)$ induces the function L_ϕ on $\mathcal{P}(\mathbb{R}^d)$

$$\mathsf{L}_\phi : \mu \to \int_{\mathbb{R}^d} \phi \,\mathrm{d}\mu \tag{4.1}$$

which clearly belongs to $\mathrm{Lip}_b(\mathcal{P}_2(\mathbb{R}^d), W_2)$ thanks to (3.5). More generally, if $\boldsymbol{\phi} = (\phi_1, \cdots, \phi_N) \in \big(\mathrm{C}^1_b(\mathbb{R}^d)\big)^N$, we denote by $\mathsf{L}_{\boldsymbol{\phi}} := (\mathsf{L}_{\phi_1}, \cdots, \mathsf{L}_{\phi_N})$ the corresponding map from $\mathcal{P}_2(\mathbb{R}^d)$ to $\mathbb{R}^N$.

Our construction is based on the algebra of C^1- cylinder functions generated by (4.1) via composition with C^1 functions and it is quite similar to the one of [17, Section 2] (see also [50]). Working in the flat space $\mathbb{R}^d$ allows for a further simplification in the structure of the tangent bundle and of corresponding vector fields.

Definition 4.1 *(C^1-Cylinder functions).* We say that a function $F : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is a C^1-*cylinder function* if there exist $N \in \mathbb{N}$, $\psi \in \mathrm{C}^1_b(\mathbb{R}^N)$ and $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_N) \in (\mathrm{C}^1_b(\mathbb{R}^d))^N$ such that

$$F(\mu) = \psi(\mathsf{L}_{\boldsymbol{\phi}}(\mu)) = \psi\big(\mathsf{L}_{\phi_1}(\mu), \cdots, \mathsf{L}_{\phi_N}(\mu)\big) \quad \text{for every } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{4.2}$$

We denote the set of such functions by $\mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$.

Remark 4.2. Notice that $\mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ is a unital subalgebra of $\mathrm{Lip}_b(\mathcal{P}_2(\mathbb{R}^d), W_2)$. One could also consider the smaller algebra $\mathrm{FC}_b^1\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ (resp. $\mathrm{FC}_c^\infty\left(\mathcal{P}_2(\mathbb{R}^d)\right)$) generated by functions as in (4.1) (resp. by functions as in (4.1) where $\phi \in \mathrm{C}_c^\infty(\mathbb{R}^d)$), thus restricting ψ to be a polynomial in (4.2). This means that every element $F \in \mathrm{FC}_b^1\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ (resp. $\mathrm{FC}_c^\infty\left(\mathcal{P}_2(\mathbb{R}^d)\right)$) can be written as

$$F = \psi \circ \mathsf{L}_{\boldsymbol{\phi}}$$

for some ψ polynomial in $\mathbb{R}^N$, $\boldsymbol{\phi} \in (\mathrm{C}_b^1(\mathbb{R}^d))^N$ (resp. $(\mathrm{C}_c^\infty(\mathbb{R}^d)^N)$ and $N \in \mathbb{N}$, $N \geq 1$. We prefer at this stage the choice of $\mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$, since it simplifies some technical points. However, Proposition 4.19 shows that using $\mathrm{FC}_b^1\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ or $\mathrm{FC}_c^\infty\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ will lead to the same conclusions.

Remark 4.3. Since for every $\boldsymbol{\phi} \in \left(\mathrm{C}_b^1(\mathbb{R}^d)\right)^N$ the range of $\mathsf{L}_{\boldsymbol{\phi}}$ is always contained in the bounded set $[-M, M]^N$ where $M := \max_{i=1,\ldots,d} \|\phi_i\|_\infty$, also functions $F = \psi \circ \mathsf{L}_{\boldsymbol{\phi}}$ with $\psi \in \mathrm{C}^1(\mathbb{R}^N)$ belong to $\mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$. Indeed it is enough to consider a function $\tilde{\psi} \in \mathrm{C}_b^1(\mathbb{R}^N)$ coinciding with ψ on $[-M, M]^N$ and equal to 0 outside $[-M-1, M+1]^N$ so that $F = \tilde{\psi} \circ \mathsf{L}_{\boldsymbol{\phi}}$. In particular every function of the form L_ϕ, $\phi \in \mathrm{C}_b^1(\mathbb{R}^d)$, belongs to $\mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$.

Let us consider the set

$$\mathcal{D} := \Big\{(\mu, x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d : x \in \mathrm{supp}(\mu)\Big\}. \tag{4.3}$$

The set $\mathcal{D}$ is a Borel set (in fact it is a G_δ): if $(r_n)_n = \mathbb{Q} \cap (0, +\infty)$ we have that $\mathcal{D} = \cap_n \mathcal{D}_n$, where

$$\mathcal{D}_n := \left\{(\mu, x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d : \mu(\mathrm{B}(x, r_n)) > 0\right\},$$

and each $\mathcal{D}_n$ is open in $\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$, being the inverse image of $(0, +\infty)$ through the lower semicontinuous map $(\mu, x) \mapsto \mu(\mathrm{B}(x, r))$.

Definition 4.4. If $F = \psi \circ \mathsf{L}_{\boldsymbol{\phi}} \in \mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ as in (4.2) for some $N \in \mathbb{N}$, $\psi \in \mathrm{C}_b^1(\mathbb{R}^N)$ and $\boldsymbol{\phi} \in (\mathrm{C}_b^1(\mathbb{R}^d))^N$, then the Wasserstein differential of F, $\mathrm{D}F : \overline{\mathcal{D}} \to \mathbb{R}^d$, is defined by

$$\mathrm{D}F(\mu, x) := \sum_{n=1}^N \partial_n \psi\left(\mathsf{L}_{\boldsymbol{\phi}}(\mu)\right) \nabla \phi_n(x), \quad (\mu, x) \in \overline{\mathcal{D}}. \tag{4.4}$$

We will also denote by $\mathrm{D}F[\mu]$ the function $x \mapsto \mathrm{D}F(\mu, x)$ and we will set

$$\|\mathrm{D}F\,[\mu]\|_{\mu}^{2} := \int_{\mathbb{R}^d} |\mathrm{D}F[\mu](x)|^2 \mathrm{d}\mu(x), \quad \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{4.5}$$

Remark 4.5. It is not difficult to check that

$$\mathrm{D}F \text{ is continuous in } \overline{\mathcal{D}} \tag{4.6}$$

with respect to the natural product (narrow and euclidean) topology of $\mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d$.

In principle $\mathrm{D}F$ (and thus $\|\mathrm{D}F\,[\mu]\|_{\mu}$) may depend on the choice of $N \in \mathbb{N}$, $\psi \in \mathrm{C}_b^1(\mathbb{R}^N)$ and $\boldsymbol{\phi} \in (\mathrm{C}_b^1(\mathbb{R}^d))^N$ used to represent F. In Proposition 4.9 we show that for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ the function $\mathrm{D}F[\mu]$ is uniquely characterized in $\mathrm{supp}(\mu)$ and $\|\mathrm{D}F\,[\mu]\|_{\mu}$ is well defined, so that $\mathrm{D}F$ is uniquely characterized by F in $\mathcal{D}$. By (4.6), $\mathrm{D}F$ is also uniquely characterized by F on $\overline{\mathcal{D}}$.

We have seen that the Wasserstein differential $\mathrm{D}F$ can be considered as a map from $\overline{\mathcal{D}}$ with values in $\mathbb{R}^d$. It is natural to introduce the measure $\boldsymbol{\mathfrak{m}} = \int \delta_\mu \otimes \mu \,\mathrm{d}\mathfrak{m}(\mu) \in \mathcal{P}(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d)$ obtained integrating the measures μ w.r.t. $\mathfrak{m}$: for every bounded Borel function $H : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}$ we have

$$\int H(\mu, x)\,\mathrm{d}\boldsymbol{\mathfrak{m}}(\mu, x) = \int_{\mathcal{P}_2(\mathbb{R}^d)} \Big(\int_{\mathbb{R}^d} H(\mu, x)\,\mathrm{d}\mu(x) \Big)\,\mathrm{d}\mathfrak{m}(\mu). \tag{4.7}$$

Since $\mathrm{supp}(\boldsymbol{\mathfrak{m}}) \subset \overline{\mathcal{D}}$, it is then clear that $\mathrm{D}F$ belongs to $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$ and

$$\|\mathrm{D}F\|^2_{L^2(\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d,\boldsymbol{\mathfrak{m}};\mathbb{R}^d)} = \int_{\mathcal{P}_2(\mathbb{R}^d)} \|\mathrm{D}F\,[\mu]\|_{\mu}^{2}\,\mathrm{d}\mathfrak{m}(\mu) = \int_{\overline{\mathcal{D}}} |\mathrm{D}F(\mu, x)|^2\,\mathrm{d}\boldsymbol{\mathfrak{m}}(\mu, x). \tag{4.8}$$

Lemma 4.6. *Let Y be a Polish space and let $G : \mathcal{P}(Y) \times Y \to [0, +\infty)$ be a bounded and continuous function. If $(\mu_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{P}(Y)$ narrowly converging to μ as $n \to +\infty$, then*

$$\lim_{n\to\infty} \int_Y G(\mu_n, y)\mathrm{d}\mu_n(y) = \int_Y G(\mu, y)\mathrm{d}\mu(y).$$

Proof. We set $g_n(x) := G(\mu_n, x)$, $g(x) := G(\mu, x)$. Since G is continuous, g_n converge uniformly to g on compact subsets of Y as $n \to \infty$. Thanks to [4, Lemma 5.2.1] $(g_n)_\sharp\mu_n$ converge narrowly to $g_\sharp\mu$ in $\mathcal{P}(\mathbb{R})$. On the other hand, the support of $(g_n)_\sharp\mu_n$ is uniformly bounded because G is bounded so that

$$\lim_{n\to\infty} \int_Y G(\mu_n, y)\mathrm{d}\mu_n(y) = \lim_{n\to\infty} \int_{\mathbb{R}} r\,\mathrm{d}((g_n)_\sharp\mu_n)(r) = \int_{\mathbb{R}} r\,\mathrm{d}(g_\sharp\mu)(r) = \int_Y G(\mu, y)\,\mathrm{d}\mu(y). \quad \square$$

Lemma 4.7. *Let $F = \psi \circ \mathsf{L}_{\boldsymbol{\phi}} \in \mathfrak{C}^1_b\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ as in (4.2) and let $(\mu_t)_{t\in[0,1]}$ be an absolutely continuous curve in $\mathcal{P}_2(\mathbb{R}^d)$. Then*

$$F(\mu_1) - F(\mu_0) = \int_0^1 \int_{\mathbb{R}^d} \langle \mathrm{D}F[\mu_t](x), v_t(x)\rangle \mathrm{d}\mu_t(x)\,\mathrm{d}t, \tag{4.9}$$

where $v_t \in L^2(\mathbb{R}^d, \mu_t; \mathbb{R}^d)$ is the Wasserstein velocity field (cf. Theorem 3.1) of $(\mu_t)_{t\in[0,1]}$ at time t and $\mathrm{D}F$ is as in (4.4).

In case the curve $(\mu_t)_{t\in[0,1]}$ admits the parametrization

$$\mu_t := \big(\boldsymbol{x}_t\big)_\sharp \mu, \quad t \in [0,1],$$

for some Borel probability measure μ in a Polish space Ω and some map $\boldsymbol{x} \in \mathrm{C}^1([0,1]; L^2(\Omega, \mu; \mathbb{R}^d))$, then $F \circ \mu \in \mathrm{C}^1([0,1])$ and

$$\frac{\mathrm{d}}{\mathrm{d}t} F(\mu_t) = \int_\Omega \langle \mathrm{D}F(\mu_t, \boldsymbol{x}_t(\omega)), \dot{\boldsymbol{x}}_t(\omega))\rangle \,\mathrm{d}\mu(\omega) \quad \textit{for every } t \in [0,1]. \tag{4.10}$$

Proof. Observe that, since F is Lipschitz continuous and $t \mapsto \mu_t$ is absolutely continuous, the map $t \mapsto F(\mu_t)$ is absolutely continuous and thus it holds

$$F(\mu_1) - F(\mu_0) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} F(\mu_t)\mathrm{d}t.$$

It is then enough to prove that

$$\frac{\mathrm{d}}{\mathrm{d}t} F(\mu_t) = \int_{\mathbb{R}^d} \langle \mathrm{D}F(\mu_t, x), v_t(x)\rangle \,\mathrm{d}\mu_t(x) \quad \text{for a.e. } t \in (0,1). \tag{4.11}$$

We have, for every $t \in A((\mu_t)_{t\in[0,1]}) \subset (0,1)$ (cf. Theorem 3.1), that

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} F(\mu_t) &= \sum_{i=1}^N \partial_i \psi(\mathsf{L}_{\boldsymbol{\phi}}(\mu_t)) \frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{R}^d} \phi_i \,\mathrm{d}\mu_t \\
&= \sum_{i=1}^N \partial_i \psi(\mathsf{L}_{\boldsymbol{\phi}}(\mu_t)) \int_{\mathbb{R}^d} \langle \nabla \phi_i, v_t(x)\rangle \,\mathrm{d}\mu_t(x) \\
&= \int_{\mathbb{R}^d} \langle \mathrm{D}F(\mu_t, x), v_t(x)\rangle \,\mathrm{d}\mu_t(x),
\end{aligned}$$

where we used Theorem 3.1. A completely analogous argument yields (4.10). □

Remark 4.8. Consider the case in which the curve $(\mu_t)_{t\in[0,1]}$ has the simple form

$$\mu_t := (\boldsymbol{i}_{\mathbb{R}^d} + tu)_\sharp \mu, \quad t \in [0,1]$$

for some map $u \in L^2(\mathbb{R}^d, \mu; \mathbb{R}^d)$, where $\boldsymbol{i}_{\mathbb{R}^d}$ denotes the identity map on $\mathbb{R}^d$. Then (4.10) yields

$$\frac{\mathrm{d}}{\mathrm{d}t} F(\mu_t) = \int_{\mathbb{R}^d} \langle \mathrm{D}F(\mu_t, x), u(x) \rangle \,\mathrm{d}\mu_t(x) \quad \text{for every } t \in [0,1],$$

and, in particular, we get

$$\lim_{t\downarrow 0} \frac{F(\mu_t) - F(\mu)}{t} = \int_{\mathbb{R}^d} \langle \mathrm{D}F(\mu, x), u(x)) \rangle \,\mathrm{d}\mu(x). \tag{4.12}$$

Proposition 4.9. *Let $F = \psi \circ \mathsf{L}_{\boldsymbol{\phi}} \in \mathfrak{C}^1_b\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ as in (4.2). Then*

$$\|\mathrm{D}F\,[\mu]\|_\mu = \operatorname{lip} F(\mu) \quad \textit{for every } \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

In particular $\|\mathrm{D}F\,[\mu]\|_\mu$ does not depend on the choice of the representation of F and $\mathrm{D}F$ just depends on F on $\overline{\mathcal{D}}$.

Proof. Let $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and let $(\mu'_n, \mu''_n) \in \mathcal{P}_2(\mathbb{R}^d)^2$ with $\mu'_n \neq \mu''_n$ be such that $(\mu'_n, \mu''_n) \to (\mu, \mu)$ in W_2 and

$$\lim_n \frac{|F(\mu'_n) - F(\mu''_n)|}{W_2(\mu'_n, \mu''_n)} = \operatorname{lip} F(\mu).$$

Let us define, for every $t \in [0,1]$, the map $\mathsf{x}^t : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ as

$$\mathsf{x}^t(x_0, x_1) := (1-t)x_0 + t x_1, \quad (x_0, x_1) \in \mathbb{R}^d \times \mathbb{R}^d.$$

Using (4.10) along $\mu^t_n := \mathsf{x}^t_\sharp \boldsymbol{\mu}_n$ for plans $\boldsymbol{\mu}_n \in \Gamma_o(\mu'_n, \mu''_n)$ (it is easy to check that $(\mu^t_n)_{t\in[0,1]}$ is Lipschitz continuous), we get

$$\begin{aligned} |F(\mu'_n) - F(\mu''_n)| &= \left| \int_0^1 \int_{\mathbb{R}^d\times\mathbb{R}^d} \langle \mathrm{D}F(\mu^t_n, \mathsf{x}^t(x_0,x_1)), x_1 - x_0 \rangle \,\mathrm{d}\boldsymbol{\mu}_n(x_0,x_1)\,\mathrm{d}t \right| \\ &\le \left(\int_0^1 \int_{\mathbb{R}^d\times\mathbb{R}^d} \left|\mathrm{D}F(\mu^t_n, \mathsf{x}^t(x_0,x_1)))\right|^2 \mathrm{d}\boldsymbol{\mu}_n\,\mathrm{d}t \right)^{\frac12} \left(\int_0^1 \int_{\mathbb{R}^d\times\mathbb{R}^d} |x_1 - x_0|^2 \,\mathrm{d}\boldsymbol{\mu}_n\,\mathrm{d}t \right)^{\frac12} \end{aligned}$$

$$= W_2(\mu'_n, \mu''_n) \left(\int_0^1 \int_{\mathbb{R}^d} \left| \mathrm{D}F(\mu_n^t, x) \right|^2 \, \mathrm{d}\mu_n^t(x) \, \mathrm{d}t \right)^{\frac{1}{2}},$$

where we used Theorem 3.1. Dividing both sides by $W_2(\mu'_n, \mu''_n)$, we obtain

$$\frac{|F(\mu'_n) - F(\mu''_n)|}{W_2(\mu'_n, \mu''_n)} \le \left(\int_0^1 \int_{\mathbb{R}^d} \left| \mathrm{D}F(\mu_n^t, x) \right|^2 \, \mathrm{d}\mu_n^t \, \mathrm{d}t \right)^{\frac{1}{2}}.$$

Observe that $\boldsymbol{\mu}_n \to \boldsymbol{\mu} := (\boldsymbol{i}_{\mathbb{R}^d}, \boldsymbol{i}_{\mathbb{R}^d})_\sharp \mu$ narrowly in $\mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ as $n \to +\infty$ so that $\mu_n^t \to \mu$ narrowly in $\mathcal{P}(\mathbb{R}^d)$ as $n \to +\infty$ for every $t \in [0,1]$. We can pass to the limit as $n \to +\infty$ the above inequality using the dominated convergence Theorem and Lemma 4.6 with

$$G(\mu, x) := \left| \sum_{n=1}^N \partial_n \psi \left(\mathsf{L}_{\boldsymbol{\phi}}(\mu) \right) \nabla \phi_n(x) \right|^2, \quad \mu \in \mathcal{P}(\mathbb{R}^d),\, x \in \mathbb{R}^d,$$

which provides a continuous and bounded extension (depending on the particular choice of ψ and $\boldsymbol{\phi}$) of $|\mathrm{D}F|^2$ to $\mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d$. We hence get

$$\operatorname{lip} F(\mu) \le \left(\int_0^1 \int_{\mathbb{R}^d} |\mathrm{D}F(\mu, x)|^2 \, \mathrm{d}\mu(x) \, \mathrm{d}t \right)^{\frac{1}{2}} = \|\mathrm{D}F\,[\mu]\|_\mu.$$

This proves one inequality. In order to prove the opposite one, it is not restrictive to assume $\|\mathrm{D}F\,[\mu]\|_\mu > 0$. Let us now consider the map $T : \operatorname{supp}(\mu) \to \mathbb{R}^d$ defined as

$$T(x) := \mathrm{D}F[\mu](x), \quad x \in \operatorname{supp}(\mu).$$

By definition of $\operatorname{Tan}_\mu(\mathcal{P}_2(\mathbb{R}^d))$, we have that $T \in \operatorname{Tan}_\mu(\mathcal{P}_2(\mathbb{R}^d))$ so that, by [4, Proposition 8.5.6], we have

$$\lim_{\varepsilon \downarrow 0} \frac{W_2(\mu, (\boldsymbol{i}_{\mathbb{R}^d} + \varepsilon T)_\sharp \mu)}{\varepsilon} = \|T\|_{L^2(\mathbb{R}^d, \mu; \mathbb{R}^d)} = \|\mathrm{D}F\,[\mu]\|_\mu.$$

Moreover, if we apply (4.12) to the curve $\mu_\varepsilon := (\boldsymbol{i}_{\mathbb{R}^d} + \varepsilon T)_\sharp \mu$, $\varepsilon \in [0,1]$, we get

$$\lim_{\varepsilon \downarrow 0} \frac{F(\mu_\varepsilon) - F(\mu)}{\varepsilon} = \int_{\mathbb{R}^d} \langle \mathrm{D}F(\mu, x), T(x) \rangle \, \mathrm{d}\mu(x) = \|\mathrm{D}F\,[\mu]\|_\mu^2,$$

thus

$$\operatorname{lip} F(\mu) \ge \lim_{\varepsilon \downarrow 0} \frac{F(\mu_\varepsilon) - F(\mu)}{W_2(\mu_\varepsilon, \mu)} = \|\mathrm{D}F\,[\mu]\|_\mu.$$

This shows the other inequality and concludes the proof. □

4.2. The density result

Recall that for a bounded Lipschitz function $F : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ the pre-Cheeger energy (cf. (2.15)) associated to $\mathfrak{m}$ is defined by

$$\mathsf{pCE}_2(F) = \int_{\mathcal{P}_2(\mathbb{R}^d)} \big(\mathrm{lip}\, F(\mu)\big)^2 \,\mathrm{d}\mathfrak{m}(\mu). \tag{4.13}$$

Thanks to Proposition 4.9, if F is a cylinder function in $\mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$, we have a nice equivalent expression

$$\mathsf{pCE}_2(F) = \int_{\mathcal{P}_2(\mathbb{R}^d)} \|\mathrm{D}F\,[\mu]\|_\mu^2 \,\mathrm{d}\mathfrak{m}(\mu) = \int |\mathrm{D}F(\mu,x)|^2 \,\mathrm{d}\boldsymbol{\mathfrak{m}}(\mu,x), \tag{4.14}$$

which shows that the restriction of pCE_2 to $\mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ is a quadratic form (thus satisfying (2.59)) induced by the bilinear form

$$\mathsf{pCE}_2(F,G) := \int \mathrm{D}F(\mu,x)\cdot \mathrm{D}G(\mu,x)\,\mathrm{d}\boldsymbol{\mathfrak{m}}(\mu,x), \quad F,G \in \mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big) \tag{4.15}$$

and coincides with the typical bilinear forms on cylinder functions used in [50,47,17,18]. It is therefore important to prove that $\mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ is dense in energy and therefore $H^{1,2}(\mathbb{W}_2)$ is a Hilbert space: this is precisely the object of our main result.

Theorem 4.10. *The algebra $\mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ is dense in 2-energy: for every $F \in D^{1,2}(\mathbb{W}_2)$ there exists a sequence $F_n \in \mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$, $n \in \mathbb{N}$, such that*

$$F_n \to F \quad \mathfrak{m}\text{-a.e.}, \quad \mathrm{lip}(F_n) \to |\mathrm{D}F|_\star \text{ in } L^2(\mathcal{P}_2(\mathbb{R}^d),\mathfrak{m}); \tag{4.16}$$

if moreover $F \in L^p(\mathcal{P}_2(\mathbb{R}^d),\mathfrak{m})$, $p \in [1,+\infty)$, then we can find a sequence $F_n \in \mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ as in (4.16) and converging to F in $L^p(\mathcal{P}_2(\mathbb{R}^d),\mathfrak{m})$.

Corollary 4.11. *$H^{1,2}(\mathbb{W}_2)$ is a separable Hilbert space and $\mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ is strongly dense in $H^{1,2}(\mathbb{W}_2)$. If $\big(\mathsf{pCE}_2, \mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)\big)$ is closable (recall Remark 2.19) then its smallest closed extension coincides with $(\mathsf{CE}_2, H^{1,2}(\mathbb{W}_2))$.*

According to the terminology introduced in [24] (see also [7]) we can say that $(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m})$ is infinitesimally Hilbertian for every positive Borel measure $\mathfrak{m}$. We devote the remaining part of this subsection to the proof this result, using Theorem 2.13.

We adopt the notation $\mathscr{A} := \mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$.

We start with a preliminary lemma, which provides a simple gradient estimate for the distance from the Dirac mass centered at 0, i.e. the quadratic moment of a measure.

Lemma 4.12. *Let $\vartheta \in \mathrm{Lip}(\mathbb{R}^d)$ be a L-Lipschitz function which is continuously differentiable in the open set $\Omega_\vartheta := \{x \in \mathbb{R}^d : \vartheta(x) \neq 0\}$. Then the map*

$$F : \mu \to \big(\mathsf{L}_{\vartheta^2}(\mu)\big)^{1/2} = \Big(\int_{\mathbb{R}^d} \vartheta^2(x)\,\mathrm{d}\mu(x)\Big)^{1/2} \tag{4.17}$$

is L-Lipschitz and belongs to $D^{1,2}(\mathbb{W}_2, \mathscr{A})$, in particular its $(2, \mathscr{A})$-relaxed gradient is bounded above by L and satisfies

$$|\mathrm{D}F|^2_{\star,\mathscr{A}}(\mu) \le \frac{1}{F^2(\mu)} \int_{\mathbb{R}^d} \vartheta^2 |\nabla\vartheta|^2\,\mathrm{d}\mu \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d) \text{ with } F(\mu) > 0. \tag{4.18}$$

Proof. Let $T \in \mathrm{C}^\infty(\mathbb{R})$ be an odd, nondecreasing truncation function satisfying

$$T(x) = x \quad \text{if } |x| \le 1/2, \quad |T(x)| = 1 \quad \text{if } |x| \ge 2, \quad |T'(x)| \le 1, \tag{4.19}$$

and let us set $T_n(x) := nT(x/n)$, $\vartheta_n := T_n \circ \vartheta$, so that ϑ_n is L-Lipschitz and continuously differentiable in Ω_ϑ, so that $\vartheta_n^2 \in \mathrm{C}^1_b(\mathbb{R}^d)$.

We define $\psi_n(r) := (r + 1/n)^{1/2}$ and $F_n := \psi_n \circ \mathsf{L}_{\vartheta_n^2}$. By construction $F_n \in \mathscr{A}$ with

$$\mathrm{D}F_n(\mu, x) = \frac{1}{F_n(\mu)} \vartheta_n(x) \nabla\vartheta_n(x),$$
$$\big(\operatorname{lip} F_n(\mu)\big)^2 = \|\mathrm{D}F_n[\mu]\|^2 = \frac{1}{F_n^2(\mu)} \int_{\mathbb{R}^d} \vartheta_n^2(x) |\nabla\vartheta_n(x)|^2\,\mathrm{d}\mu(x) \le L^2. \tag{4.20}$$

Since $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ is a length space we deduce that F_n is L-Lipschitz. On the other hand $\lim_{n\to\infty} F_n(\mu) = F(\mu)$ pointwise everywhere, so that F is L-Lipschitz as well, it belongs to $D^{1,2}(\mathbb{W}_2, \mathscr{A})$ and $|\mathrm{D}F|_{\star,\mathscr{A}} \le L$. Passing eventually to the limit as $n \to \infty$ in (4.20) for μ in the open set $\{\mu \in \mathcal{P}_2(\mathbb{R}^d) : F(\mu) > 0\}$ we get (4.18). □

Selecting $\vartheta(x) := |x|$ and applying the first part of Lemma 4.12 we immediately get the following corollary.

Corollary 4.13. *The function $\mathsf{m}_2(\cdot)$ as in (3.6) belongs to $D^{1,2}(\mathbb{W}_2, \mathscr{A})$ with*

$$|\mathsf{D}\mathsf{m}_2|_{\star,\mathscr{A}}(\mu) \le 1 \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{4.21}$$

We now use m_2 for localizing gradient estimates in $\mathcal{P}_2(\mathbb{R}^d)$.

Lemma 4.14. *Let F_n be a sequence of functions in $D^{1,2}(\mathbb{W}_2, \mathscr{A}) \cap L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ such that F_n and $|\mathrm{D}F_n|_{\star,\mathscr{A}}$ are uniformly bounded in every bounded set of $\mathcal{P}_2(\mathbb{R}^d)$ and let F, G be Borel functions in $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$, G nonnegative. If*

$$\lim_{n\to\infty} F_n(\mu) = F(\mu), \quad \limsup_{n\to\infty} |\mathrm{D}F_n|_{\star,\mathscr{A}}(\mu) \le G(\mu) \quad \mathfrak{m}\text{-}a.e. \text{ in } \mathcal{P}_2(\mathbb{R}^d), \tag{4.22}$$

then $F \in H^{1,2}(\mathbb{W}_2, \mathscr{A})$ and $|\mathrm{D}F|_{\star,\mathscr{A}} \le G$.

Proof. Let us consider a smooth nonincreasing function $\theta \in \mathrm{C}^\infty[0, +\infty)$ such that

$$\theta(r) = 1 \quad \text{if } 0 \le r \le 1, \quad \theta(r) = 0 \quad \text{if } r \ge 2, \quad |\theta'(r)| \le 2 \tag{4.23}$$

and set

$$\chi_n(\mu) := \theta\big(\mathsf{m}_2(\mu)/n\big). \tag{4.24}$$

By Corollary 4.13 we have

$$\chi_n \in H^{1,2}(\mathbb{W}_2, \mathscr{A}), \quad |\mathrm{D}\chi_n|_{\star,\mathscr{A}} \le 2/n, \quad |\mathrm{D}\chi_n|_{\star,\mathscr{A}}(\mu) = 0 \text{ if } \mathsf{m}_2(\mu) \le n \text{ or } \mathsf{m}_2(\mu) \ge 2n. \tag{4.25}$$

Thanks to the Leibniz rule, setting $F_{n,m}(\mu) := F_n(\mu)\chi_m^2(\mu)$ and $G_n := |\mathrm{D}F_n|_{\star,\mathscr{A}}$, we have

$$F_{n,m} \in \mathrm{D}^{1,2}(\mathbb{W}_2, \mathscr{A}), \quad |\mathrm{D}F_{n,m}|_{\star,\mathscr{A}}(\mu) \le G_n(\mu)\chi_m^2(\mu) + 4/mF_n(\mu)\chi_m(\mu). \tag{4.26}$$

Since for every $m \in \mathbb{N}$ the sequence $n \mapsto G_n\chi_m^2$ is uniformly bounded, we can find an increasing subsequence $k \mapsto n(k)$ such that $k \mapsto G_{n(k)}\chi_m^2$ is weakly* convergent in $L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ and we denote by $\tilde{G}_m$ is weak* limit. By Fatou's lemma, for every Borel set $B \subset \mathcal{P}_2(\mathbb{R}^d)$ we get

$$\begin{aligned}
\int_B \tilde{G}_m \,\mathrm{d}\mathfrak{m} &= \lim_{k\to\infty} \int_B G_{n(k)}(\mu)\chi_m^2(\mu)\,\mathrm{d}\mathfrak{m}(\mu) \\
&\le \int_B \limsup_{k\to\infty} \Big(G_{n(k)}(\mu)\chi_m^2(\mu)\Big)\,\mathrm{d}\mathfrak{m}(\mu) \\
&\le \int_B G^2\chi_m^2 \,\mathrm{d}\mathfrak{m}
\end{aligned}$$

so that we deduce

$$\tilde{G}_m \le G^2\chi_m^2 \quad \mathfrak{m}\text{-a.e. in } \mathcal{P}_2(\mathbb{R}^d), \text{ for every } m \in \mathbb{N}. \tag{4.27}$$

On the other hand, passing to the limit in (4.26) along the subsequence $n(k)$ and recalling that $\lim_{k\to\infty} F_{n(k),m} = F\chi_m^2$ $\mathfrak{m}$-a.e. we get

$$|\mathrm{D}(F\chi_m^2)|_{\star,\mathscr{A}}(\mu) \le \tilde{G}_m(\mu) + \frac{4}{m}F(\mu)\chi_m(\mu) \le G(\mu)\chi_m^2(\mu) + \frac{4}{m}F(\mu)\chi_m(\mu) \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{4.28}$$

We eventually pass to the limit as $m \to \infty$ concluding the proof of the Lemma. □

We now derive a natural estimate, extending (4.4) to the case of quadratically coercive functions whose gradient has a linear growth.

Lemma 4.15. *Let $\phi \in \mathrm{C}^1(\mathbb{R}^d)$ be satisfying the growth conditions*

$$\phi(x) \ge A|x|^2 - B, \quad |\nabla\phi(x)| \le C(|x|+1) \quad \text{for every } x \in \mathbb{R}^d \tag{4.29}$$

for given positive constants $A, B, C > 0$ and let $\zeta : \mathbb{R} \to \mathbb{R}$ be a C^1 nondecreasing function whose derivative has compact support. Then the function $F(\mu) := \zeta \circ \mathsf{L}_\phi$ is Lipschitz in $\mathcal{P}_2(\mathbb{R}^d)$, it belongs to $H^{1,2}(\mathbb{W}_2, \mathscr{A})$, and

$$|\mathrm{D}F|_{\star,\mathscr{A}}(\mu) \le \zeta'(\mathsf{L}_\phi(\mu))\Big(\int_{\mathbb{R}^d} |\nabla\phi(x)|^2 \,\mathrm{d}\mu(x)\Big)^{1/2}. \tag{4.30}$$

Proof. We set $\zeta_a(z) := (z+a)^{1/2}$ and $\vartheta_a := \zeta_a \circ \phi$, with $a := A + B$, so that

$$\vartheta_a \in \mathrm{C}^1(\mathbb{R}^d), \quad \vartheta_a \ge \big(A(|x|^2+1)\big)^{1/2}, \quad |\nabla\vartheta_a(x)| = \frac{|\nabla\phi(x)|}{2(\phi(x)+a)^{1/2}} \le L, \quad L := A^{-1/2}C$$

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

We can then apply Lemma 4.12, observing that

$$\big(\mathsf{L}_{\vartheta_a^2}(\mu)\big)^{1/2} = \zeta_a\big(\mathsf{L}_\phi(\mu)\big);$$

we deduce that $F_a = \zeta_a \circ \mathsf{L}_\phi$ is L-Lipschitz, it belongs to $D^{1,2}(\mathbb{W}_2, \mathscr{A})$ and satisfies (recall (4.18))

$$|\mathrm{D}F_a|_{\star,\mathscr{A}}(\mu) \le \frac{1}{2F_a(\mu)}\Big(\int_{\mathbb{R}^d} |\nabla(\vartheta_a^2)|^2 \,\mathrm{d}\mu\Big)^{1/2} = \frac{1}{2F_a(\mu)}\Big(\int_{\mathbb{R}^d} |\nabla\phi|^2 \,\mathrm{d}\mu\Big)^{1/2}. \tag{4.31}$$

We eventually observe that $F = \psi_a \circ F_a$, where $\psi_a(z) = \zeta(z^2 - a)$ is a C^1 Lipschitz function since ζ' has compact support. By the chain rule in Theorem 2.3(7) we get that

$$|\mathrm{D}F|_{\star,\mathscr{A}} = |\psi_a' \circ F_a||\mathrm{D}F_a|_{\star,\mathscr{A}} = 2F_a\zeta'(\mathsf{L}_\phi)|\mathrm{D}F_a|_{\star,\mathscr{A}}.$$

Then (4.31) yields (4.30). □

We collect in the following definition some useful tools and notation we will extensively use.

Definition 4.16. We denote by $\kappa \in \mathrm{C}_c^\infty(\mathbb{R}^d)$ a smooth function satisfying $\operatorname{supp}\kappa = \overline{\mathrm{B}(0,1)}$, $\kappa(x) \geq 0$ for every $x \in \mathbb{R}^d$ and $\kappa(x) > 0$ for every $x \in \mathrm{B}(0,1)$, $\int_{\mathbb{R}^d} \kappa \,\mathrm{d}\mathscr{L}^d = 1$ and $\kappa(-x) = \kappa(x)$ for every $x \in \mathbb{R}^d$.

For every $0 < \varepsilon < 1$ we define the family of associated mollifiers

$$\kappa_\varepsilon(x) := \frac{1}{\varepsilon^d}\kappa(x/\varepsilon) \quad x \in \mathbb{R}^d, \tag{4.32}$$

and for every $\sigma \in \mathcal{P}_2(\mathbb{R}^d)$ we define

$$\sigma_\varepsilon := \sigma * \kappa_\varepsilon, \tag{4.33}$$

$$\hat{\sigma}_\varepsilon := \frac{\sigma_\varepsilon \llcorner \mathrm{B}(0,1/\varepsilon) + \varepsilon^{d+3}\mathscr{L}^d \llcorner \mathrm{B}(0,1/\varepsilon)}{\sigma_\varepsilon(\mathrm{B}(0,1/\varepsilon)) + \varepsilon^{d+3}\mathscr{L}^d(\mathrm{B}(0,1/\varepsilon))}. \tag{4.34}$$

For every $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ we eventually define the continuous functions $W_\nu, W_\nu^\varepsilon, F_\nu^\varepsilon : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ as

$$W_\nu(\mu) := W_2(\mu,\nu), \quad W_\nu^\varepsilon(\mu) := W_{\hat{\nu}_\varepsilon}(\mu_\varepsilon), \quad F_\nu^\varepsilon(\mu) := \frac{1}{2}(W_\nu^\varepsilon(\mu))^2, \quad \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{4.35}$$

Notice that $\sigma_\varepsilon, \hat{\sigma}_\varepsilon \in \mathcal{P}_2^r(\mathbb{R}^d)$, $\operatorname{supp}\hat{\sigma}_\varepsilon = \overline{\mathrm{B}(0,1/\varepsilon)}$ and $W_2(\sigma_\varepsilon,\sigma) \to 0$, $W_2(\hat{\sigma}_\varepsilon,\sigma) \to 0$ as $\varepsilon \downarrow 0$. Moreover, if $\sigma, \sigma' \in \mathcal{P}_2(\mathbb{R}^d)$, we have

$$W_2(\sigma_\varepsilon, \sigma'_\varepsilon) \leq W_2(\sigma,\sigma') \quad \text{for every } 0 < \varepsilon < 1 \tag{4.36}$$

and it is easy to check that, if we set

$$C_\varepsilon := \mathsf{m}_2(\kappa_\varepsilon \mathscr{L}^d), \tag{4.37}$$

then we have

$$\mathsf{m}_2(\mu_\varepsilon) \leq \mathsf{m}_2(\mu) + C_\varepsilon \quad \text{for every } 0 < \varepsilon < 1. \tag{4.38}$$

Proposition 4.17. *Let $\nu \in \mathcal{P}_2(\mathbb{R}^d)$, $\varepsilon \in (0,1)$ and let $\zeta : \mathbb{R} \to \mathbb{R}$ be a C^1 nondecreasing function whose derivative has compact support. With the notation of Definition 4.16 we have*

$$|\mathrm{D}(\zeta \circ F_\nu^\varepsilon)|_{\star,\mathscr{A}}(\mu) \leq \zeta'(F_\nu^\varepsilon(\mu))\Big(\int_{\mathbb{R}^d} |x - \nabla(\varphi_\varepsilon^* * \kappa_\varepsilon)(x)|^2 \,\mathrm{d}\mu(x)\Big)^{1/2} \tag{4.39}$$

for $\mathfrak{m}$-a.e. $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, *where* $\varphi_\varepsilon^* = \Phi^*(\hat{\nu}_\varepsilon, \mu_\varepsilon)$ *as in Theorem 3.2.*

Proof. Let $\mathcal{G} := \{\mu^h\}_{h\in\mathbb{N}}$ be a dense and countable set in $\mathcal{P}_2(\mathbb{R}^d)$ and let us set, for every $h \in \mathbb{N}$, $\varphi_{\varepsilon,h} := \Phi(\hat{\nu}_\varepsilon, \mu_\varepsilon^h)$, ${\varphi_{\varepsilon,h}}^* := \Phi^*(\hat{\nu}_\varepsilon, \mu_\varepsilon^h)$ (see Theorem 3.2),

$$a_{\varepsilon,h} := \int\limits_{\mathrm{B}(0,1/\varepsilon)} \left(\frac{1}{2}|y|^2 - \varphi_{\varepsilon,h}(y)\right) \mathrm{d}\hat{\nu}_\varepsilon(y), \quad u_{\varepsilon,h}(x) := \frac{1}{2}|x|^2 - {\varphi_{\varepsilon,h}}^*(x) + a_{\varepsilon,h}, \quad x \in \mathbb{R}^d$$

and

$$G_{\varepsilon,k}(\mu) := \max_{1\le h\le k} \int\limits_{\mathbb{R}^d} u_{\varepsilon,h} \mathrm{d}\mu_\varepsilon, \quad \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

We first observe that $\varphi^*_{\varepsilon,h}(x)$ is $1/\varepsilon$-Lipschitz (cf. Theorem 3.2), so that $|\varphi^*_{\varepsilon,h}(x)| \le |x|/\varepsilon$ and

$$u_{\varepsilon,h}(x) \ge \frac{1}{2}|x|^2 - \frac{1}{\varepsilon}|x| + a_{\varepsilon,h} \ge \frac{1}{4}|x|^2 - \frac{1}{\varepsilon^2} + a_{\varepsilon,h}, \tag{4.40}$$

$$u_{\varepsilon,h}(x) \le |x|^2 + \frac{1}{\varepsilon^2} + a_{\varepsilon,h}. \tag{4.41}$$

Claim 1. It holds

$$\lim_{k\to+\infty} G_{\varepsilon,k}(\mu) = F_\nu^\varepsilon(\mu) \quad \text{for every } \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

Proof of claim 1. Since $G_{\varepsilon,k+1}(\mu) \ge G_{\varepsilon,k}(\mu)$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, we have that

$$\lim_{k\to+\infty} G_{\varepsilon,k}(\mu) = \sup_k G_{\varepsilon,k}(\mu) = \sup_h \int\limits_{\mathbb{R}^d} u_{\varepsilon,h} \mathrm{d}\mu_\varepsilon \quad \text{for every } \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

By the definition of $\varphi_{\varepsilon,h}$ and ${\varphi_{\varepsilon,h}}^*$ (see (3.9)) we have that

$$\frac{1}{2}|x|^2 - \varphi^*_{\varepsilon,h}(x) + \frac{1}{2}|y|^2 - \varphi_{\varepsilon,h}(y) \le \frac{1}{2}|x-y|^2 \quad \text{for every } x \in \mathbb{R}^d,\ y \in \mathrm{B}(0,1/\varepsilon),$$

so that for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $h \in \mathbb{N}$, we get

$$\begin{aligned}
\int\limits_{\mathbb{R}^d} u_{\varepsilon,h} \mathrm{d}\mu_\varepsilon &= \int\limits_{\mathbb{R}^d} \left(\frac{1}{2}|x|^2 - \varphi^*_{\varepsilon,h}(x)\right) \mathrm{d}\mu_\varepsilon + \int\limits_{\mathrm{B}(0,1/\varepsilon)} \left(\frac{1}{2}|y|^2 - \varphi_{\varepsilon,h}(y)\right) \mathrm{d}\hat{\nu}_\varepsilon(y) \\
&\le \frac{1}{2} \int\limits_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2 \,\mathrm{d}\boldsymbol{\gamma}(x,y) \\
&= \frac{1}{2} W_2^2(\mu_\varepsilon, \hat{\nu}_\varepsilon)
\end{aligned}$$

$$= F_\nu^\varepsilon(\mu),$$

where $\boldsymbol{\gamma} \in \Gamma_o(\mu_\varepsilon, \hat{\nu}_\varepsilon)$. This proves that $\sup_k G_{\varepsilon,k}(\mu) \leq F_\nu^\varepsilon(\mu)$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. If $\mu \in \mathcal{G}$, then we can find $h \in \mathbb{N}$ such that $\mu = \mu^h$ so that, by definition of $\varphi_{\varepsilon,h}$ and $\varphi_{\varepsilon,h}^*$, we obtain that

$$\begin{aligned}\int_{\mathbb{R}^d} u_{\varepsilon,h} \mathrm{d}\mu_\varepsilon &= \int_{\mathbb{R}^d} \left(\frac{1}{2}|x|^2 - \varphi_{\varepsilon,h}^*(x)\right) \mathrm{d}\mu_\varepsilon + \int_{\mathrm{B}(0,1/\varepsilon)} \left(\frac{1}{2}|y|^2 - \varphi_{\varepsilon,h}(y)\right) \mathrm{d}\hat{\nu}_\varepsilon(y) \\ &= \frac{1}{2} W_2^2(\mu_\varepsilon^h, \hat{\nu}_\varepsilon).\end{aligned}$$

Hence, if $\mu \in \mathcal{G}$, then $\sup_k G_{\varepsilon,k}(\mu) = F_\nu^\varepsilon(\mu)$.

Let now $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$ and $h \in \mathbb{N}$ and observe that

$$\begin{aligned}\int_{\mathbb{R}^d} u_{\varepsilon,h} \mathrm{d}\mu_\varepsilon - \int_{\mathbb{R}^d} u_{\varepsilon,h} \mathrm{d}\mu'_\varepsilon &= \frac{1}{2}\mathsf{m}_2^2(\mu_\varepsilon) - \frac{1}{2}\mathsf{m}_2^2(\mu'_\varepsilon) - \int_{\mathbb{R}^d} {\varphi_{\varepsilon,h}}^* \mathrm{d}(\mu_\varepsilon - \mu'_\varepsilon) \\ &\leq \frac{1}{2}\Big(\mathsf{m}_2(\mu_\varepsilon) + \mathsf{m}_2(\mu'_\varepsilon)\Big) W_2(\mu_\varepsilon, \mu'_\varepsilon) + \frac{1}{\varepsilon} W_2(\mu_\varepsilon, \mu'_\varepsilon) \\ &\leq \frac{1}{2}\Big(\mathsf{m}_2(\mu) + \mathsf{m}_2(\mu') + 2C_\varepsilon\Big) W_2(\mu, \mu') + \frac{1}{\varepsilon} W_2(\mu, \mu') \\ &\leq \Big(\mathsf{m}_2(\mu) + \mathsf{m}_2(\mu') + C_\varepsilon + \frac{1}{\varepsilon}\Big) W_2(\mu, \mu'),\end{aligned}$$

where we used (4.36), (4.38), the fact that $\varphi_{\varepsilon,h}^*$ is $1/\varepsilon$-Lipschitz continuous and (3.5). We hence deduce that for every $k \in \mathbb{N}$

$$\begin{aligned}&|G_{\varepsilon,k}(\mu) - G_{\varepsilon,k}(\mu')| \leq \left(\mathsf{m}_2(\mu) + \mathsf{m}_2(\mu') + \frac{1}{\varepsilon} + C_\varepsilon\right) W_2(\mu, \mu') \\ &\qquad \text{for every } \mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d).\end{aligned} \tag{4.42}$$

Choosing $\mu' \in \mathcal{G}$ and passing to the limit as $k \to +\infty$ we get from (4.42)

$$\begin{aligned}&\left|\lim_{k \to +\infty} G_{\varepsilon,k}(\mu) - F_\nu^\varepsilon(\mu')\right| \leq \left(\mathsf{m}_2(\mu) + \mathsf{m}_2(\mu') + C_\varepsilon + \frac{1}{\varepsilon}\right) W_2(\mu, \mu') \\ &\qquad \text{for every } \mu \in \mathcal{P}_2(\mathbb{R}^d),\ \mu' \in \mathcal{G}.\end{aligned}$$

Using the density of $\mathcal{G}$ and the continuity of $\mu' \mapsto F_\nu^\varepsilon(\mu')$ we deduce that

$$\lim_{k \to +\infty} G_{\varepsilon,k}(\mu) = F_\nu^\varepsilon(\mu) \quad \text{ for every } \mu \in \mathcal{P}_2(\mathbb{R}^d)$$

proving the first claim.

Claim 2. If $H_{\varepsilon,k} := \zeta \circ G_{\varepsilon,k}$ and $\tilde{u}_{\varepsilon,h} := u_{\varepsilon,h} * \kappa_\varepsilon \in \mathrm{C}^1(\mathbb{R}^d)$ it holds

$$
\begin{aligned}
|\mathrm{D}H_{\varepsilon,k}|^2_{\star,\mathscr{A}}(\mu) &\le \big(\zeta'(G_{\varepsilon,k}(\mu))\big)^2 \int_{\mathbb{R}^d} |\nabla \tilde u_{\varepsilon,h}|^2 \,\mathrm{d}\mu(x) \\
&= \big(\zeta'(G_{\varepsilon,k}(\mu))\big)^2 \int_{\mathbb{R}^d} |x - \nabla({\varphi_{\varepsilon,h}}^* * \kappa_\varepsilon)(x)|^2 \,\mathrm{d}\mu(x),
\end{aligned}
$$

for $\mathfrak{m}$-a.e. $\mu \in {B_{\varepsilon,h}}^k$, where ${B_{\varepsilon,h}}^k := \{\mu \in \mathcal{P}_2(\mathbb{R}^d) \mid G_{\varepsilon,k}(\mu) = \int_{\mathbb{R}^d} u_{\varepsilon,h} \mathrm{d}\mu_\varepsilon\}$, $h \in \{1,\dots,k\}$.

Proof of claim 2. For every $h \in \mathbb{N}$, (4.40) yields

$$
\tilde u_{\varepsilon,h}(x) \ge \frac{1}{4}|x|^2 - \frac{1}{\varepsilon^2} + a_{\varepsilon,h}, \quad |\nabla \tilde u_{\varepsilon,h}(x)| \le |x| + \frac{1}{\varepsilon}; \tag{4.43}
$$

where we used that

$$
|x|^2 * \kappa_\varepsilon \ge |x * \kappa_\varepsilon|^2 = |x|^2, \quad \nabla u_{\varepsilon,h}(x) = x - \nabla {\varphi_{\varepsilon,h}}^*(x), \quad \nabla u_{\varepsilon,h} * \kappa_\varepsilon = x - \nabla {\varphi_{\varepsilon,h}}^* * \kappa_\varepsilon.
$$

Since the map $\ell_{\varepsilon,h} : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ defined as $\ell_{\varepsilon,h}(\mu) := \int_{\mathbb{R}^d} u_{\varepsilon,h} \mathrm{d}\mu_\varepsilon$ satisfies

$$
\ell_{\varepsilon,h}(\mu) = \int_{\mathbb{R}^d} (u_{\varepsilon,h} * \kappa_\varepsilon) \mathrm{d}\mu = \mathsf{L}_{\tilde u_{\varepsilon,h}}(\mu), \quad \mu \in \mathcal{P}_2(\mathbb{R}^d),
$$

Lemma 4.15 and the above estimates yield

$$
|\mathrm{D}(\zeta \circ \ell_{\varepsilon,h})|_\star(\mu) \le \zeta'(\ell_{\varepsilon,h}(\mu)) \Big(\int_{\mathbb{R}^d} |\nabla \tilde u_{\varepsilon,h}|^2 \,\mathrm{d}\mu \Big)^{1/2} \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d).
$$

Since ζ is nondecreasing, $H_{\varepsilon,k}$ can be written as

$$
H_{\varepsilon,k}(\mu) = \max_{1 \le h \le k} (\zeta \circ \ell_{\varepsilon,h})(\mu), \quad \mu \in \mathcal{P}_2(\mathbb{R}^d),
$$

so that we can apply Theorem 2.3 (8) and conclude the proof of the second claim.

Claim 3. Let $(h_n)_n \subset \mathbb{N}$ be a non-decreasing sequence and let $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. If $\lim_n \int_{\mathbb{R}^d} u_{\varepsilon,h_n} \mathrm{d}\mu_\varepsilon = F^\varepsilon_\nu(\mu)$, then

$$
\lim_n \int_{\mathbb{R}^d} |x - \nabla({\varphi_{\varepsilon,h_n}}^* * \kappa_\varepsilon)(x)|^2 \,\mathrm{d}\mu(x) = \int_{\mathbb{R}^d} |x - \nabla(\varphi^*_\varepsilon * \kappa_\varepsilon)(x)|^2 \,\mathrm{d}\mu(x),
$$

where $\varphi^*_\varepsilon = \Phi^*(\hat\nu_\varepsilon, \mu_\varepsilon)$.

Proof of claim 3. Let us set for every $n \in \mathbb{N}$

$$
\phi_{\varepsilon,n} := \varphi_{\varepsilon,h_n}, \quad \psi_{\varepsilon,n} = \phi^*_{\varepsilon,n} := \varphi^*_{\varepsilon,h_n}.
$$

We will show that from any (non relabeled) increasing subsequence it is possible to extract a further subsequence $j \mapsto n(j)$ such that

$$\lim_j \int_{\mathbb{R}^d} \left| x - \nabla(\phi^*_{\varepsilon,n(j)} * \kappa_\varepsilon)(x) \right|^2 \mathrm{d}\mu(x) = \int_{\mathbb{R}^d} |x - \nabla(\varphi^*_\varepsilon * \kappa_\varepsilon)(x)|^2 \,\mathrm{d}\mu(x).$$

By Theorem 3.2, we have that, for every $n \in \mathbb{N}$, $\phi^*_{\varepsilon,n} : \mathbb{R}^d \to \mathbb{R}$ is convex and $1/\varepsilon$-Lipschitz continuous with $\phi^*_{\varepsilon,n}(0) = 0$, $\phi_{\varepsilon,n} : \mathrm{B}(0, 1/\varepsilon) \to \mathbb{R}$ is convex and continuous and

$$\phi_{\varepsilon,n}(x) = \sup_{y \in \mathbb{R}^d} \langle x, y\rangle - \phi^*_{\varepsilon,n}(y) \quad \text{for every } x \in \mathbb{R}^d.$$

Moreover, since $F^\varepsilon_\nu \geq 0$ and $|\phi^*_{\varepsilon,n}(x)| \leq \varepsilon^{-1}|x|$, for n sufficiently large we have

$$a_{\varepsilon,h_n} = \int_{\mathbb{R}^d} \left(u_{\varepsilon,h_n}(x) + \phi^*_{\varepsilon,n}(x) - \frac{1}{2}|x|^2 \right) \mathrm{d}\mu_\varepsilon(x) \geq -C(\varepsilon, \mu) \tag{4.44}$$

with $C(\varepsilon, \mu) := 1 + \frac{1}{2}\mathsf{m}_2^2(\mu_\varepsilon) + \frac{1}{\varepsilon}\mathsf{m}_2(\mu_\varepsilon)$. It follows that

$$\int_{\mathrm{B}(0,1/\varepsilon)} \phi_{\varepsilon,n} \,\mathrm{d}\hat{\nu}_\varepsilon = \frac{1}{2}\mathsf{m}_2(\hat{\nu}^\varepsilon) - a_{\varepsilon,h_n} \leq C'(\varepsilon, \mu, \nu)$$

where $C'(\varepsilon, \mu, \nu) := C(\varepsilon, \mu) + \frac{1}{2}\mathsf{m}_2^2(\hat{\nu}_\varepsilon)$. Since $\hat{\nu}^\varepsilon \geq \frac{\varepsilon^{d+3}}{1+\varepsilon^3\omega_d} \mathscr{L}^d \llcorner \mathrm{B}(0, 1/\varepsilon)$ we can find a constant $I = I(\varepsilon, \mu, \nu)$ such that $\int_{\mathrm{B}(0,1/\varepsilon)} \phi_{\varepsilon,n} \,\mathrm{d}x \leq I$ for sufficiently large n.

Thus by Lemma 3.4, we get the existence of a subsequence $j \mapsto n(j)$ and two convex continuous functions $\phi^*_\varepsilon : \mathbb{R}^d \to \mathbb{R}$ and $\phi_\varepsilon : \mathrm{B}(0, 1/\varepsilon) \to \mathbb{R}$ such that points (i), (ii), (iii) and conclusions of Lemma 3.4 hold. By points (i) and (ii) we can use Fatou Lemma and the dominated convergence Theorem to conclude that

$$\liminf_j \int_{\mathrm{B}(0,1/\varepsilon)} \phi_{\varepsilon,n(j)} \mathrm{d}\hat{\nu}_\varepsilon \geq \int_{\mathrm{B}(0,1/\varepsilon)} \phi_\varepsilon \mathrm{d}\hat{\nu}_\varepsilon, \quad \lim_j \int_{\mathbb{R}^d} \phi^*_{\varepsilon,n(j)} \mathrm{d}\mu_\varepsilon = \int_{\mathbb{R}^d} \phi^*_\varepsilon \mathrm{d}\mu_\varepsilon.$$

We thus deduce that

$$\begin{aligned}\int_{\mathbb{R}^d} \left(\frac{1}{2}|x|^2 - \phi^*_\varepsilon(x)\right) \mathrm{d}\mu_\varepsilon(x) + \int_{\mathrm{B}(0,1/\varepsilon)} \left(\frac{1}{2}|y|^2 - \phi_\varepsilon(y)\right) \mathrm{d}\hat{\nu}_\varepsilon(y) &\geq \limsup_j \int_{\mathbb{R}^d} u_{\varepsilon,h_{n(j)}} \mathrm{d}\mu_\varepsilon \\ &= F^\varepsilon_\nu(\mu)\end{aligned}$$

proving that

$$\int_{\mathrm{B}(0,1/\varepsilon)} \phi_\varepsilon \,\mathrm{d}\hat{\nu}_\varepsilon + \int_{\mathbb{R}^d} \phi_\varepsilon^* \,\mathrm{d}\mu_\varepsilon = \frac{1}{2}\mathsf{m}_2^2(\hat{\nu}_\varepsilon) + \frac{1}{2}\mathsf{m}_2^2(\mu_\varepsilon) - \frac{1}{2}W_2^2(\hat{\nu}_\varepsilon, \mu_\varepsilon).$$

By the uniqueness part of Theorem 3.2 we deduce that $\phi_\varepsilon = \varphi_\varepsilon = \Phi(\hat{\nu}_\varepsilon, \mu_\varepsilon)$ and $\phi_\varepsilon^* = \varphi_\varepsilon^* = \Phi^*(\hat{\nu}_\varepsilon, \mu_\varepsilon)$. Finally, the a.e. convergence of the gradient of $\phi_{\varepsilon,n}^*$ to the gradient of ϕ_ε^* given by point (iii) in Lemma 3.4 gives that $\nabla(\phi_{\varepsilon,n(j)}^* * \kappa_\varepsilon) \to \nabla(\phi_\varepsilon^* * \kappa_\varepsilon)$ pointwise everywhere. Moreover, since for every $x \in \mathbb{R}^d$ we have $x * \kappa_\varepsilon = x$ and

$$\left|x - \nabla(\varphi_{\varepsilon,n(j)}^* * \kappa_\varepsilon)(x)\right|^2 \le \big(|x| + 1/\varepsilon\big)^2 \in L^1(\mathbb{R}^d, \mu),$$

we can use the dominated convergence Theorem to conclude that

$$\lim_j \int_{\mathbb{R}^d} \left|x - \nabla(\phi_{\varepsilon,n(j)}^* * \kappa_\varepsilon)(x)\right|^2 \mathrm{d}\mu(x) = \int_{\mathbb{R}^d} |x - \nabla(\phi_\varepsilon^* * \kappa_\varepsilon)(x)|^2 \,\mathrm{d}\mu(x).$$

This concludes the proof of the third claim.

Claim 4. It holds

$$\limsup_k |\mathrm{D}H_{\varepsilon,k}|_{\star,\mathscr{A}}(\mu) \le \zeta'(F_\nu^\varepsilon(\mu))\Big(\int_{\mathbb{R}^d} |x - \nabla(\varphi_\varepsilon^* * \kappa_\varepsilon)(x)|^2 \,\mathrm{d}\mu(x)\Big)^{1/2}$$

for $\mathfrak{m}$-a.e. $\mu \in \mathcal{P}_2(\mathbb{R}^d)$,

where $\varphi_\varepsilon^* = \Phi^*(\hat{\nu}_\varepsilon, \mu_\varepsilon)$.

Proof of claim 4. Let $B_\varepsilon \subset \mathcal{P}_2(\mathbb{R}^d)$ be defined as

$$B_\varepsilon := \bigcap_k \bigcup_{h=1}^k A_{\varepsilon,h}^k,$$

where $A_{\varepsilon,h}^k$ is the full $\mathfrak{m}$-measure subset of ${B_{\varepsilon,h}}^k$ where claim 2 holds. Notice that B_ε has full $\mathfrak{m}$-measure. Let $\mu \in B_\varepsilon$ be fixed and let us pick a non-decreasing sequence $k \mapsto h_k$ such that

$$G_{\varepsilon,k}(\mu) = \int_{\mathbb{R}^d} u_{\varepsilon,h_k} \mathrm{d}\mu_\varepsilon.$$

By claim 1 we know that $G_{\varepsilon,k}(\mu) \to F_\nu^\varepsilon(\mu)$ so that we can apply claim 4 and conclude that

$$\begin{aligned} &\zeta'(F_\nu^\varepsilon(\mu)) \int_{\mathbb{R}^d} |x - \nabla(\varphi_\varepsilon^* * \kappa_\varepsilon)(x)|^2 \,\mathrm{d}\mu(x) \\ &= \lim_k \zeta'(G_{\varepsilon,k}(\mu)) \int_{\mathbb{R}^d} \left|x - \nabla(\varphi_{\varepsilon,h_k}^* * \kappa_\varepsilon)(x)\right|^2 \mathrm{d}\mu(x). \end{aligned} \tag{4.45}$$

By claim 2, the right hand side is greater than $\limsup_k |\mathrm{D}H_{\varepsilon,k}|^2_{\star,\mathscr{A}}(\mu)$; this concludes the proof of the fourth claim.

Conclusion. We conclude the proof applying Lemma 4.14 with $(H_{\varepsilon,k})_k$ in the role of $(F_n)_n$, $F := \zeta \circ F^\varepsilon_\nu$ and G given by

$$G(\mu) := \zeta'(F^\varepsilon_\nu(\mu))\Big(\int_{\mathbb{R}^d} |x - \nabla(\varphi^*_\varepsilon * \kappa_\varepsilon)(x)|^2 \,\mathrm{d}\mu(x)\Big)^{1/2}, \quad \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

We check that the hypotheses of Lemma 4.14 are satisfied: by Claim 2 we have that $H_{\varepsilon,k} \in D^{1,2}(\mathbb{W}_2, \mathscr{A})$ and it is also in $L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ since ζ is uniformly bounded. Notice that, by (4.43), for every $R > 0$ it holds that

$$\Big(\int_{\mathbb{R}^d} |\nabla \tilde{u}_{\varepsilon,h}(x)|^2 \,\mathrm{d}\mu(x)\Big)^{1/2} \le R + 1/\varepsilon \quad \text{whenever } \mathsf{m}_2(\mu) \le R. \tag{4.46}$$

This gives, also using Claim 2, that $|\mathrm{D}H_{\varepsilon,k}|_{\star,\mathscr{A}}$ is uniformly bounded on bounded subsets of $\mathcal{P}_2(\mathbb{R}^d)$ (recall that ζ' is uniformly bounded). It is also clear that $H_{\varepsilon,k}$ is uniformly bounded on bounded subsets of $\mathcal{P}_2(\mathbb{R}^d)$ since it is uniformly bounded by the infinity norm of ζ.

The function F, being bounded again by the infinity norm of ζ, belongs to $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$. The same holds for G: using (4.45) and passing to the limit the estimate in (4.46) we see that G is uniformly bounded, having ζ' compact support.

By Claim 1 and Claim 4 we have

$$\lim_{k\to+\infty} H_{\varepsilon,k}(\mu) = F(\mu), \quad \limsup_{k\to+\infty} |\mathrm{D}H_{\varepsilon,k}|_{\star,\mathscr{A}}(\mu) \le G(\mu) \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

By Lemma 4.14 we get (4.39). □

We still keep the notation of Definition 4.16.

Corollary 4.18. *Let $\nu \in \mathcal{P}_2(\mathbb{R}^d)$. Then*

$$|\mathrm{D}W_\nu|_{\star,\mathscr{A}}(\mu) \le 1 \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{4.47}$$

Proof. First of all we prove that for every $0 < \varepsilon < 1$, it holds

$$\int_{\mathbb{R}^d} |x - \nabla(\varphi^*_\varepsilon * \kappa_\varepsilon)(x)|^2 \,\mathrm{d}\mu(x) \le W_2^2(\mu_\varepsilon, \hat{\nu}_\varepsilon) \quad \text{for every } \mu \in \mathcal{P}_2(\mathbb{R}^d), \tag{4.48}$$

where $\varphi^*_\varepsilon = \Phi^*(\hat{\nu}_\varepsilon, \mu_\varepsilon)$ as in Theorem 3.2. Since

$$|x - \nabla(\varphi^*_\varepsilon * \kappa_\varepsilon)(x)|^2 \le |x - \nabla\varphi^*_\varepsilon(x)|^2 * \kappa_\varepsilon(x) \quad \text{for every } x \in \mathbb{R}^d,$$

we get that

$$\begin{aligned}\int_{\mathbb{R}^d} |x - \nabla(\varphi_\varepsilon^* * \kappa_\varepsilon)(x)|^2 \, \mathrm{d}\mu(x) &\leq \int_{\mathbb{R}^d} \Big(|x - \nabla\varphi_\varepsilon^*(x)|^2 * \kappa_\varepsilon(x) \Big) \, \mathrm{d}\mu(x) \\ &= \int_{\mathbb{R}^d} |x - \nabla\varphi_\varepsilon^*(x)|^2 \, \mathrm{d}\mu_\varepsilon(x) \\ &= W_2^2(\mu_\varepsilon, \hat{\nu}_\varepsilon),\end{aligned}$$

for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, where the last equality comes from Theorem 3.2. This proves (4.48). It follows from Proposition 4.17 that, for every nondecreasing function $\zeta \in \mathrm{C}^1(\mathbb{R})$ whose derivative has compact support, it holds

$$|\mathrm{D}\,(\zeta \circ F_\nu^\varepsilon)|_{\star,\mathscr{A}}(\mu) \leq \zeta'(F_\nu^\varepsilon(\mu))\sqrt{2F_\nu^\varepsilon(\mu)} \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{4.49}$$

Let us now consider a sequence of continuous and compactly supported functions $\alpha_n : \mathbb{R} \to \mathbb{R}$ such that

$$0 \leq \alpha_n(s) \uparrow \frac{\chi_{(0,+\infty)}(s)}{1+s^2} \leq 1, \quad \text{for every } s \in \mathbb{R}$$

and let us define $\zeta_n(s) : \mathbb{R} \to \mathbb{R}$ as

$$\zeta_n(s) = \int_0^s \alpha_n(r)\mathrm{d}r, \quad s \in \mathbb{R}.$$

Then, for every $n \in \mathbb{N}$, $\zeta_n : \mathbb{R} \to \mathbb{R}$ is a C^1 nondecreasing function whose derivative has compact support so that we can plug it into (4.49) in place of ζ and we see that

$$|\mathrm{D}\,(\zeta_n \circ F_\nu^\varepsilon)|_{\star,\mathscr{A}}(\mu) \leq \zeta_n'(F_\nu^\varepsilon(\mu))\sqrt{2F_\nu^\varepsilon(\mu)} \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{4.50}$$

Observe that $\zeta_n(s) \to \arctan(s)\chi_{(0,+\infty)}$ and $\zeta_n'(s) \to \frac{\chi_{(0,+\infty)}(s)}{1+s^2}$ for every $s \in \mathbb{R}$ and the r.h.s. of (4.50) is uniformly bounded. Using Theorem 2.3(1)-(3) we can thus pass to the limit as $n \to +\infty$ and we obtain

$$|\mathrm{D}(\vartheta \circ W_\nu^\varepsilon)|_{\star,\mathscr{A}}(\mu) \leq \vartheta'(W_\nu^\varepsilon(\mu)) \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d),$$

where $\vartheta : \mathbb{R} \to \mathbb{R}$ is defined as $\vartheta(s) := \arctan(s^2/2)\chi_{(0,+\infty)}(s)$, $s \in \mathbb{R}$. We can thus apply Lemma 2.11 and conclude that

$$|\mathrm{D}W_\nu^\varepsilon|_{\star,\mathscr{A}} \leq 1 \quad \mathfrak{m}\text{-a.e. and for every } 0 < \varepsilon < 1\ . \tag{4.51}$$

Choosing $\varepsilon = 1/k$, we have $\lim_{k\to+\infty} W_\nu^{1/k}(\mu) = W_\nu(\mu)$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$; using Theorem 2.3 (1)-(3), we obtain (4.47). $\square$

The **proof of Theorem 4.10** then easily follows by Corollary 4.18 and Theorem 2.13.

We conclude this section with a simple but useful density property, which shows the possibility to use smaller algebra of cylinder functions to operate in $H^{1,2}(\mathbb{W}_2)$.

Proposition 4.19. *Let $\mathscr{F}$ be a subset of $\mathrm{C}_b^1(\mathbb{R}^d)$ satisfying the following property: for every $f \in \mathrm{C}_b^1(\mathbb{R}^d)$ there exists a sequence $f_n \in \mathscr{F}$, $n \in \mathbb{N}$, such that*

$$\sup_n \|f_n\|_\infty + \|\nabla f_n\|_\infty < \infty, \qquad \lim_{n\to\infty} \int_{\mathbb{R}^d} |f_n - f| + |\nabla(f_n - f)| \,\mathrm{d}\mu = 0 \tag{4.52}$$

for $\mathfrak{m}$-a.e. $\mu \in \mathcal{P}_2(\mathbb{R}^d)$.

Then the algebra $\mathscr{A} \subset \mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ generated by the set of cylinder functions $\big\{\mathsf{L}_f : f \in \mathscr{F}\big\}$ is dense in $H^{1,2}(\mathbb{W}_2)$ and satisfies the strong approximation property of Theorem 4.10.

In particular the algebra $\mathrm{FC}_b^\infty\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ generated by $\big\{\mathsf{L}_f : f \in \mathrm{C}_c^\infty(\mathbb{R}^d)\big\}$ is strongly dense in $H^{1,2}(\mathbb{W}_2)$ and satisfies the approximation property of Theorem 4.10.

Proof. Thanks to Theorem 4.10 and a simple diagonal argument, it is sufficient to prove that for every cylinder function $F \in \mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ there exists a sequence $F_n \in \mathscr{A}$ such that

$$F_n \to F \text{ in } L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}) \quad \text{and} \quad \mathsf{pCE}_2(F_n - F) \to 0 \quad \text{as } n \to \infty. \tag{4.53}$$

In the case $F = \mathsf{L}_f$ with $f \in \mathrm{C}_b^1(\mathbb{R}^d)$, (4.52) and Lebesgue Dominated Convergence Theorem show that we can find a sequence $f_n \in \mathscr{F}$ such that, setting $F_n := \mathsf{L}_{f_n}$, we have

$$\int_{\mathcal{P}_2(\mathbb{R}^d)} |F_n - F|^2 \,\mathrm{d}\mathfrak{m} = \int_{\mathcal{P}_2(\mathbb{R}^d)} \Big| \int_{\mathbb{R}^d} (f_n(x) - f(x)) \,\mathrm{d}\mu(x) \Big|^2 \mathrm{d}\mathfrak{m}(\mu) \to 0 \quad \text{as } n \to \infty,$$

$$\mathsf{pCE}_2(F_n - F) = \int_{\mathcal{P}_2(\mathbb{R}^d)} \int |\nabla f_n(x) - \nabla f(x)|^2 \,\mathrm{d}\mu(x) \,\mathrm{d}\mathfrak{m}(x) \to 0 \quad \text{as } n \to \infty.$$

Let us now consider a general $F = \psi \circ \mathsf{L}_{\boldsymbol{f}}$ as in (4.2), where $\boldsymbol{f} = (f_1, \cdots, f_N)$ is a vector of functions in $\mathrm{C}_b^1(\mathbb{R}^d)$ and $\psi \in \mathrm{C}_b^1(\mathbb{R}^N)$. If we consider $\tilde{\boldsymbol{f}} := (1, f_1, \ldots, f_N)$ and $\tilde{\psi} \in \mathrm{C}_b^1(\mathbb{R}^{N+1})$ defined as

$$\tilde{\psi}(x_0, x_1, \ldots, x_n) := \psi(0)x_0 - \psi(0) + \psi(x_1, x_2, \ldots, x_N), \quad (x_0, x_1, \ldots, x_N) \in \mathbb{R}^{N+1},$$

we have that $\tilde{\psi}(0) = 0$ and $\tilde{\psi} \circ \mathsf{L}_{\tilde{\boldsymbol{f}}} = F$. For this reason we can always suppose that $f_1 \equiv 1$ and $\psi(0) = 0$. It is also not restrictive to assume that ψ is a polynomial with

$\psi(0) = 0$: in fact, setting $R := \sup_{\mathbb{R}^d,\, 1\le k\le N} \Big(|f_k| + |\nabla f_k|\Big)$, we can find a sequence of polynomials $(P_h)_h$ in $\mathbb{R}^N$ such that

$$P_h(0) = 0, \qquad \sup_{|z|\le R} |P_h(z) - \psi(z)| + |\nabla P_h(z) - \nabla\psi(z)| \to 0 \quad \text{as } h \to \infty. \tag{4.54}$$

It follows that $F_h := P_h \circ \mathsf{L}_{\boldsymbol{f}}$ satisfies

$$\lim_{h\to\infty} \sup_{\mathcal{P}_2(\mathbb{R}^d)} \Big(|F_h(\mu) - F(\mu)| + \|\mathrm{D}F_h[\mu] - \mathrm{D}F[\mu]\|_\mu\Big) = 0. \tag{4.55}$$

Let us consider sequences $(f_{k,n})_{n\in\mathbb{N}}$, $k = 1, \cdots, N$, approximating f_k as in (4.52). In particular, there exists $R > 0$ such that $\sup_{\mathbb{R}^d} \Big(|f_{k,n}| + |\nabla f_{k,n}| + |f_k| + |\nabla f_k|\Big) \le R$ for every $n \in \mathbb{N},\ k \in \{1, \cdots, N\}$. If ψ is a polynomial in $\mathbb{R}^N$ with $\psi(0) = 0$ then the function $F_n := \psi \circ \mathsf{L}_{\boldsymbol{f}_n}$ belongs to $\mathscr{A}$ (cf. Remark 4.3), where $\boldsymbol{f}_n = (f_{1,n}, f_{2,n}, \ldots, f_{N,n})$. Denoting by L the maximum of the Lipschitz constants of ψ and $\partial_k\psi$ in the cube $[-R, R]^N$ with respect to the ∞-norm, it is easy to see that

$$\begin{aligned}
|F_n(\mu) - F(\mu)| &= \Big|\psi(\mathsf{L}_{\boldsymbol{f}_n}(\mu)) - \psi(\mathsf{L}_{\boldsymbol{f}_n}(\mu)))\Big| \le L \sup_k |\mathsf{L}_{f_{k,n}}(\mu) - \mathsf{L}_{f_k}(\mu)| \to 0,\\
\|\mathrm{D}F_n[\mu] - \mathrm{D}F[\mu]\|_\mu &= \Big\| \sum_k \Big(\partial_k\psi(\mathsf{L}_{\boldsymbol{f}_n}(\mu))\nabla f_{k,n} - \partial_k\psi(\mathsf{L}_{\boldsymbol{f}}(\mu))\nabla f_k\Big)\Big\|_\mu\\
&\le \sum_k \Big\|\partial_k\psi(\mathsf{L}_{\boldsymbol{f}_n}(\mu))\nabla f_{k,n} - \partial_k\psi(\mathsf{L}_{\boldsymbol{f}_n}(\mu))\nabla f_k\Big\|_\mu\\
&\quad + \sum_k \Big\|\Big(\partial_k\psi(\mathsf{L}_{\boldsymbol{f}_n}(\mu)) - \partial_k\psi(\mathsf{L}_{\boldsymbol{f}}(\mu))\Big)\nabla f_k\Big\|_\mu\\
&\le L \sum_k \Big(\Big\|\nabla f_{k,n} - \nabla f_k\Big\|_\mu + R\Big|\langle f_{k,n} - f_k, \mu\rangle\Big|\Big).
\end{aligned}$$

Both terms are uniformly bounded w.r.t. μ and n, and converge to 0 as $n \to \infty$. We deduce that (4.53) holds. $\square$

Remark 4.20 *(Polynomials).* If there exists a radius $R > 0$ such that $\mathrm{supp}(\mu) \subset \overline{\mathrm{B}(0, R)}$ for $\mathfrak{m}$-a.e. μ then we can also choose subsets $\mathscr{F}$ of $\mathrm{C}^1(\mathbb{R}^d)$ in Proposition 4.19. An interesting example is provided by the collection $\mathscr{F}$ of all the polynomials. In this case the algebra $\mathscr{A}$ is the set of functionals

$$\mu \mapsto \int_{(\mathbb{R}^d)^k} P(x_1, \cdots, x_k)\, \mathrm{d}\mu^{\otimes k}(x_1, \cdots, x_k), \quad P \text{ polynomial in } (\mathbb{R}^d)^k, \quad k \in \mathbb{N}.$$

5. Calculus rules

We now show that the Cheeger energy can be expressed in terms of an appropriate notion of (relaxed) Wasserstein gradient, also depending on $\mathfrak{m}$, which enjoys useful calculus rules.

Theorem 5.1 *($\mathfrak{m}$-Wasserstein differential). For every $F \in D^{1,2}(\mathbb{W}_2)$ there exists a unique vector field $\mathrm{D}_{\mathfrak{m}}F \in L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathbf{m}; \mathbb{R}^d)$ (the $\mathfrak{m}$-Wasserstein differential of F) such that for every sequence $F_n \in \mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$, $n \in \mathbb{N}$, satisfying (4.16) we have*

$$\mathrm{D}F_n \to \mathrm{D}_{\mathfrak{m}}F \quad \text{strongly in } L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathbf{m}; \mathbb{R}^d). \tag{5.1}$$

Moreover:

(a) *The map $F \mapsto \mathrm{D}_{\mathfrak{m}}F$ from $D^{1,2}(\mathbb{W}_2)$ to $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathbf{m}; \mathbb{R}^d)$ is linear and for every $F, G \in D^{1,2}(\mathbb{W}_2)$ we have*

$$\begin{aligned} \mathsf{CE}_2(F,G) &= \int \mathrm{D}_{\mathfrak{m}}F(\mu,x) \cdot \mathrm{D}_{\mathfrak{m}}G(\mu,x) \,\mathrm{d}\mathbf{m}(\mu,x), \\ \mathsf{CE}_2(F) &= \int |\mathrm{D}_{\mathfrak{m}}F(\mu,x)|^2 \,\mathrm{d}\mathbf{m}(\mu,x), \end{aligned} \tag{5.2}$$

where $\mathsf{CE}_2(\cdot,\cdot)$ denotes the quadratic form associated to $\mathsf{CE}_2(\cdot)$ as in Remark 2.18.

(b) *The map $F \mapsto (F, \mathrm{D}_{\mathfrak{m}}F)$ is a linear isometric (thus continuous) immersion of $H^{1,2}(\mathbb{W}_2)$ into $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}) \times L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathbf{m}; \mathbb{R}^d)$.*

(c) *The graph of $\mathrm{D}_{\mathfrak{m}}$ in $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}) \times L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathbf{m}; \mathbb{R}^d)$ is (weakly) closed: for every sequence $F_n \in H^{1,2}(\mathbb{W}_2)$*

$$\left.\begin{aligned} F_n \rightharpoonup F &\text{ in } L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}) \\ \mathrm{D}_{\mathfrak{m}}F_n \rightharpoonup \boldsymbol{G} &\text{ in } L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathbf{m}; \mathbb{R}^d) \end{aligned}\right\} \quad \Rightarrow \quad F \in H^{1,2}(\mathbb{W}_2),\ \boldsymbol{G} = \mathrm{D}_{\mathfrak{m}}F. \tag{5.3}$$

Proof. The proof uses well known arguments of the theory of quadratic forms. If F_n, $n \in \mathbb{N}$, is a sequence in $\mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$, then for every $m, n \in \mathbb{N}$ we have

$$\frac{1}{4}\mathsf{pCE}_2(F_m - F_n) = \frac{1}{2}\Big(\mathsf{pCE}_2(F_m) + \mathsf{pCE}(F_n)\Big) - \mathsf{pCE}_2\Big(\frac{1}{2}(F_m + F_n)\Big). \tag{5.4}$$

If (4.16) holds, we can pass to the limit as $m, n \to \infty$, observing that $\lim_{m,n\to\infty} \frac{1}{2}(F_m + F_n) = F$, and therefore by (2.16) $\liminf_{m,n\to\infty} \mathsf{pCE}_2\Big(\frac{1}{2}(F_m + F_n)\Big) \geq \mathsf{CE}_2(F)$; we thus obtain

$$\limsup_{m,n\to\infty} \frac{1}{4}\mathsf{pCE}_2(F_m - F_n) = \limsup_{m,n\to\infty} \frac{1}{4}\int |\mathrm{D}F_m(\mu,x) - \mathrm{D}F_n(\mu,x)|^2 \,\mathrm{d}\mathbf{m}(\mu,x) \leq 0 \tag{5.5}$$

which shows that $m \mapsto \mathrm{D}F_m$ is a Cauchy sequence in $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathfrak{m}; \mathbb{R}^d)$ and therefore converges to some element $\boldsymbol{V}$.

If $\tilde{F}_n$ is another sequence satisfying (4.16), we can use the identity

$$\frac{1}{4}\mathsf{pCE}_2(F_n - \tilde{F}_n) = \frac{1}{2}\Big(\mathsf{pCE}_2(F_n) + \mathsf{pCE}(\tilde{F}_n)\Big) - \mathsf{pCE}_2\Big(\frac{1}{2}(F_n + \tilde{F}_n)\Big) \tag{5.6}$$

and the same argument to conclude that $\lim_{n\to\infty} \mathsf{pCE}_2(F_n - \tilde{F}_n) = 0$, so that the limit $\boldsymbol{V}$ is independent of the approximating sequence and we are authorized to call it $\mathrm{D}_{\mathfrak{m}}F$.

Concerning claim (a), the linearity of $\mathrm{D}_{\mathfrak{m}}$ follows immediately from the linearity of D as a map from $\mathfrak{C}^1_b\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ to $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathfrak{m}; \mathbb{R}^d)$.

If $F, G \in D^{1,2}(\mathbb{W}_2)$ and $(F_n)_n, (G_n)_n \subset \mathfrak{C}^1_b\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ are sequences satisfying (4.16) for F and G respectively, we can see that $\mathsf{pCE}_2(F_n, G_n) \to \mathsf{CE}_2(F, G)$; indeed

$$\begin{aligned}\mathsf{pCE}_2(F_n, G_n) &= \frac{1}{2}\mathsf{pCE}_2(F_n + G_n) - \frac{1}{2}\mathsf{pCE}_2(F_n) - \frac{1}{2}\mathsf{pCE}_2(G_n),\\ &= -\frac{1}{2}\mathsf{pCE}_2(F_n - G_n) + \frac{1}{2}\mathsf{pCE}_2(F_n) + \frac{1}{2}\mathsf{pCE}_2(G_n).\end{aligned}$$

Passing the first equality to the $\liminf_n$, the second one to the $\limsup_n$ and using (2.16), we get that $\mathsf{pCE}_2(F_n, G_n) \to \mathsf{CE}_2(F, G)$. Passing then to the limit in (4.15) we immediately see that

$$\mathsf{CE}_2(F, G) = \int \mathrm{D}_{\mathfrak{m}}F(\mu, x) \cdot \mathrm{D}_{\mathfrak{m}}G(\mu, x)\,\mathrm{d}\mathfrak{m}(\mu, x) \tag{5.7}$$

which, together with (2.61), shows that $F \mapsto (F, \mathrm{D}_{\mathfrak{m}}F)$ is an isometry from $H^{1,2}(\mathbb{W}_2)$ into $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}) \times L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathfrak{m}; \mathbb{R}^d)$ (claim (b)).

Claim (c) then follows by claim (b) and the fact that $H^{1,2}(\mathbb{W}_2)$ is a Hilbert space. □

Let us now collect a few properties of $\mathrm{D}_{\mathfrak{m}}F$, which follow by the corresponding metric versions of Theorem 2.3 and the approximation property of Theorem 5.1.

Proposition 5.2 *(Calculus properties of $\mathrm{D}_{\mathfrak{m}}F$). The $\mathfrak{m}$-Wasserstein differential satisfies the following properties:*

(a) *(Minimal relaxed gradient and pointwise Lipschitz constant) For every $F \in D^{1,2}(\mathbb{W}_2)$ we have*

$$\|\mathrm{D}_{\mathfrak{m}}F[\mu]\|^2_\mu = \int |\mathrm{D}_{\mathfrak{m}}F(\mu, x)|^2\,\mathrm{d}\mu(x) = |\mathrm{D}F|^2_\star(\mu) \quad \textit{for } \mathfrak{m}\textit{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{5.8}$$

In particular we have the pointwise Rademacher property*: for every $F \in \mathrm{Lip}_b(\mathcal{P}_2(\mathbb{R}^d))$*

$$\|\mathrm{D}_{\mathfrak{m}}F[\mu]\|_\mu^2 = \int |\mathrm{D}_{\mathfrak{m}}F(\mu,x)|^2\,\mathrm{d}\mu(x) \le \big|\mathrm{D}F\big|^2(\mu) \quad \textit{for } \mathfrak{m}\textit{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d), \tag{5.9}$$

and if $F \in \mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$

$$\int |\mathrm{D}_{\mathfrak{m}}F(\mu,x)|^2\,\mathrm{d}\mu(x) \le \int |\mathrm{D}F(\mu,x)|^2\,\mathrm{d}\mu(x) \quad \textit{for } \mathfrak{m}\textit{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{5.10}$$

(b) *(Leibniz rule) If* $F, G \in L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}) \cap D^{1,2}(\mathbb{W}_2)$*, then* $H := FG \in D^{1,2}(\mathbb{W}_2)$ *and*

$$\mathrm{D}_{\mathfrak{m}}H(\mu,x) = F(\mu)\mathrm{D}_{\mathfrak{m}}G(\mu,x) + G(\mu)\mathrm{D}_{\mathfrak{m}}F(\mu,x) \quad \textit{for } \boldsymbol{\mathfrak{m}}\textit{-a.e. } (\mu,x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d. \tag{5.11}$$

(c) *(Locality) If* $F \in D^{1,2}(\mathbb{W}_2)$ *then for any* $\mathscr{L}^1$*-negligible Borel subset* $N \subset \mathbb{R}$ *we have*

$$\mathrm{D}_{\mathfrak{m}}F[\mu] = 0 \quad \textit{in } L^2(\mathbb{R}^d, \mu; \mathbb{R}^d) \quad \mathfrak{m}\textit{-a.e. on } F^{-1}(N). \tag{5.12}$$

(d) *(Truncations) If* $F_j \in D^{1,2}(\mathbb{W}_2)$*,* $j = 1, \cdots, J$*, then also the functions*

$$F_+ := \max(F_1, \cdots, F_J) \textit{ and } F_- := \min(F_1, \cdots, F_J)$$

belong to $D^{1,2}(\mathbb{W}_2)$ *and*

$$\mathrm{D}_{\mathfrak{m}}F_+ = \mathrm{D}_{\mathfrak{m}}F_j \quad \boldsymbol{\mathfrak{m}}\textit{-a.e. on } \{(\mu,x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d : F_+(\mu) = F_j(\mu)\}, \tag{5.13}$$
$$\mathrm{D}_{\mathfrak{m}}F_- = \mathrm{D}_{\mathfrak{m}}F_j \quad \boldsymbol{\mathfrak{m}}\textit{-a.e. on } \{(\mu,x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d : F_-(\mu) = F_j(\mu)\}. \tag{5.14}$$

(e) *(Chain rule) If* $F \in D^{1,2}(\mathbb{W}_2)$ *and* $\phi \in \mathrm{Lip}(\mathbb{R})$ *then* $\phi \circ F \in D^{1,2}(\mathbb{W}_2)$ *and*

$$\mathrm{D}_{\mathfrak{m}}(\phi \circ F) = \phi'(F)\,\mathrm{D}_{\mathfrak{m}}F \quad \boldsymbol{\mathfrak{m}}\textit{-a.e. on } \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d. \tag{5.15}$$

Remark 5.3. Notice that the product in (5.15) is well defined since there exists a $\mathscr{L}^1$-negligible Borel set $N \subset \mathbb{R}$ such that ϕ is differentiable in $\mathbb{R} \setminus N$ and $\mathrm{D}_{\mathfrak{m}}F$ vanishes $\mathfrak{m}$-a.e. in $F^{-1}(N)$ thanks to the locality property (5.12).

Proof. Claim (a) is an immediate consequence of the fact that (4.16) yields $\mathrm{lip}\,F_n \to |\mathrm{D}F|_\star$ strongly in $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$; up to extracting a suitable (not relabeled) subsequence we get $\int |\mathrm{D}F_n[\mu]|^2\,\mathrm{d}\mu \to |\mathrm{D}F|_\star^2(\mu)$ for $\mathfrak{m}$-a.e. μ. On the other hand since by (5.1) $\mathrm{D}F_n \to \mathrm{D}_{\mathfrak{m}}F$ in $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$, then $|\mathrm{D}F_n|^2 \to |\mathrm{D}_{\mathfrak{m}}F|^2$ in $L^1(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}})$; indeed

$$\begin{aligned}\int \Big||\mathrm{D}F_n(\mu,x)|^2 - |\mathrm{D}_{\mathfrak{m}}F(\mu,x)|^2\Big|\,\mathrm{d}\boldsymbol{\mathfrak{m}}(\mu,x) &\le \int \big((|\mathrm{D}F_n| + |\mathrm{D}_{\mathfrak{m}}F|)|\mathrm{D}F_n - \mathrm{D}_{\mathfrak{m}}F|\big)\,\mathrm{d}\boldsymbol{\mathfrak{m}} \\ &\le \left(\int (|\mathrm{D}F_n| + |\mathrm{D}_{\mathfrak{m}}F|)^2\,\mathrm{d}\boldsymbol{\mathfrak{m}}\right)^{1/2} \\ &\quad \cdot \|\mathrm{D}F_n - \mathrm{D}_{\mathfrak{m}}F\|_{L^2(\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)}\end{aligned}$$

so that

$$\int \Big| |\mathrm{D}F_n(\mu,x)|^2 - |\mathrm{D}_{\mathfrak{m}}F(\mu,x)|^2 \Big| \,\mathrm{d}\mathfrak{m}(\mu,x) \to 0 \quad \text{as } n \to +\infty. \tag{5.16}$$

Hence Fubini's Theorem yields, up to extracting a suitable subsequence,

$$\int |\mathrm{D}F_n(\mu,x)|^2 \,\mathrm{d}\mu \to \int |\mathrm{D}_{\mathfrak{m}}F(\mu,x)|^2 \,\mathrm{d}\mu \quad \text{for } \mathfrak{m}\text{-a.e.}\mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{5.17}$$

(5.9) and (5.10) then follows by the general properties of the minimal relaxed gradients (recall Remark 2.8).

Claim (c) follows by (2.8) and (5.8).

Claim (d) is just a consequence of the locality property (5.12).

Claim (e) is true if $\phi \in \mathrm{C}^1_b(\mathbb{R})$ just by passing to the limit in the corresponding formula for a cylinder function. In fact if $F_n \in \mathfrak{C}^1_b\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ is a sequence as in (4.16) and (5.1) we have

$$\mathrm{D}(\phi \circ F_n) = (\phi' \circ F_n)\mathrm{D}F_n \quad \text{in } \mathcal{D}. \tag{5.18}$$

Since ϕ' is bounded and continuous we get

$$\mathrm{D}(\phi\circ F_n) \to \boldsymbol{G} = (\phi'\circ F)\mathrm{D}_{\mathfrak{m}}F \quad \text{strongly in } L^2(\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d) \quad \text{as } n \to \infty. \tag{5.19}$$

Integrating w.r.t. $\boldsymbol{\mathfrak{m}}$ and recalling (5.8) and Theorem 2.3(7) we get

$$\int |\boldsymbol{G}|^2 \,\mathrm{d}\boldsymbol{\mathfrak{m}} = \int |\phi'(F(\mu))|^2 |\mathrm{D}_{\mathfrak{m}}F(\mu,x)|^2 \,\mathrm{d}\boldsymbol{\mathfrak{m}}(\mu,x) = \int |\phi'(F(\mu))|\,|\mathrm{D}F|^2_\star \,\mathrm{d}\mathfrak{m} = \mathsf{CE}_2(\phi\circ F)$$

so that

$$\lim_{n\to\infty} \mathsf{pCE}_2(\phi \circ F_n) = \mathsf{CE}_2(\phi \circ F).$$

We conclude by Theorem 5.1 that $\boldsymbol{G} = (\phi' \circ F)\mathrm{D}_m F$ coincides with $\mathrm{D}_{\mathfrak{m}}(\phi \circ F)$.

Let us now consider the case of a general Lipschitz function ϕ; by truncation and Claim (d) it is not restrictive to assume that ϕ is also bounded. We can find a sequence $\phi_n \in \mathrm{C}^1_b(\mathbb{R})$ such that $\sup_{\mathbb{R}} |\phi_n| + |\phi_n'| \le L < \infty$, $\phi_n \to \phi$ uniformly, and $\phi_n'(x) \to \phi'(x)$ for every $x \in \mathbb{R} \setminus N$ for a Borel set N with $\mathscr{L}^1(N) = 0$. We have

$$\mathrm{D}_{\mathfrak{m}}(\phi_n \circ F) = \phi_n'(F)\mathrm{D}_{\mathfrak{m}}F \quad \mathfrak{m}\text{-a.e. in } \mathcal{P}_2(\mathbb{R}^d). \tag{5.20}$$

Setting $\tilde{N} := \{(\mu,x) \in \overline{\mathcal{D}} : F(\mu) \in N\}$, Fubini's Theorem and the locality property (5.12) yield $\mathrm{D}_{\mathfrak{m}}F(\mu,x) = 0$ for $\boldsymbol{\mathfrak{m}}$-a.e. $(\mu,x) \in \tilde{N}$. On the other hand $\phi_n'(F(\mu)) \to \phi'(F(\mu))$ for every $(\mu,x) \in \mathcal{D} \setminus \tilde{N}$; since ϕ_n' is uniformly bounded, we deduce that

$$\phi_n'(F)\mathrm{D}_{\mathfrak{m}}F \to \phi'(F)\mathrm{D}_{\mathfrak{m}}F \quad \text{strongly in } L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d). \tag{5.21}$$

We conclude by Theorem 5.1(b) that $\mathrm{D}_{\mathfrak{m}}(\phi \circ F) = \phi'(F)\mathrm{D}_{\mathfrak{m}}F$.

Claim (b) follows by claim (e); indeed, since $F, G \in L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$, we can find a constant $M > 0$ such that

$$|F|(\mu) \le M, \quad |G|(\mu) \le M, \quad |F+G|(\mu) \le M \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

Let $\phi \in \mathrm{Lip}(\mathbb{R})$ be such that $\phi(x) = x^2$ for every $x \in [-M-1, M+1]$; then we have

$$\begin{aligned}
\mathrm{D}_{\mathfrak{m}}FG &= \frac{1}{2}\mathrm{D}_{\mathfrak{m}}((F+G)^2) - \frac{1}{2}\mathrm{D}_{\mathfrak{m}}(F^2) - \frac{1}{2}\mathrm{D}_{\mathfrak{m}}(G^2) \\
&= \frac{1}{2}\mathrm{D}_{\mathfrak{m}}(\phi \circ (F+G)) - \frac{1}{2}\mathrm{D}_{\mathfrak{m}}(\phi \circ F) - \frac{1}{2}\mathrm{D}_{\mathfrak{m}}(\phi \circ G) \\
&= \frac{1}{2}\phi'(F+G)\mathrm{D}_{\mathfrak{m}}(F+G) - \frac{1}{2}\phi'(F)\mathrm{D}_{\mathfrak{m}}F - \frac{1}{2}\phi'(G)\mathrm{D}_{\mathfrak{m}}G \\
&= (F+G)\mathrm{D}_{\mathfrak{m}}(F+G) - F\mathrm{D}_{\mathfrak{m}}F - G\mathrm{D}_{\mathfrak{m}}G \\
&= F\mathrm{D}_{\mathfrak{m}}G + G\mathrm{D}_{\mathfrak{m}}F
\end{aligned}$$

for $\boldsymbol{\mathfrak{m}}$-a.e. $(\mu, x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$. $\square$

Corollary 5.4. CE_2 *is a local Dirichlet form in* $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ *[13, 3.1.1] enjoying* Γ-*calculus with Carré du champs* Γ *given by*

$$\Gamma(F,G)[\mu] := \int \mathrm{D}_{\mathfrak{m}}F(\mu,x) \cdot \mathrm{D}_{\mathfrak{m}}G(\mu,x)\,\mathrm{d}\mu(x) \quad \textit{for } \mathfrak{m}\textit{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{5.22}$$

In particular, for every $F, G \in H^{1,2}(\mathbb{W}_2)$ *we have*

$$\begin{aligned}
\mathsf{CE}_2(F,G) &= \int_{\mathcal{P}_2(\mathbb{R}^d)} \Gamma(F,G)[\mu]\,\mathrm{d}\mathfrak{m}(\mu) = \int \mathrm{D}_{\mathfrak{m}}F(\mu,x) \cdot \mathrm{D}_{\mathfrak{m}}G(\mu,x)\,\mathrm{d}\boldsymbol{\mathfrak{m}}(\mu,x), \\
\mathsf{CE}_2(F) &= \int_{\mathcal{P}_2(\mathbb{R}^d)} \Gamma(F,F)\,\mathrm{d}\mathfrak{m}(\mu) = \int |\mathrm{D}_{\mathfrak{m}}F(\mu,x)|^2\,\mathrm{d}\boldsymbol{\mathfrak{m}}(\mu,x).
\end{aligned} \tag{5.23}$$

Proof. The fact that CE_2 is a Dirichlet form follows by the truncation property (5.15) with $\phi(r) := r \wedge 1$. Since $\mathsf{CE}_2(1) = 0$, the same property with $\phi(r) = |r|$ also shows that CE_2 is local (see [13, Corollary 5.1.4]).

Using the Leibniz rule (5.11) one can also easily show that the Γ-operator (5.22) is the Carré du champ associated to $\frac{1}{2}\mathsf{CE}_2$ [13, Definition 4.1.2]. $\square$

5.1. Tangent bundle, residual differentials and relaxation

In general we cannot guarantee that $\mathsf{CE}_2(F)$ coincides with $\mathsf{pCE}_2(F)$ if $F \in \mathfrak{C}_b^1(\mathcal{P}_2(\mathbb{R}^d))$, or, equivalently, that $\mathrm{D}_\mathfrak{m}F = \mathrm{D}F$: this property corresponds to the closability of pCE_2. We can however investigate the relations between $\mathrm{D}F$ and $\mathrm{D}_\mathfrak{m}F$: two useful tools are represented by the closure of the graph of D and by the collection of all the weak limits of Wasserstein differentials along vanishing sequences.

Definition 5.5 *(Multivalued gradient).* We denote by $\mathbf{G} \subset L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}) \times L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$ the closure of the linear space $\big\{(F, \mathrm{D}F) : F \in \mathfrak{C}_b^1(\mathcal{P}_2(\mathbb{R}^d))\big\}$. The multivalued gradient $\mathbf{D}_\mathfrak{m} : H^{1,2}(\mathbb{W}_2) \rightrightarrows L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$ is the operator whose graph is $\mathbf{G}$.

It is clear that $\mathbf{G}$ is a closed vector subspace of $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}) \times L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$, which can also be obtained as the weak closure of $\big\{(F, \mathrm{D}F) : F \in \mathfrak{C}_b^1(\mathcal{P}_2(\mathbb{R}^d))\big\}$. Thus $\boldsymbol{V} \in \mathbf{D}_\mathfrak{m}F$ if and only if there exists a sequence $F_n \in \mathfrak{C}_b^1(\mathcal{P}_2(\mathbb{R}^d))$ such that

$$F_n \to F \text{ in } L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}), \quad \mathrm{D}F_n \rightharpoonup \boldsymbol{V} \text{ in } L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d). \tag{5.24}$$

Definition 5.6 *(Residual gradients).* The set of residual gradients $\mathrm{G}_0 \subset L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$ is defined as

$$\begin{aligned}\mathrm{G}_0 := \Big\{ \boldsymbol{V} \in L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d) : \text{ there exists } (F_n)_{n\in\mathbb{N}} \subset \mathfrak{C}_b^1(\mathcal{P}_2(\mathbb{R}^d)) : \\ F_n \to 0 \text{ in } L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}),\ \mathrm{D}F_n \rightharpoonup \boldsymbol{V} \text{ in } L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d) \Big\}.\end{aligned} \tag{5.25}$$

The notion of residual gradient is known in the literature, see e.g. [30, Section 1.2]. Notice that pCE_2 is closable if and only if G_0 is trivial and that G_0 contains all the vector fields that are limits of gradients of vanishing sequence of functions (see also Lemma 5.9(1)). A third important space is the L^2 tangent bundle of $\mathcal{P}_2(\mathbb{R}^d)$. In the following, given a Borel map $\boldsymbol{G} \in \mathcal{L}^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$, we denote, for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, by $\boldsymbol{G}[\mu]$ the map $x \mapsto \boldsymbol{G}(\mu, x)$.

Definition 5.7. We denote by $\mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \boldsymbol{\mathfrak{m}})$ the subspace of $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$ of vector fields $\boldsymbol{V}$ satisfying

$$\boldsymbol{V}[\mu] \in \mathrm{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d) \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{5.26}$$

Lemma 5.8. $\mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \boldsymbol{\mathfrak{m}})$ *is a closed subspace of* $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$ *which is a* $L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ *module:*

$$\textit{for every } \boldsymbol{V} \in \mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \boldsymbol{\mathfrak{m}}),\ H \in L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}) : \quad H\boldsymbol{V} \in \mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \boldsymbol{\mathfrak{m}}). \tag{5.27}$$

For every $F \in H^{1,2}(\mathbb{W}_2)$ *(resp.* $F \in \mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$*)* $\mathrm{D}_{\mathfrak{m}} F \in \mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ *(resp.* $\mathrm{D}F \in \mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$*). Finally, if* $\mathscr{C} \subset \mathrm{C}_c^\infty(\mathbb{R}^d)$ *is a countable set dense in* $\mathrm{C}_c^\infty(\mathbb{R}^d)$ *with respect to the Lipschitz norm* $\|\zeta\|_{\mathrm{Lip}} := \sup_{\mathbb{R}^d} |\zeta| + |\nabla\zeta|$ *and* $\mathscr{L}$ *is a countable set dense in* $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ *then the set*

$$\mathscr{T} = \mathrm{span}\Big\{H\nabla\zeta : H \in \mathscr{L},\ \zeta \in \mathscr{C}\Big\} \quad \textit{is dense in } \mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}). \tag{5.28}$$

Proof. Let $(\boldsymbol{V}_n)_{n\in\mathbb{N}}$ be a sequence in $\mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ strongly converging to $\boldsymbol{V}$ in L^2; it is not restrictive to assume that $\boldsymbol{V}_n$ are Borel maps satisfying $\boldsymbol{V}_n[\mu] \in \mathrm{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}$ for a $\mathfrak{m}$-negligible set of $\mathcal{P}_2(\mathbb{R}^d)$. Up to extracting a suitable subsequence, we can also assume that $\sum_{n=1}^\infty \|\boldsymbol{V}_n - \boldsymbol{V}\|_{L^2}^2 < \infty$. Applying Fubini's Theorem it follows that

$$\int_{\mathcal{P}_2(\mathbb{R}^d)} \Big(\sum_{n=1}^\infty \int_{\mathbb{R}^d} |\boldsymbol{V}_n[\mu](x) - \boldsymbol{V}[\mu](x)|^2 \,\mathrm{d}\mu(x)\Big) \,\mathrm{d}\mathfrak{m} < +\infty$$

so that there exists a $\mathfrak{m}$-negligible set $\mathcal{N}' \supset \mathcal{N}$ such that

$$\sum_{n=1}^\infty \int_{\mathbb{R}^d} |\boldsymbol{V}_n[\mu](x) - \boldsymbol{V}[\mu](x)|^2 \,\mathrm{d}\mu(x) < \infty \quad \text{for every } \mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}';$$

and this implies that $\boldsymbol{V}_n[\mu] \to \boldsymbol{V}[\mu]$ strongly in $L^2(\mathcal{P}_2(\mathbb{R}^d), \mu; \mathbb{R}^d)$, so that $\boldsymbol{V}[\mu] \in \mathrm{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}'$.

(5.27) is obvious. Since for every $F = \mathsf{L}_\phi$, $\phi \in \mathrm{C}_b^1$ $\mathrm{D}F[\mu] = \nabla\phi \in \mathrm{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, it is immediate to check that $\mathrm{D}F \in \mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ for every cylinder function. The closure property of $\mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ then yields the analogous conclusion for the Wasserstein differential of $\mathrm{D}_{\mathfrak{m}} F$ of a Sobolev function $F \in H^{1,2}(\mathbb{W}_2)$.

Let us eventually consider (5.28): it is sufficient to prove that any $\boldsymbol{V} \in \mathscr{T}^\perp$ belongs to $\big(\mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})\big)^\perp$, where $\perp$ denotes the orthogonal complement in the Hilbert space $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathfrak{m}; \mathbb{R}^d)$. If $\boldsymbol{V} \in \mathscr{T}^\perp$ is a Borel vector field, then

$$\int_{\mathcal{P}_2(\mathbb{R}^d)} \Big(\int \langle \nabla\zeta, \boldsymbol{V}(\mu, x)\rangle \,\mathrm{d}\mu(x)\Big) H(\mu) \,\mathrm{d}\mathfrak{m}(\mu) = 0$$

for every $\zeta \in \mathscr{C},\ H \in \mathscr{L}$. Since $\mathscr{L}$ is dense in $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ we have for every $\zeta \in \mathscr{C}$

$$\int \langle \nabla\zeta, \boldsymbol{V}(\mu, x)\rangle \,\mathrm{d}\mu(x) = 0 \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d)$$

Since $\mathscr{C}$ is countable, we can find a $\mathfrak{m}$-negligible set $\mathcal{N} \subset \mathcal{P}_2(\mathbb{R}^d)$ such that

$$\int \langle \nabla\zeta, \boldsymbol{V}(\mu, x)\rangle \,\mathrm{d}\mu(x) = 0 \quad \text{for every } \zeta \in \mathscr{C} \text{ and every } \mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}$$

which shows that $\boldsymbol{V}[\mu] \in \Big(\operatorname{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)\Big)^\perp$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}$, so that for every $\boldsymbol{W} \in \operatorname{Tan}(\mathcal{P}_2(\mathbb{R}^d), \boldsymbol{\mathfrak{m}})$

$$\int \langle \boldsymbol{V}(\mu, x), \boldsymbol{W}(\mu, x)\rangle \,\mathrm{d}\boldsymbol{\mathfrak{m}} = \int_{\mathcal{P}_2(\mathbb{R}^d)} \Big(\int_{\mathbb{R}^d} \langle \boldsymbol{V}[\mu](x), \boldsymbol{W}[\mu](x)\rangle \,\mathrm{d}\mu(x)\Big) \,\mathrm{d}\mathfrak{m}(\mu) = 0. \quad \square$$

Let us collect a few simple properties of G_0.

Lemma 5.9. *Let* G_0 *be as in* (5.25).

(1) G_0 *is a closed subspace of* $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$ *and coincides with the set*

$$\mathbf{D}_{\mathfrak{m}} 0 = \big\{\boldsymbol{V} \in L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d) : (0, \boldsymbol{V}) \in \mathrm{G}\big\}. \tag{5.29}$$

(2) *For every* $\boldsymbol{V} \in \mathrm{G}_0$ *there exists a sequence* $F_n \in \mathfrak{C}^1_b\big(\mathcal{P}_2(\mathbb{R}^d)\big)$, $n \in \mathbb{N}$, *such that*

$$F_n \to 0 \text{ in } L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}),\ \mathrm{D}F_n \to \boldsymbol{V} \text{ strongly in } L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d). \tag{5.30}$$

Every element $\boldsymbol{V} \in \mathrm{G}_0$ *is therefore characterized by the property*

$$\begin{aligned} &\forall \varepsilon > 0 \ \ \exists F \in \mathfrak{C}^1_b\big(\mathcal{P}_2(\mathbb{R}^d)\big) : \\ &\quad \|F\|_{L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})} \le \varepsilon, \quad \|\mathrm{D}F - \boldsymbol{V}\|_{L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)} \le \varepsilon. \end{aligned} \tag{5.31}$$

(3) G_0 *satisfies the locality property*

$$\text{for every } \boldsymbol{V} \in \mathrm{G}_0,\ H \in L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}) : \quad H\boldsymbol{V} \in \mathrm{G}_0. \tag{5.32}$$

Proof. We have already observed that G is a closed vector space, coinciding with the weak closure of $\big\{(F, \mathrm{D}F) : F \in \mathfrak{C}^1_b\big(\mathcal{P}_2(\mathbb{R}^d)\big)\big\}$; in view of (5.24), (5.25) precisely characterizes the elements $\boldsymbol{V}$ for which $(0, \boldsymbol{V}) \in \mathrm{G}$. Therefore the first two claims are obvious.

Let us eventually prove the last claim. We first consider the case when $H \in \mathfrak{C}^1_b\big(\mathcal{P}_2(\mathbb{R}^d)\big)$. If $\boldsymbol{V} \in \mathrm{G}_0$ we can find a sequence $F_n \in \mathfrak{C}^1_b\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ such that (5.30) holds. Setting $G_n := HF_n$, since H is bounded we clearly have $G_n \to 0$ strongly in $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$; moreover, by the Leibniz rule we get

$$\mathrm{D}G_n = H\mathrm{D}F_n + F_n\mathrm{D}H \to H\boldsymbol{V} \tag{5.33}$$

since $\mathrm{D}H \in L^\infty(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$ and $F_n \to 0$ strongly in $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$. We deduce that $H\boldsymbol{V} \in \mathrm{G}_0$ as well.

If now H is a function in $L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ we can find by (2.3) a uniformly bounded sequence $H_n \in \mathfrak{C}^1_b\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ converging to H $\mathfrak{m}$-a.e. in $\mathcal{P}_2(\mathbb{R}^d)$, so that $H_n\boldsymbol{V} \to H\boldsymbol{V}$ in $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$. Being G_0 a closed subspace and $H_n\boldsymbol{V} \in \mathrm{G}_0$ by the previous step, we deduce that $H\boldsymbol{V} \in \mathrm{G}_0$. $\square$

We now define

$$\begin{aligned} \mathrm{T} := &\ \mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \boldsymbol{\mathfrak{m}}) \cap \mathrm{G}_0^{\perp} \\ = &\Big\{ \boldsymbol{V} \in \mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \boldsymbol{\mathfrak{m}}) : \langle \boldsymbol{V}, \boldsymbol{W} \rangle_{L^2} = 0 \text{ for every } \boldsymbol{W} \in \mathrm{G}_0 \Big\}, \end{aligned} \tag{5.34}$$

where $\perp$ denotes the orthogonal complement in the Hilbert space $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$. We can now obtain our main structure result.

Theorem 5.10. *For every $F \in H^{1,2}(\mathbb{W}_2)$ we have $\mathrm{D}_{\mathfrak{m}} F \in \mathrm{T}$ and for every $\boldsymbol{V} \in \mathrm{G}_0$ we have the pointwise orthogonality property*

$$\int_{\mathbb{R}^d} \mathrm{D}_{\mathfrak{m}} F(\mu, x) \cdot \boldsymbol{V}(\mu, x) \, \mathrm{d}\mu(x) = 0 \quad \textit{for } \mathfrak{m}\textit{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{5.35}$$

If $\boldsymbol{V} \in \mathbf{D}_{\mathfrak{m}} F$ then $\boldsymbol{V} - \mathrm{D}_{\mathfrak{m}} F \in \mathrm{G}_0$. In particular for every $F \in \mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ $\mathrm{D}F - \mathrm{D}_{\mathfrak{m}} F \in \mathrm{G}_0$ and for every $G \in H^{1,2}(\mathbb{W}_2)$

$$\begin{aligned} &\int_{\mathbb{R}^d} \mathrm{D}_{\mathfrak{m}} F(\mu, x) \cdot \mathrm{D}_{\mathfrak{m}} G(\mu, x) \, \mathrm{d}\mu(x) = \int_{\mathbb{R}^d} \mathrm{D}F(\mu, x) \cdot \mathrm{D}_{\mathfrak{m}} G(\mu, x) \, \mathrm{d}\mu(x) \\ &\qquad \textit{for } \mathfrak{m}\textit{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d). \end{aligned} \tag{5.36}$$

Finally, for every $F \in H^{1,2}(\mathbb{W}_2)$, $\mathrm{D}_{\mathfrak{m}} F$ is the element of minimal L^2-norm in $\mathbf{D}_{\mathfrak{m}} F$.

Proof. Let us first observe that if $F_n \in \mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ satisfies (5.30) and $\tilde{F}_n \in \mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ satisfies (5.1), we have $F_n + \tilde{F}_n \to F$ strongly in $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$, with $\mathrm{D}(F_n + \tilde{F}_n) \to \mathrm{D}_{\mathfrak{m}} F + \boldsymbol{V}$, so that the lower semicontinuity of the Cheeger energy with respect to L^2 convergence yields together with (5.2) that

$$\mathsf{CE}_2(F) = \int |\mathrm{D}_{\mathfrak{m}} F|^2 \, \mathrm{d}\boldsymbol{\mathfrak{m}} \le \int |\mathrm{D}_{\mathfrak{m}} F + \boldsymbol{V}|^2 \, \mathrm{d}\boldsymbol{\mathfrak{m}}. \tag{5.37}$$

Since $\boldsymbol{V}$ is arbitrary in G_0 we deduce that

$$\int \mathrm{D}_{\mathfrak{m}} F \cdot \boldsymbol{V} \, \mathrm{d}\boldsymbol{\mathfrak{m}} = 0 \quad \text{for every } \boldsymbol{V} \in \mathrm{G}_0.$$

Replacing $\boldsymbol{V}$ with $H\boldsymbol{V}$, $H \in L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ we get

$$\int_{\mathcal{P}_2(\mathbb{R}^d)} \Big(\int_{\mathbb{R}^d} \mathrm{D}_{\mathfrak{m}} F \cdot \boldsymbol{V} \, \mathrm{d}\mu(x) \Big) H(\mu) \, \mathrm{d}\mathfrak{m}(\mu) = 0 \quad \text{for every } \boldsymbol{V} \in \mathrm{G}_0, \ H \in L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}), \tag{5.38}$$

which yields (5.35).

If now $F_n \in \mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ converges strongly to F with $\mathrm{D}F \rightharpoonup \boldsymbol{G}$, selecting $\tilde{F}_n$ as above, we have $F_n - \tilde{F}_n \to 0$ strongly in $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ and $\mathrm{D}(F_n - \tilde{F}_n) \rightharpoonup \boldsymbol{G} - \mathrm{D}_\mathfrak{m} F$ weakly in $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$, so that $\boldsymbol{G} - \mathrm{D}_\mathfrak{m} F \in \mathrm{G}_0$. By (5.37) we conclude that $\mathrm{D}_\mathfrak{m} F$ is the element of minimal norm in $\mathbf{D}_\mathfrak{m} F = \mathrm{D}_\mathfrak{m} F + \mathrm{G}_0$. $\square$

We can give a "pointwise" interpretation of the orthogonality properties of the previous Theorem. Let us select an orthonormal set $\mathrm{O}_0 := \{\boldsymbol{V}_n : n \in \mathbb{N}\} \subset \mathcal{L}^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$ dense in G_0 (we are thus assuming that $\boldsymbol{V}_n$ are Borel vector fields everywhere defined). Since

$$\int\limits_{\mathcal{P}_2(\mathbb{R}^d)} \Big(\int\limits_{\mathbb{R}^d} |\boldsymbol{V}_n(\mu, x)|^2 \,\mathrm{d}\mu(x) \Big) \,\mathrm{d}\mathfrak{m}(\mu) = 1$$

we deduce that there exists a $\mathfrak{m}$-negligible set $\mathcal{N} \subset \mathcal{P}_2(\mathbb{R}^d)$ such that

$$\int\limits_{\mathbb{R}^d} |\boldsymbol{V}_n(\mu, x)|^2 \,\mathrm{d}\mu(x) < \infty \quad \text{for every } n \in \mathbb{N},\ \mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}. \tag{5.39}$$

We thus define $\mathrm{G}_0[\mu] := \overline{\mathrm{span}\{\boldsymbol{V}_n[\mu] : n \in \mathbb{N}\}} \subset L^2(\mathbb{R}^d, \mu; \mathbb{R}^d)$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}$ and $\mathrm{T}[\mu] := \big(\mathrm{G}_0[\mu]\big)^{\perp} \cap \mathrm{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$, where here $\perp$ denotes the orthogonal complement in the Hilbert space $L^2(\mathbb{R}^d, \mu; \mathbb{R}^d)$.

Theorem 5.11. *Let $F \in H^{1,2}(\mathbb{W}_2)$ and $\boldsymbol{V} \in L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$.*

(1) *$\boldsymbol{V}$ belongs to G_0 if and only if, for $\mathfrak{m}$-a.e. μ, $\boldsymbol{V}[\mu] \in \mathrm{G}_0[\mu]$.*
(2) *$\boldsymbol{V}$ belongs to T if and only if, for $\mathfrak{m}$-a.e. μ, $\boldsymbol{V}[\mu] \in \mathrm{T}[\mu]$.*
(3) *$\mathrm{D}_\mathfrak{m} F[\mu] \in \mathrm{T}[\mu]$ for $\mathfrak{m}$-a.e. μ.*
(4) *If $F \in \mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ then, for $\mathfrak{m}$-a.e. $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\mathrm{D}_\mathfrak{m} F[\mu]$ is the $L^2(\mathbb{R}^d, \mu)$-orthogonal projection of $\mathrm{D}F[\mu]$ on $\mathrm{T}[\mu]$.*

Proof. If $\boldsymbol{V} \in \mathrm{G}_0$ we can write $\boldsymbol{V} = \lim_{N\to\infty} \boldsymbol{V}^N$ in $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$ where $\boldsymbol{V}^N = \sum_{n=1}^N u_n \boldsymbol{V}_n$ is the orthogonal projection of $\boldsymbol{V}$ on the space generated by $\{\boldsymbol{V}_1, \cdots, \boldsymbol{V}_N\}$, with $u_n := \langle V, V_n \rangle$. Clearly $\boldsymbol{V}^N[\mu] \in \mathrm{G}_0[\mu]$ for every $N \in \mathbb{N}$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}$. Moreover we can find a subsequence, not relabeled, and a $\mathfrak{m}$-negligible set $\mathcal{N}' \supset \mathcal{N}$ such that $\boldsymbol{V}^N[\mu] \to \boldsymbol{V}[\mu]$ in $L^2(\mathbb{R}^d, \mu; \mathbb{R}^d)$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}'$, so that $\boldsymbol{V}[\mu] \in \mathrm{G}_0[\mu]$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}'$.

Let now $\boldsymbol{V} \in L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$ be a vector field such that $\boldsymbol{V}[\mu] \in \mathrm{G}_0[\mu]$ for $\mathfrak{m}$-a.e. $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Since G_0 is a closed subspace, in order to show that $\boldsymbol{V} \in \mathrm{G}_0$ it is sufficient to prove that the scalar product with every element $\boldsymbol{W} \in \mathrm{G}_0^{\perp}$ vanishes.

If $\boldsymbol{W} \in \mathrm{G}_0^\perp$ then for every $H \in L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ and every $n \in \mathbb{N}$ we get

$$\int_{\mathcal{P}_2(\mathbb{R}^d)} \Big(\int_{\mathbb{R}^d} \boldsymbol{W} \cdot \boldsymbol{V}_n \, \mathrm{d}\mu(x) \Big) H(\mu) \, \mathrm{d}\mathfrak{m}(\mu) = 0,$$

since $H\boldsymbol{V}_n \in \mathrm{G}_0$ by (5.32). Being H arbitrary, we find that there exists a $\mathfrak{m}$-negligible set $\mathcal{N}'' \subset \mathcal{P}_2(\mathbb{R}^d)$ such that

$$\int_{\mathbb{R}^d} \boldsymbol{W}[\mu] \cdot \boldsymbol{V}_n[\mu] \, \mathrm{d}\mu = 0 \quad \text{for every } n \in \mathbb{N}, \ \mu \in \mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{N}'',$$

so that $\boldsymbol{W}[\mu] \in \big(\mathrm{G}_0[\mu]\big)^\perp$ for $\mathfrak{m}$-a.e. $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. We then deduce that

$$\int_{\mathbb{R}^d} \boldsymbol{W}[\mu] \cdot \boldsymbol{V}[\mu] \, \mathrm{d}\mu = 0 \quad \text{for } \mathfrak{m}\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{R}^d),$$

and therefore

$$\langle \boldsymbol{W}, \boldsymbol{V} \rangle_{L^2} = \int_{\mathcal{P}_2(\mathbb{R}^d)} \Big(\int_{\mathbb{R}^d} \boldsymbol{W} \cdot \boldsymbol{V} \, \mathrm{d}\mu(x) \Big) \, \mathrm{d}\mathfrak{m}(\mu) = 0.$$

The previous argument also shows that a vector field $\boldsymbol{V}$ belongs to $\mathrm{G}_0^\perp$ if and only if $\boldsymbol{V}[\mu] \in \big(\mathrm{G}_0[\mu]\big)^\perp$ for $\mathfrak{m}$-a.e. $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. This fact, together with the very definition of $\mathrm{Tan}(\mathcal{P}_2(\mathbb{R}^d), \boldsymbol{\mathfrak{m}})$ (5.26), yields claim (2).

Claim (3) just follows by Theorem 5.10, since (5.35) shows that, for every $F \in H^{1,2}(\mathbb{W}_2)$, $\mathrm{D}_\mathfrak{m} F[\mu] \in \mathrm{T}[\mu]$ for $\mathfrak{m}$-a.e. $\mu \in \mathcal{P}_2(\mathbb{R}^d)$.

If $F \in \mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)$, combining claim 1 and Theorem 5.10, we see that $\mathrm{D}F[\mu] - \mathrm{D}_\mathfrak{m} F[\mu] \in \mathrm{G}_0[\mu] \subset \big(\mathrm{T}[\mu]\big)^\perp$ $\mathfrak{m}$-a.e., so that $\mathrm{D}_\mathfrak{m} F[\mu]$ is the $L^2(\mathbb{R}^d, \mu; \mathbb{R}^d)$-orthogonal projection of $\mathrm{D}F[\mu]$ on $\mathrm{T}[\mu]$, as stated in Claim (4). □

We can now interpret the above results in terms of the nonsmooth tangent and cotangent structures introduced and developed by Gigli in [25]. Since we are in the Hilbertian case, we can identify the cotangent module $L^2(T^*\mathcal{P}_2(\mathbb{R}^d))$ and dual tangent module $L^2(T\mathcal{P}_2(\mathbb{R}^d))$ with the Hilbert space T defined by (5.34). Let us report a useful characterization of the cotangent module $L^2(T^*X)$ [26, Theorem 4.1.1] for a general metric measure space $(X, \mathsf{d}, \mathfrak{m})$.

Theorem 5.12. *Let $(X, \mathsf{d}, \mathfrak{m})$ be a metric measure space. Then there exists a unique pair $((\mathcal{M}, \|\cdot\|_\mathcal{M}, \cdot_\mathcal{M}, |\cdot|_\mathcal{M}), \mathsf{diff})$ such that $(\mathcal{M}, \|\cdot\|_\mathcal{M}, \cdot_\mathcal{M}, |\cdot|_\mathcal{M})$ is a $L^2(X, \mathfrak{m})$-normed $L^\infty(X, \mathfrak{m})$ module (cf. [26, Definition 3.1.1]) and $\mathsf{diff} : D^{1,2}(X, \mathsf{d}, \mathfrak{m}) \to \mathcal{M}$ is a linear operator such that*

(i) $|\mathsf{diff}(f)|_{\mathcal{M}} = |\mathrm{D}f|_\star$ $\mathfrak{m}$-*a.e. in* X *for every* $f \in D^{1,2}(X, \mathsf{d}, \mathfrak{m})$.
(ii) $\mathcal{M}$ *is generated by* $\{\mathsf{diff}(f) : f \in D^{1,2}(X, \mathsf{d}, \mathfrak{m})\}$.

Uniqueness is intended in the following sense: if $((\tilde{\mathcal{M}}, \|\cdot\|_{\tilde{\mathcal{M}}}, \cdot_{\tilde{\mathcal{M}}}, |\cdot|_{\tilde{\mathcal{M}}}), \tilde{\mathsf{diff}})$ *is another pair with the above properties, then there exists a unique module isomorphism* $\mathcal{J} : \mathcal{M} \to \tilde{\mathcal{M}}$ *such that* $\tilde{\mathsf{diff}} = \mathcal{J} \circ \mathsf{diff}$.

We thus have the following result.

Theorem 5.13. *There exists a unique module isomorphism* $\mathcal{I} : \mathrm{T} \to L^2(T^*\mathcal{P}_2(\mathbb{R}^d)) \cong L^2(T\mathcal{P}_2(\mathbb{R}^d))$ *such that* $\mathcal{I} \circ \mathrm{D}_{\mathfrak{m}}$ *coincides with the abstract differential operator taking values in* $L^2(T^*\mathcal{P}_2(\mathbb{R}^d))$ *as in [25, Definition 2.2.2].*

Proof. It is enough to show that T (with an appropriate module structure) and the map $\mathrm{D}_{\mathfrak{m}}$ satisfy the properties listed in Theorem 5.12.
If as $\|\cdot\|_{\mathrm{T}}$ we take the $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$ norm, it is clear that $(\mathrm{T}, \|\cdot\|_{\mathrm{T}})$ is a Banach space, being closed by Lemma 5.8. The pointwise product $\cdot_{\mathrm{T}} : L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}) \times \mathrm{T} \to \mathrm{T}$ is well defined by (5.27) and (5.32), bilinear and associative in $L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ by definition. Defining the pointwise norm $|\cdot|_{\mathrm{T}}$ as the map sending $\boldsymbol{V} \in \mathrm{T}$ to $\|\boldsymbol{V}[\mu]\|_\mu$, we immediately have that $\|\boldsymbol{V}\|_{\mathrm{T}} = \||\boldsymbol{V}|_{\mathrm{T}}\|_{L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})}$ and $|H \cdot_{\mathrm{T}} \boldsymbol{V}|_{\mathrm{T}} = |H||\boldsymbol{V}|_{\mathrm{T}}$ $\mathfrak{m}$-a.e. in $\mathcal{P}_2(\mathbb{R}^d)$ for every $\boldsymbol{V} \in \mathrm{T}$ and every $H \in L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$. This shows that $(\mathrm{T}, \|\cdot\|_{\mathrm{T}}, \cdot_{\mathrm{T}}, |\cdot|_{\mathrm{T}})$ is a $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$-normed $L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ module. Taking as diff the map $\mathrm{D}_{\mathfrak{m}} : D^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m}) \to \mathrm{T}$, we see that it is well defined and linear by Theorem 5.1 and Theorem 5.10. Property (i) of Theorem 5.12 follows by (5.8). Finally property (ii) of Theorem 5.12, meaning that ([26, Definition 3.1.13]) T coincides with the $\|\cdot\|_{\mathrm{T}}$-closure of

$$\mathrm{T}_0 := \operatorname{span}\left\{H\mathrm{D}_{\mathfrak{m}}F : H \in L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m}),\ F \in D^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m})\right\},$$

follows by (5.28) and the definition of T. Indeed, let $\mathscr{L} \subset L^\infty(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ be a dense subset of $L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ and $\mathscr{C}$ be a dense subset of $\mathrm{C}_c^\infty(\mathbb{R}^d)$ with respect to the Lipschtiz norm as in Lemma 5.8. If $\boldsymbol{V} \in \mathrm{T}$, in particular $\boldsymbol{V} \in \operatorname{Tan}(\mathcal{P}_2(\mathbb{R}^d), \boldsymbol{\mathfrak{m}})$ so that we can find by (5.28) numbers $(N_n)_n \subset \mathbb{N}$, $(\{\alpha_n^i\}_{i=1}^{N_n})_n \subset \mathbb{R}$ and functions $(\{H_n^i\}_{i=1}^{N_n})_n \subset \mathscr{L}$, $(\{\zeta_n^i\}_{i=1}^{N_n})_n \subset \mathscr{C}$ such that the sequence

$$\boldsymbol{V}_n(\mu, x) := \sum_{i=1}^{N_n} \alpha_n^i H_n^i(\mu) \nabla \zeta_n^i(x), \quad (\mu, x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d\, n \in \mathbb{N}$$

converges to $\boldsymbol{V}$ in $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \boldsymbol{\mathfrak{m}}; \mathbb{R}^d)$. Consider now the functions $F_n^i := \mathsf{L}_{\zeta_n^i}$ and the sequence

$$\boldsymbol{V}'_n(\mu, x) := \sum_{i=1}^{N_n} \alpha_n^i H_n^i(\mu) \mathrm{D}_{\mathfrak{m}} F_n^i, \quad (\mu, x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d\, n \in \mathbb{N}.$$

It is clear that $(\boldsymbol{V}'_n)_n \subset \mathrm{T}_0$; by Theorem 5.11, $\boldsymbol{V}'_n$ is the orthogonal projection of $\boldsymbol{V}_n$ on T for every $n \in \mathbb{N}$, so that $\boldsymbol{V}'_n$ converges to $\boldsymbol{V}$ in $L^2(\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d, \mathfrak{m}; \mathbb{R}^d)$.

Theorem 5.12 gives thus the existence of a unique module isomorphism $\mathcal{I} : \mathrm{T} \to L^2(T^*\mathcal{P}_2(\mathbb{R}^d))$.

Finally, notice that $L^2(T^*\mathcal{P}_2(\mathbb{R}^d)) \cong L^2(T\mathcal{P}_2(\mathbb{R}^d))$ since $(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m})$ is infinitesimally Hilbertian by Corollary 4.11 (see also [26, Theorem 4.3]). $\square$

5.2. Examples

Isometric embedding of Euclidean Sobolev spaces

Let Ω be a Lipschitz bounded open set in $\mathbb{R}^d$. For every $\omega \in \Omega$ let us consider the Dirac mass δ_ω concentrated at ω. The map $\iota : \omega \mapsto \delta_\omega$ is an isometry between $\mathbb{R}^d$ and $\iota(\mathbb{R}^d) \subset \mathcal{P}_2(\mathbb{R}^d)$. Setting $\mathfrak{m} := \iota_\sharp \mathscr{L}^d \llcorner \Omega$ we easily see that $H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m})$ is isomorphic to $H^{1,2}(\Omega)$.

In this case only Dirac masses are involved and cylinder functions are of the form $F(\delta_\omega) = \psi(\boldsymbol{\phi}(\omega))$, so that the Wasserstein gradient reduces to the usual gradient of $\psi \circ \boldsymbol{\phi}$.

Another isometric embedding is also possible: we fix a reference measure $\lambda \in \mathcal{P}_2(\mathbb{R}^d)$ symmetric w.r.t. the origin and we consider the map $\iota : \Omega \to \mathcal{P}_2(\mathbb{R}^d)$ given by

$$\iota(\omega) := \lambda(\cdot - \omega) = (\mathsf{t}_\omega)_\sharp \lambda, \quad \mathsf{t}_\omega(x) := x + \omega, \quad \omega \in \Omega. \tag{5.40}$$

To every function $F : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ corresponds a map $\hat{F} : \Omega \to \mathbb{R}$ defined as

$$\hat{F}(\omega) := F((\mathsf{t}_\omega)_\sharp \lambda). \tag{5.41}$$

In the case of a cylinder function as in (4.2) we get

$$\hat{F}(\omega) = \psi\Big(\int \phi_1(x+\omega) \,\mathrm{d}\lambda(x), \cdots, \int \phi_N(x+\omega) \,\mathrm{d}\lambda(x) \Big) = \psi\Big(\phi_1 * \lambda(\omega), \cdots, \phi_N * \lambda(\omega) \Big). \tag{5.42}$$

In this case (identifying $\iota(\omega)$ with ω) we have

$$\mathrm{D}F(\omega, x) = \sum_{j=1}^N \partial\psi_j(\phi_1 * \lambda(\omega), \cdots, \phi_N * \lambda(\omega)) \nabla\phi_j(x) \tag{5.43}$$

and

$$\|\mathrm{D}F[\omega]\|_\omega^2 = \int_{\mathbb{R}^d} \Big| \sum_{j=1}^N \partial\psi_j(\phi_1 * \lambda(\omega), \cdots, \phi_N * \lambda(\omega)) \nabla\phi_j(x+\omega) \Big|^2 \mathrm{d}\lambda(x). \tag{5.44}$$

On the other hand, ι is an isometry of $\mathbb{R}^d$ into $\mathcal{P}_2(\mathbb{R}^d)$, so that the space $H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m})$ is still isomorphic to $H^{1,2}(\Omega)$. It follows that the $\mathfrak{m}$-Wasserstein gradient of F is

$$\mathrm{D}_{\mathfrak{m}} F(\omega, x) = \sum_{j=1}^{N} \partial\psi_j(\phi_1 * \lambda(\omega), \cdots, \phi_N * \lambda(\omega)) \nabla\phi_j * \lambda(\omega) \tag{5.45}$$

independent of x and the minimal relaxed gradient is

$$|\mathrm{D}_{\mathfrak{m}} F|_\star^2(\omega) = \Big| \sum_{j=1}^{N} \partial\psi_j(\phi_1 * \lambda(\omega), \cdots, \phi_N * \lambda(\omega)) \nabla\phi_j * \lambda(\omega) \Big|^2 \tag{5.46}$$

Gaussian distributions

Let now $\kappa = N(\omega, \Sigma) := (\det(2\pi\Sigma))^{-1/2} \mathrm{e}^{-\frac{1}{2}\langle \omega, \Sigma^{-1}\omega\rangle} \mathscr{L}^d$ be a Gaussian measure with mean ω and covariance matrix $\Sigma \in \mathrm{Sym}^+(d)$, the space of symmetric and positive definite $d \times d$-matrices; we consider the set

$$\mathcal{N}^d := \Big\{ \mathcal{N}(\omega, \Sigma) : \omega \in \mathbb{R}^d,\ \Sigma \in \mathrm{Sym}^+(d) \Big\}, \tag{5.47}$$

endowed with the Wasserstein distance and a finite positive Borel measure $\mathfrak{m}$ concentrated on $\mathcal{N}^d$. Since

$$W_2^2(N(\omega_1, \Sigma_1), N(\omega_2, \Sigma_2)) = |\omega_1 - \omega_2|^2 + \mathrm{tr}\Sigma_1 + \mathrm{tr}\Sigma_2 - 2\mathrm{tr}\Big(\Sigma_1^{1/2}\Sigma_2\Sigma_1^{1/2}\Big)^{1/2}, \tag{5.48}$$

$H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \mathfrak{m})$ is isometric to $H^{1,2}(U, \mathsf{d}, \hat{\mathfrak{m}})$ where $U = \mathbb{R}^d \times \overline{\mathrm{Sym}^+(d)} \subset \mathbb{R}^d \times \mathbb{R}^{d\times d}$ endowed with the distance d induced by the formula (5.48) and $\hat{\mathfrak{m}}$ is the measure induced by $\mathfrak{m}$.

The closable case

Following [17] (here in the simpler setting of the Euclidean space, but see Section 6.2 below), we assume that $\mathfrak{m}$ has no atoms and the following integration by parts formula: for every $G \in \mathrm{FC}_c^\infty\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ and $w \in \mathrm{C}_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$ there exists $\mathrm{D}_w^* G \in L^2(\mathcal{P}_2(\mathbb{R}^d), \mathfrak{m})$ such that for every $F \in \mathrm{FC}_c^\infty\big(\mathcal{P}_2(\mathbb{R}^d)\big)$ it holds

$$\int_{\mathcal{P}_2(\mathbb{R}^d)} \left(\int_{\mathbb{R}^d} \mathrm{D}F(\mu, x) \cdot w(x)\, \mathrm{d}\mu(x) \right) G(\mu)\, \mathrm{d}\mathfrak{m}(\mu) = \int_{\mathcal{P}_2(\mathbb{R}^d)} \mathrm{D}_w^* G(\mu) F(\mu)\, \mathrm{d}\mathfrak{m}(\mu).$$

This equality implies that $\mathrm{G}_0 = \{0\}$ i.e. that pCE_2 is closable. We notice that the measure $\mathfrak{m}$ induced by the immersion in the space of delta measures considered at the beginning of this section satisfies the integration by parts formula above (see also Example 5.4 in [17]). In [17], in case the base space is a compact Riemannian manifold, are reported important examples of measures $\mathfrak{m}$ satisfying the (Riemannian analogue of the) integration by parts formula: the normalized mixed Poisson measure (Example 5.11 in [17] and [1,41]), the entropic measure over $\mathbb{S}^1$ (Example 5.15 in [17] and [50], see also the multidimensional case [47]) and the Malliavin–Shavgulidze image measure (Example 5.18 in [17] and [33]).

6. Extensions to Riemannian manifolds and Hilbert spaces

The aim of this Section is to extend the density result stated in Theorem 4.10 from the finite dimensional and flat space $\mathbb{R}^d$ to Riemannian manifolds and (possibly infinite dimensional) Hilbert spaces. Our first step deals with manifold embedded in some Euclidean space $\mathbb{R}^d$ and in fact we will consider more general closed subsets of $\mathbb{R}^d$.

6.1. Intrinsic Wasserstein spaces on closed subsets of $\mathbb{R}^d$

In this subsection we denote by ϱ the Euclidean distance on $\mathbb{R}^d$. $\mathcal{P}_2(\mathbb{R}^d)$ still denotes the subset of Borel probability measures on $\mathbb{R}^d$ with finite second ϱ-moment and W_2 is the Wasserstein distance on $\mathcal{P}_2(\mathbb{R}^d)$ induced by ϱ.

We assume that $C \subset \mathbb{R}^d$ is a closed set and that σ is a distance on C such that (C, σ) is a complete and separable metric space and

$$\varrho(x_1, x_2) \le \sigma(x_1, x_2) \le \varrho_{C,\ell}(x_1, x_2) \quad \text{for every } x_1, x_2 \in C, \tag{6.1}$$

where $\varrho_{C,\ell}$ is defined as in (2.65) with respect to the distance $\mathsf{d} := \varrho$. Since the topology induced by σ is stronger than the Euclidean topology and they are both Polish topologies, the Borel sets of (C, σ) coincide with the Borel sets of C as a subset of the Euclidean space $\mathbb{R}^d$. This means that every Borel probability measure on $\mathbb{R}^d$ with support contained in C can be identified with a Borel probability measure in (C, σ). Conversely any probability measure on (C, σ) extends to a probability measure on $\mathbb{R}^d$. We can thus denote unambiguously by $\mathcal{P}(C)$ the set of Borel probability measures on C and by $\mathcal{P}_{2,\sigma}(C)$ the elements of $\mathcal{P}(C)$ with finite second σ-moment.

$\mathcal{P}_{2,\sigma}(C)$ can be identified with the subset of $\mathcal{P}_2(\mathbb{R}^d)$

$$\left\{ \mu \in \mathcal{P}_2(\mathbb{R}^d) : \mathrm{supp}(\mu) \subset C, \quad \int_C \sigma^2(x_0, x) \, \mathrm{d}\mu(x) < +\infty \text{ for some } x_0 \in C \right\}.$$

We will denote by $\iota : C \to \mathbb{R}^d$ the inclusion map; $\boldsymbol{\iota} : \mathcal{P}_{2,\sigma}(C) \to \mathcal{P}_2(\mathbb{R}^d)$ is the corresponding continuous injection given by $\boldsymbol{\iota}(\mu) := \iota_\sharp \mu$, which may be identified with the inclusion map of $\mathcal{P}_{2,\sigma}(C)$ into $\mathcal{P}_{2,\sigma}(\mathbb{R}^d)$.

Since $(\mathcal{P}_2(C), W_{2,\sigma})$ is a complete and separable metric space and the topology induced by $W_{2,\sigma}$ is stronger than the topology induced by W_2, we deduce that $\mathcal{P}_{2,\sigma}(C)$ is a Lusin (and therefore Borel) subset of $\mathcal{P}_2(\mathbb{R}^d)$.

If $\mathfrak{m}$ is a finite and positive Borel measure on $\mathcal{P}_{2,\sigma}(C)$, $\boldsymbol{\iota}_\sharp \mathfrak{m}$ is the Borel measure in $\mathcal{P}_2(\mathbb{R}^d)$ which is concentrated on $\mathcal{P}_{2,\sigma}(C)$ and satisfies $\boldsymbol{\iota}_\sharp \mathfrak{m}(Z) = \mathfrak{m}(Z \cap \mathcal{P}_{2,\sigma}(C))$ for every Borel set $Z \subset \mathcal{P}_2(\mathbb{R}^d)$.

In a similar way, if $F : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is a Borel (or $\boldsymbol{\iota}_\sharp \mathfrak{m}$- measurable) map, we will set $\boldsymbol{\iota}^* F := F \circ \boldsymbol{\iota} : \mathcal{P}_{2,\sigma}(C) \to \mathbb{R}$.

Theorem 6.1. *We have $H^{1,2}(\mathcal{P}_{2,\sigma}(C), W_{2,\sigma}, \mathfrak{m}) \cong H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \boldsymbol{\iota}_\sharp \mathfrak{m})$ with equal minimal relaxed gradient, meaning that*

$$|\mathrm{D}(\boldsymbol{\iota}^* F)|_\star = \boldsymbol{\iota}^* \left(|\mathrm{D}F|_\star\right) \text{ for every } F \in H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \boldsymbol{\iota}_\sharp \mathfrak{m}). \tag{6.2}$$

In particular $H^{1,2}(\mathcal{P}_{2,\sigma}(C), W_{2,\sigma}, \mathfrak{m})$ is a Hilbert space and the algebra of cylinder functions $\boldsymbol{\iota}^*\big(\mathfrak{C}_b^1\big(\mathcal{P}_2(\mathbb{R}^d)\big)\big)$ is dense in $H^{1,2}(\mathcal{P}_{2,\sigma}(C), W_{2,\sigma}, \mathfrak{m})$ in the sense of (4.16).

Proof. We want to apply Theorem 2.24 where $X := \mathcal{P}_2(\mathbb{R}^d)$, $\mathsf{d} := W_2$, $Y := \mathcal{P}_{2,\sigma}(C)$, and $\delta := W_{2,\sigma}$. The first assumption of Condition (A), $\boldsymbol{\iota}_\sharp \mathfrak{m}(\mathcal{P}_2(\mathbb{R}^d) \setminus \mathcal{P}_{2,\sigma}(C)) = 0$, is clearly satisfied by construction.

In order to prove (2.71) we consider a W_2-Lipschitz curve $\mu : [0, \ell] \to \mathcal{P}_2(\mathbb{R}^d)$ parametrized by the W_2-arc-length such that $\mu_s \in \mathcal{P}_{2,\sigma}(C)$ for $\mathscr{L}^1$-a.e. $s \in [0, \ell]$. Since the map μ is continuous in $\mathcal{P}_2(\mathbb{R}^d)$, C is a closed set, and $\mu_s(\mathbb{R}^d \setminus C) = 0$ for $\mathscr{L}^1$-a.e. $s \in [0, \ell]$, we conclude that $\mu_s(\mathbb{R}^d \setminus C) = 0$ for every $s \in [0, \ell]$.

By [4, Theorem 8.2.1, Theorem 8.3.1]) there exists a measure $\boldsymbol{\eta} \in \mathcal{P}(\mathrm{C}([0, \ell]; \mathbb{R}^d))$ concentrated on absolutely continuous curves such that $(\mathsf{e}_t)_\sharp(\boldsymbol{\eta}) = \mu_t$ for every $t \in [0, \ell]$ and

$$\int |\gamma'(t)|^2 \,\mathrm{d}\boldsymbol{\eta}(\gamma) = \int |\dot\gamma|_\varrho^2(t) \,\mathrm{d}\boldsymbol{\eta}(\gamma) = 1 \quad \text{for a.e. } t \in [0, \ell]. \tag{6.3}$$

Let us also consider the function $\zeta(x) := \mathrm{dist}(x, C) \wedge 1$, $x \in \mathbb{R}^d$, where $\mathrm{dist}(x, C) := \min_{z \in C} \varrho(x, z)$. ζ is a bounded Lipschitz function which vanishes precisely on C. Fubini's Theorem yields

$$\int \Big(\int_0^\ell \zeta(\gamma(t)) \,\mathrm{d}t \Big) \,\mathrm{d}\boldsymbol{\eta}(\gamma) = \int_0^\ell \int \zeta(\mathsf{e}_t(\gamma)) \,\mathrm{d}\boldsymbol{\eta}(\gamma) \,\mathrm{d}t = \int_0^\ell \int_{\mathbb{R}^d} \zeta \,\mathrm{d}\mu_t \,\mathrm{d}t = 0$$

since $\int \zeta(x) \,\mathrm{d}\mu_t = 0$ for $\mathscr{L}^1$-a.e. $t \in (0, \ell)$.

It follows that $\int_0^\ell \zeta(\gamma(t)) \,\mathrm{d}t = 0$ for $\boldsymbol{\eta}$-a.e. γ, so that the set of $t \in [0, \ell]$ for which $\gamma(t) \in C$ is dense in $[0, \ell]$. Being C closed, we conclude that γ takes values in C for $\boldsymbol{\eta}$-a.e. γ.

We can now estimate the $W_{2,\sigma}$ distance between the two measures μ_{t_0} and μ_{t_1}, where $0 \le t_0 < t_1 \le \ell$:

$$\begin{aligned} W_{2,\sigma}^2(\mu_{t_0}, \mu_{t_1}) &\le \int \sigma^2(\gamma(t_0), \gamma(t_1)) \,\mathrm{d}\boldsymbol{\eta}(\gamma) \\ &\le \int \Big(\int_{t_0}^{t_1} |\dot\gamma|_\sigma(s) \,\mathrm{d}s \Big)^2 \,\mathrm{d}\boldsymbol{\eta}(\gamma) \end{aligned}$$

$$
\begin{aligned}
&= \int \Big(\int_{t_0}^{t_1} |\dot\gamma|_\varrho(s)\,\mathrm{d}s \Big)^2 \,\mathrm{d}\boldsymbol{\eta}(\gamma) \\
&\le (t_1 - t_0) \int \int_{t_0}^{t_1} |\dot\gamma|_\varrho^2 \,\mathrm{d}s\,\mathrm{d}\boldsymbol{\eta}(\gamma) \\
&= (t_1 - t_0) \int_{t_0}^{t_1} \int |\dot\gamma|_\varrho^2 \,\mathrm{d}\boldsymbol{\eta}(\gamma)\,\mathrm{d}s \\
&= (t_1 - t_0)^2,
\end{aligned}
$$

where we have used that $(\mathsf{e}_{t_0}, \mathsf{e}_{t_1})_\sharp \boldsymbol{\eta} \in \Gamma(\mu_{t_0}, \mu_{t_1})$, (6.1) and Remark 2.23 to say that $|\dot\gamma|_\varrho(s) = |\dot\gamma|_\sigma(s)$.

Choosing $t_0 \in [0, \ell]$ such that $\mu_{t_0} \in \mathcal{P}_{2,\sigma}(C)$ we deduce that $\mu_{t_1} \in \mathcal{P}_{2,\sigma}(C)$ as well for every $t_1 \in [0, \ell]$. This concludes the proof of property (A).

Condition (B) corresponds to

$$
W_2(\mu_0, \mu_1) \le W_{2,\sigma}(\mu_0, \mu_1) \le (W_2)_{Y,\ell}(\mu_0, \mu_1) \quad \text{for every } \mu_0, \mu_1 \in Y = \mathcal{P}_{2,\sigma}(C), \tag{6.4}
$$

where $(W_2)_{Y,\ell}(\mu_0, \mu_1)$ is defined as in (2.65) with W_2 in place of d. The first inequality immediately follows by (6.1); to prove the second one, we use (2.66) and the above estimate with $t_0 = 0$ and $t_1 = \ell$ for a W_2-Lipschitz curve $\mu : [0, \ell] \to Y$ such that $|\dot\mu|_{W_2} = 1$ a.e. in $[0, \ell]$ with $\mu|_{t=0} = \mu_0$ and $\mu|_{t=\ell} = \mu_1$. Taking the infimum w.r.t. ℓ we obtain (6.4). □

6.2. Wasserstein Sobolev space on complete Riemannian manifolds

In this subsection, we will briefly discuss the case of the Sobolev space $H^{1,2}(\mathcal{P}_2(\mathbb{M}), W_{2,\mathsf{d}_\mathbb{M}}, \mathfrak{m})$ where $(\mathbb{M}, \mathsf{d}_\mathbb{M})$ is a smooth and complete Riemannian manifold endowed with the canonical Riemannian distance $\mathsf{d}_\mathbb{M}$ (inducing the Wasserstein distance $W_{2,\mathsf{d}_\mathbb{M}}$) and $\mathfrak{m}$ is a finite and positive Borel measure on $\mathcal{P}_2(\mathbb{M})$. We will denote by $\mathscr{A}$ the unital algebra generated by $\big\{\mathsf{L}_f : f \in \mathrm{C}_c^1(\mathbb{M})\big\}$.

Theorem 6.2. *$H^{1,2}(\mathcal{P}_2(\mathbb{M}), W_{2,\mathsf{d}_\mathbb{M}}, \mathfrak{m})$ is a Hilbert space and the algebra $\mathscr{A}$ is (strongly) dense: for every $F \in H^{1,2}(\mathcal{P}_2(\mathbb{M}), W_{2,\mathsf{d}_\mathbb{M}}, \mathfrak{m})$ there exists a sequence $F_n \in \mathscr{A}$, $n \in \mathbb{N}$ such that*

$$
F_n \to F, \quad \mathrm{lip}(F_n) \to |\mathrm{D}F|_\star \quad \textit{strongly in } L^2(\mathcal{P}_2(\mathbb{M}), \mathfrak{m}). \tag{6.5}
$$

Proof. By Nash isometric embedding Theorem [36] we can find a dimension d, and an isometric embedding $\jmath : \mathbb{M} \to \jmath(\mathbb{M}) \subset \mathbb{R}^d$. On $M := \jmath(\mathbb{M})$ we can define the (Riemannian) metric d_M inherited by $\mathsf{d}_\mathbb{M}$: $\mathsf{d}_M(\jmath(x), \jmath(y)) = \mathsf{d}_\mathbb{M}(x, y)$ so that $\jmath$ is an isometry and

(M, d_M) is a complete and separable metric space. We denote by $\boldsymbol{\jmath} := \jmath_\sharp$ the corresponding isometry between $(\mathcal{P}_2(\mathbb{M}), W_{2,\mathsf{d}_\mathbb{M}})$ and $(\mathcal{P}_2(M), W_{2,\mathsf{d}_M})$ and we also set $\tilde{\mathfrak{m}} := \boldsymbol{\jmath}_\sharp \mathfrak{m}$ which is a positive and finite Borel measure on $\mathcal{P}_2(M)$.

It is clear that the map $\boldsymbol{\jmath}^* : F \mapsto F \circ \boldsymbol{\jmath}$ induces a linear isometric isomorphism between $H^{1,2}(\mathcal{P}_2(M), W_{2,\mathsf{d}_M}, \tilde{\mathfrak{m}})$ and $H^{1,2}(\mathcal{P}_2(\mathbb{M}), W_{2,\mathsf{d}_\mathbb{M}}, \mathfrak{m})$.

Since $\mathbb{M}$ is complete and $\jmath$ is an embedding, M is a closed subset of $\mathbb{R}^d$ and d_M induces on M the relative topology of $\mathbb{R}^d$. Since $\jmath$ is isometric, we also have

$$\varrho(y_1, y_2) \le \mathsf{d}_M(y_1, y_2) = \varrho_{M,\ell}(y_1, y_2) \quad \text{for every } y_1, y_2 \in M, \tag{6.6}$$

where $\varrho_{M,\ell}$ is as in (2.65) and ϱ denotes the Euclidean distance on $\mathbb{R}^d$.

As in Section 6.1, we can introduce the inclusion map $\iota : M \to \mathbb{R}^d$ and the corresponding $\boldsymbol{\iota} = \iota_\sharp : \mathcal{P}_{2,\mathsf{d}_M}(M) \to \mathcal{P}_2(\mathbb{R}^d)$. By Theorem 6.1 we have that the map $\boldsymbol{\iota}^* : F \mapsto F \circ \boldsymbol{\iota}$ provides a linear isometric isomorphism between $H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \boldsymbol{\iota}_\sharp \tilde{\mathfrak{m}})$ and $H^{1,2}(\mathcal{P}_{2,\mathsf{d}_M}(M), W_{2,\mathsf{d}_M}, \tilde{\mathfrak{m}})$ satisfying (6.2); we conclude that the map $\boldsymbol{\kappa}^* := \boldsymbol{\jmath}^* \circ \boldsymbol{\iota}^* = (\boldsymbol{\iota} \circ \boldsymbol{\jmath})^*$ is a linear isometric isomorphism between $H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \boldsymbol{\kappa}_\sharp \mathfrak{m})$ (notice that $\boldsymbol{\kappa}_\sharp = \boldsymbol{\iota}_\sharp \circ \boldsymbol{\jmath}_\sharp$) and $H^{1,2}(\mathcal{P}_{2,\mathsf{d}_\mathbb{M}}(M), W_{2,\mathsf{d}_\mathbb{M}}, \mathfrak{m})$ satisfying

$$|\mathrm{D}(\boldsymbol{\kappa}^* F)|_\star = \boldsymbol{\kappa}^* \left(|\mathrm{D}F|_\star\right) \text{ for every } F \in H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \boldsymbol{\kappa}_\sharp \mathfrak{m}). \tag{6.7}$$

This property in particular yields the Hilbertianity of $H^{1,2}(\mathcal{P}_2(\mathbb{M}), W_{2,\mathsf{d}_\mathbb{M}}, \mathfrak{m})$.

In order to prove that $\mathscr{A}$ is dense in $H^{1,2}(\mathcal{P}_2(\mathbb{M}), W_{2,\mathsf{d}_\mathbb{M}}, \mathfrak{m})$ we consider the algebra $\tilde{\mathscr{A}}$ generated by $\left\{\mathsf{L}_{\tilde{f}} : \tilde{f} \in \mathrm{C}_c^\infty(\mathbb{R}^d)\right\}$; Proposition 4.19 shows that $\tilde{\mathscr{A}}$ is strongly dense in $H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \tilde{\mathfrak{m}})$, so that $\mathscr{A}' := \boldsymbol{\kappa}^*(\tilde{\mathscr{A}})$ is strongly dense in $H^{1,2}(\mathcal{P}_{2,\mathsf{d}_\mathbb{M}}(\mathbb{M}), W_{2,\mathsf{d}_\mathbb{M}}, \mathfrak{m})$.

$\mathscr{A}'$ is generated by functions of the form $\boldsymbol{\kappa}^* \mathsf{L}_{\tilde{f}}$, $\tilde{f} \in \mathrm{C}_c^\infty(\mathbb{R}^d)$. Since

$$\boldsymbol{\kappa}^* \mathsf{L}_{\tilde{f}}(\mu) = \mathsf{L}_{\tilde{f}}(\boldsymbol{\kappa}(\mu)) = \int_{\mathbb{R}^d} \tilde{f}(\kappa(x)) \,\mathrm{d}\mu(x) \quad \text{for every } \mu \in \mathcal{P}_{2,\mathsf{d}_\mathbb{M}}(\mathbb{M}),$$

where $\kappa = \iota \circ \jmath$, we see that $\mathscr{A}'$ is generated by functions of the form $\mathsf{L}_{\tilde{f} \circ \kappa}$, so that $\mathscr{A}' \subset \mathscr{A}$ and a fortiori $\mathscr{A}$ is strongly dense in $H^{1,2}(\mathcal{P}_{2,\mathsf{d}_\mathbb{M}}(\mathbb{M}), W_{2,\mathsf{d}_\mathbb{M}}, \mathfrak{m})$ as well.

To prove (6.5) (involving the asymptotic Lipschitz constants of functions in $\mathscr{A}$ with respect to the Riemannian metric) we observe that for every $\tilde{F} \in \tilde{\mathscr{A}}$ [44, Lemma 3.1.14]

$$\boldsymbol{\kappa}^* \tilde{F} \in \mathscr{A}' \subset \mathscr{A}, \quad \boldsymbol{\kappa}^*(\mathrm{lip}_{W_2} \tilde{F}) \ge \mathrm{lip}_{W_{2,\mathsf{d}_\mathbb{M}}} \boldsymbol{\kappa}^* \tilde{F}. \tag{6.8}$$

Let now $F = \boldsymbol{\kappa}^* \tilde{F} \in H^{1,2}(\mathcal{P}_2(\mathbb{M}), W_{2,\mathsf{d}_\mathbb{M}}, \mathfrak{m})$ with $\tilde{F} \in H^{1,2}(\mathcal{P}_2(\mathbb{R}^d), W_2, \tilde{\mathfrak{m}})$; there exists a sequence $\tilde{F}_n \in \tilde{\mathscr{A}}$ such that

$$\tilde{F}_n \to \tilde{F}, \quad \mathrm{lip}_{W_2} \tilde{F}_n \to |\mathrm{D}\tilde{F}|_\star \quad \text{in } L^2(\mathcal{P}_2(\mathbb{R}^d), \tilde{\mathfrak{m}}).$$

Applying the linear isometric isomorphism $\boldsymbol{\kappa}^*$, we deduce that the sequence $\boldsymbol{\kappa}^* F_n \in \mathscr{A}'$ satisfies

$$\boldsymbol{\kappa}^* \tilde{F}_n \to F, \quad \boldsymbol{\kappa}^*\big(\operatorname{lip}_{W_2} \tilde{F}_n\big) \to \boldsymbol{\kappa}^*\big(|\mathrm{D}\tilde{F}|_\star\big) = |\mathrm{D}F|_\star \quad \text{in } L^2(\mathcal{P}_{2,\mathsf{d}_{\mathbb{M}}}(\mathbb{M}), \mathfrak{m}). \tag{6.9}$$

Up to extracting a suitable (not relabeled) subsequence and using (6.8), we can suppose that $\operatorname{lip}_{W_{2,\mathsf{d}_{\mathbb{M}}}} \boldsymbol{\kappa}^* \tilde{F}_n$ converges weakly in $L^2(\mathcal{P}_2(\mathbb{M}), W_{2,\mathsf{d}_{\mathbb{M}}})$ to some $G \in L^2(\mathcal{P}_2(\mathbb{M}), W_{2,\mathsf{d}_{\mathbb{M}}})$ relaxed gradient of F. (6.8) and (6.9) also yield

$$\begin{aligned}\int G^2 \,\mathrm{d}\mathfrak{m} &\le \limsup_{n\to\infty} \int (\operatorname{lip}_{W_{2,\mathsf{d}_{\mathbb{M}}}} \boldsymbol{\kappa}^* F_n)^2 \,\mathrm{d}\mathfrak{m} \le \limsup_{n\to\infty} \int \Big(\boldsymbol{\kappa}^*(\operatorname{lip}_{W_2} \tilde{F}_n)\Big)^2 \,\mathrm{d}\mathfrak{m} \\ &= \int |\mathrm{D}F|_\star^2 \,\mathrm{d}\mathfrak{m},\end{aligned}$$

showing that $G = |\mathrm{D}F|_\star$ and $\operatorname{lip}_{W_{2,\mathsf{d}_{\mathbb{M}}}} \boldsymbol{\kappa}^* F_n \to |\mathrm{D}F|_\star$ strongly in $L^2(\mathcal{P}_{2,\mathsf{d}_{\mathbb{M}}}(\mathbb{M}), \mathfrak{m})$. □

Remark 6.3. Arguing as in Section 4.1 it is immediate to see that the restriction of pCE_2 to the algebra $\mathrm{FC}_c^\infty(\mathcal{P}_2(\mathbb{M}))$ is a quadratic form and coincides with the pre-Dirichlet forms considered in [50,47,17,18]. If $\big(\mathsf{pCE}_2, \mathrm{FC}_c^\infty(\mathcal{P}_2(\mathbb{M}))\big)$ is closable then $\big(\mathsf{CE}_2, H^{1,2}(\mathcal{P}_2(\mathbb{M}), W_{2,\mathsf{d}_{\mathbb{M}}}, \mathfrak{m})\big)$ is a Dirichlet form which coincides with the smallest closed extension of $\big(\mathsf{pCE}_2, \mathrm{FC}_c^\infty(\mathcal{P}_2(\mathbb{M}))\big)$, it satisfies the so-called Rademacher property (see Proposition 5.2 and [17]) and it is quasi-regular (see Remark 2.19).

In particular it is possible to improve the result [17, Theorem 2.10]. Referring to the notation and the formula enumeration of that paper, one can immediately obtain that Lipschitz functions belong to $\mathcal{F}_0$ and the estimate (2.16) holds just assuming that $\big(\mathsf{pCE}_2, \mathrm{FC}_c^\infty(\mathcal{P}_2(\mathbb{M}))\big)$ is closable.

6.3. Wasserstein Sobolev space on Hilbert spaces

In this last section we will consider the case of a separable Hilbert space $(\mathbb{H}, |\cdot|)$; as usual, the space $\mathcal{P}_2(\mathbb{H})$ will be endowed with the Wasserstein distance $\mathbb{W}_2$ induced by the Hilbertian norm of $\mathbb{H}$ and we will assume that $\mathfrak{m}$ is a finite and positive Borel measure on $\mathcal{P}_2(\mathbb{H})$.

We select a complete orthonormal system $E := (e_n)_{n\in\mathbb{N}}$ and the collection of maps $\pi_d : \mathbb{H} \to \mathbb{R}^d$, $d \in \mathbb{N}$, given by

$$\pi^d(x) := \big(\langle x, e_1\rangle, \cdots, \langle x, e_d\rangle\big). \tag{6.10}$$

The adjoint map $\pi^{d*} : \mathbb{R}^d \to \mathbb{H}$ is given by

$$\pi^{d*}(y_1, \cdots, y_d) := \sum_{j=1}^d y_j\, e_j. \tag{6.11}$$

The map $\hat{\pi}^d := \pi^{d*} \circ \pi^d$ is the orthogonal projection of $\mathbb{H}$ onto span$\{e_1, \cdots, e_d\}$. We say that a function $\phi : \mathbb{H} \to \mathbb{R}$ belongs to $\mathrm{C}_b^1(\mathbb{H}, E)$ if it can be written as

$$\phi := \varphi \circ \pi^d \quad \text{for some } d \in \mathbb{N},\ \varphi \in \mathrm{C}_b^1(\mathbb{R}^d). \tag{6.12}$$

If $\phi \in \mathrm{C}_b^1(\mathbb{H}, E)$ then it belongs to $\mathrm{C}_b^1(\mathbb{H})$ and its gradient $\nabla\phi$ can be written as

$$\nabla\phi = \pi^{d*} \circ \nabla\varphi \circ \pi^d, \quad \nabla\phi(x) = \sum_{j=1}^{d} \partial_j \varphi(\pi^d(x)) e_j. \tag{6.13}$$

We then consider the algebra $\mathrm{FC}_b^1(\mathcal{P}_2(\mathbb{H}))$ generated by $\left\{\mathsf{L}_\phi : \phi \in \mathrm{C}_b^1(\mathbb{H}, E)\right\}$. For every $F \in \mathrm{FC}_b^1(\mathcal{P}_2(\mathbb{H}))$ we can find $N \in \mathbb{N}$, a polynomial $\psi : \mathbb{R}^N \to \mathbb{R}$ and functions $\phi_n \in \mathrm{C}_b^1(\mathbb{H}, E)$, $n = 1, \cdots, N$, such that

$$F(\mu) = (\psi \circ \mathsf{L}_{\boldsymbol{\phi}})(\mu). \tag{6.14}$$

As in (4.4) we can set

$$\mathrm{D}F(\mu, x) := \sum_{n=1}^{N} \partial_n \psi(\mathsf{L}_{\boldsymbol{\phi}}(\mu)) \nabla\phi_n(x). \tag{6.15}$$

It is also easy to check that a function F belongs to $\mathrm{FC}_b^1(\mathcal{P}_2(\mathbb{H}))$ if and only if there exists $d \in \mathbb{N}$ and $\tilde{F} \in \mathrm{FC}_b^1\left(\mathcal{P}_2(\mathbb{R}^d)\right)$ such that

$$F(\mu) = \tilde{F}(\pi^d_\sharp(\mu)) \quad \text{for every } \mu \in \mathcal{P}_2(\mathbb{H}), \tag{6.16}$$

so that

$$\mathrm{D}F(\mu, x) = \pi^{d*}\Big(\mathrm{D}\tilde{F}(\pi^d_\sharp\mu, \pi^d(x))\Big), \quad \|\mathrm{D}F[\mu]\|_\mu = \|\mathrm{D}\tilde{F}(\pi^d_\sharp\mu)\|_{\pi^d_\sharp\mu}. \tag{6.17}$$

By Proposition 4.9 and using (6.17) it is not difficult to check that

$$\|\mathrm{D}F[\mu]\|_\mu = \operatorname{lip} F(\mu) \quad \text{for every } \mu \in \mathcal{P}_2(\mathbb{H}). \tag{6.18}$$

Adapting in an obvious way the definitions in (4.14) and (4.15) to the Hilbertian framework, we have the following result.

Theorem 6.4. *$H^{1,2}(\mathcal{P}_2(\mathbb{H}), W_2, \mathfrak{m})$ is a Hilbert space and the algebra $\mathrm{FC}_b^1(\mathcal{P}_2(\mathbb{H}))$ is (strongly) dense: for every $F \in H^{1,2}(\mathcal{P}_2(\mathbb{H}), W_2, \mathfrak{m})$ there exists a sequence $F_n \in \mathrm{FC}_b^1(\mathcal{P}_2(\mathbb{H}))$, $n \in \mathbb{N}$ such that*

$$F_n \to F, \quad \operatorname{lip}(F_n) \to |\mathrm{D}F|_\star \quad \textit{strongly in } L^2(\mathcal{P}_2(\mathbb{H}), \mathfrak{m}). \tag{6.19}$$

Proof. Let us set $\mathscr{A} := \mathrm{FC}^1_b(\mathcal{P}_2(\mathbb{H}))$; we use Theorem 2.13 and we want to prove that for every $\nu \in \mathcal{P}_2(\mathbb{H})$ the function

$$F(\mu) := W_2(\nu, \mu) \quad \text{satisfies} \quad |\mathrm{D}F|_{\star,\mathscr{A}} \leq 1 \quad \mathfrak{m}\text{-a.e.}. \tag{6.20}$$

We split the proof in two steps.

Step 1: it is sufficient to prove that, for every $h \in \mathbb{N}$, the function $F_h : \mathcal{P}_2(\mathbb{H}) \to \mathbb{R}$

$$F_h(\mu) := W_2(\hat{\pi}^h_\sharp \nu, \hat{\pi}^h_\sharp \mu) \quad \text{satisfies} \quad |\mathrm{D}F_h|_{\star,\mathscr{A}} \leq 1 \quad \mathfrak{m}\text{-a.e.} \tag{6.21}$$

In fact, using the continuity property of the Wasserstein distance, it is clear that for every $\mu \in \mathcal{P}_2(\mathbb{H})$

$$\lim_{h\to\infty} F_h(\mu) = F(\mu), \tag{6.22}$$

so that it is enough to apply Theorem 2.3(1)-(3) to obtain (6.20).

Step 2: Let $h \in \mathbb{N}$ be fixed and let us denote by $W_{2,h}$ the Wasserstein distance on $\mathcal{P}_2(\mathbb{R}^h)$; it is easy to check that

$$W_{2,h}(\pi^h_\sharp \mu_0, \pi^h_\sharp \mu_1) = W_2(\hat{\pi}^h_\sharp \mu_0, \hat{\pi}^h_\sharp \mu_1) \quad \text{for every } \mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{H}).$$

Thus, if we define the function $\tilde{F}_h : \mathcal{P}_2(\mathbb{R}^h) \to \mathbb{R}$ as

$$\tilde{F}_h(\mu) := W_{2,h}(\pi^h_\sharp \nu, \mu)$$

we get that

$$F_h(\mu) = \tilde{F}_h(\pi^h_\sharp \mu).$$

We also introduce the measure $\mathfrak{m}_h$ on $\mathcal{P}_2(\mathbb{R}^h)$ given by the push-forward of $\mathfrak{m}$ through the (1-Lipschitz) map $P^h : \mathcal{P}_2(\mathbb{H}) \to \mathcal{P}_2(\mathbb{R}^h)$ defined as $P^h(\mu) := \pi^h_\sharp \mu$. By Theorem 4.10 applied to $H^{1,2}(\mathcal{P}_2(\mathbb{R}^h), W_{2,h}, \mathfrak{m}_h)$, we can find a sequence of cylinder functions $\tilde{F}_{h,n} \in \mathrm{FC}^1_b\left(\mathcal{P}_2(\mathbb{R}^h)\right)$, $n \in \mathbb{N}$, such that

$$\tilde{F}_{h,n} \to \tilde{F}_h \text{ in } \mathfrak{m}_h\text{-measure}, \tag{6.23}$$

$$\mathrm{lip}_{\mathcal{P}_2(\mathbb{R}^h)} \tilde{F}_{h,n} \to g_h \quad \text{in } L^2(\mathcal{P}_2(\mathbb{R}^h), \mathfrak{m}_h) \quad \text{with } g_h \leq 1 \ \mathfrak{m}_h\text{-a.e.} \tag{6.24}$$

We thus consider the functions $F_{h,n} \in \mathrm{FC}^1_b(\mathcal{P}_2(\mathbb{H}))$ defined as in (6.16) by

$$F_{h,n}(\mu) := \tilde{F}_{h,n}(\pi^h_\sharp \mu) = \tilde{F}_{h,n}(P^h(\mu)) \quad \text{for every } \mu \in \mathcal{P}_2(\mathbb{H}). \tag{6.25}$$

Since for every $\varepsilon > 0$

$$\mathfrak{m}\Big(\big\{\mu : |F_{h,n}(\mu) - F_h(\mu)| > \varepsilon\big\}\Big) = \mathfrak{m}\Big(\big\{\mu : |\tilde{F}_{h,n}(P^h(\mu)) - \tilde{F}_h(P^h(\mu))| > \varepsilon\big\}\Big)$$
$$= \mathfrak{m}_h\Big(\big\{\mu : |\tilde{F}_{h,n}(\mu) - \tilde{F}_h(\mu)| > \varepsilon\big\}\Big),$$

(6.23) yields that $F_{h,n} \to F_h$ in $\mathfrak{m}$-measure as $n \to \infty$.

On the other hand, (6.17) yields

$$\operatorname{lip} F_{h,n}(\mu) = \operatorname{lip}_{\mathcal{P}_2(\mathbb{R}^h)} \tilde{F}_{h,n}(P^h(\mu))$$

so that

$$\operatorname{lip} F_{h,n} \to g_h \circ P^h \quad \text{in } L^2(\mathcal{P}_2(\mathbb{H}), \mathfrak{m})$$

and $g_h \circ P^h \leq 1$ $\mathfrak{m}$-a.e. in $\mathcal{P}_2(\mathbb{H})$. By Theorem 2.3(1)-(3), we obtain (6.21), concluding the proof. □

Remark 6.5. We observe that the results in Sections 4.2 and 5.1 can be extended to $\mathcal{P}_2(\mathbb{M})$ and $\mathcal{P}_2(\mathbb{H})$ in an analogous way.

Data availability

No data was used for the research described in the article.

References

[1] Sergio Albeverio, Yuri Kondratiev, Michael Röckner, Analysis and geometry on configuration spaces, J. Funct. Anal. 154 (2) (1998) 444–500.

[2] Luigi Ambrosio, Massimo Fornasier, Marco Morandotti, Giuseppe Savaré, Spatially inhomogeneous evolutionary games, Commun. Pure Appl. Math. 74 (7) (2021) 1353–1402.

[3] Luigi Ambrosio, Wilfrid Gangbo, Hamiltonian ODEs in the Wasserstein space of probability measures, Commun. Pure Appl. Math. 61 (1) (2008) 18–53.

[4] Luigi Ambrosio, Nicola Gigli, Giuseppe Savaré, Gradient Flows in Metric Spaces and in the Space of Probability Measures, second edition, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2008.

[5] Luigi Ambrosio, Nicola Gigli, Giuseppe Savaré, Density of Lipschitz functions and equivalence of weak gradients in metric measure spaces, Rev. Mat. Iberoam. 29 (3) (2013) 969–996.

[6] Luigi Ambrosio, Nicola Gigli, Giuseppe Savaré, Density of Lipschitz functions and equivalence of weak gradients in metric measure spaces, Rev. Mat. Iberoam. (2013) 969–986.

[7] Luigi Ambrosio, Nicola Gigli, Giuseppe Savaré, Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below, Invent. Math. 195 (2) (2014) 289–391.

[8] Luigi Ambrosio, Nicola Gigli, Giuseppe Savaré, Metric measure spaces with Riemannian Ricci curvature bounded from below, Duke Math. J. (2014) 1405–1490.

[9] Luigi Ambrosio, Federico Stra, Dario Trevisan, Weak and strong convergence of derivations and stability of flows with respect to MGH convergence, J. Funct. Anal. 272 (3) (2017) 1182–1229.

[10] Hedi Attouch, Variational Convergence for Functions and Operators, Pitman (Advanced Publishing Program), Boston, MA, 1984.

[11] Erhan Bayraktar, Andrea Cosso, Huyên Pham, Randomized dynamic programming principle and Feynman-Kac representation for optimal control of McKean-Vlasov dynamics, Trans. Am. Math. Soc. 370 (3) (2016) 2115–2160.

[12] Anders Björn, Jana Björn, Nonlinear Potential Theory on Metric Spaces, EMS Tracts in Mathematics, vol. 17, European Mathematical Society (EMS), Zürich, 2011.
[13] Nicolas Bouleau, Francis Hirsch, Dirichlet Forms and Analysis on Wiener Space, De Gruyter Studies in Mathematics, vol. 14, De Gruyter, 1991.
[14] Pierre Cardaliaguet, Marco Cirant, Alessio Porretta, Splitting methods and short time existence for the master equations in mean field games, J. Eur. Math. Soc. (2022).
[15] Jeff Cheeger, Differentiability of Lipschitz functions on metric measure spaces, Geom. Funct. Anal. 9 (3) (1999) 428–517.
[16] Gianni Dal Maso, An Introduction to Γ-Convergence, Progress in Nonlinear Differential Equations and Their Applications, vol. 8, Birkhäuser, Boston, 1993.
[17] Lorenzo Dello Schiavo, A Rademacher-type theorem on L^2-Wasserstein spaces over closed Riemannian manifolds, J. Funct. Anal. 278 (6) (2020) 108397.
[18] Lorenzo Dello Schiavo, The Dirichlet–Ferguson diffusion on the space of probability measures over a closed Riemannian manifold, Ann. Probab. 50 (2) (2022) 591–648.
[19] Lorenzo Dello Schiavo, Kohei Suzuki, Rademacher-type theorems and Sobolev-to-Lipschitz properties for strongly local Dirichlet spaces, J. Funct. Anal. 281 (11) (2021) 109234.
[20] Simone Di Marino, Nicola Gigli, Enrico Pasqualetto, Elefterios Soultanis, Infinitesimal Hilbertianity of locally CAT(κ)-spaces, J. Geom. Anal. 31 (8) (2021) 7621–7685.
[21] Simone Di Marino, Danka Lučić, Enrico Pasqualetto, A short proof of the infinitesimal Hilbertianity of the weighted Euclidean space, C. R. Math. Acad. Sci. Paris 358 (7) (2020) 817–825.
[22] Ivar Ekeland, Roger Temam, Analyse Convexe et Problèmes Variationnels, Dunod, Gauthier-Villars, Paris, 1974.
[23] Wilfrid Gangbo, Adrian Tudorascu, On differentiability in the Wasserstein space and well-posedness for Hamilton–Jacobi equations, J. Math. Pures Appl. 125 (2019) 119–174.
[24] Nicola Gigli, On the differential structure of metric measure spaces and applications, Mem. Am. Math. Soc. 236 (1113) (2015), vi+91.
[25] Nicola Gigli, Nonsmooth differential geometry—an approach tailored for spaces with Ricci curvature bounded from below, Mem. Am. Math. Soc. 251 (1196) (2018), v+161.
[26] Nicola Gigli, Enrico Pasqualetto, Lectures on Nonsmooth Differential Geometry, SISSA Springer Series, vol. 2, Springer, Cham, 2020.
[27] Juha Heinonen, Pekka Koskela, Quasiconformal maps in metric spaces with controlled geometry, Acta Math. 181 (1) (1998) 1–61.
[28] Juha Heinonen, Pekka Koskela, Nageswari Shanmugalingam, Jeremy T. Tyson, Sobolev Spaces on Metric Measure Spaces, New Mathematical Monographs, vol. 27, Cambridge University Press, Cambridge, 2015.
[29] Pekka Koskela, Paul MacManus, Quasiconformal mappings and Sobolev spaces, Stud. Math. 131 (1) (1998) 1–17.
[30] Jean Louet, Optimal transport problems with gradient penalization, Theses, Université Paris Sud - Paris XI, July 2014.
[31] Danka Lučić, Enrico Pasqualetto, Infinitesimal Hilbertianity of weighted Riemannian manifolds, Can. Math. Bull. 63 (1) (2020) 118–140.
[32] Zhi-Ming Ma, Michael Röckner, Introduction to the Theory of (Non-symmetric) Dirichlet Forms, Springer, New York, 1992.
[33] Malliavin Marie-Paule, Paul Malliavin, An infinitesimally quasi-invariant measure on the group of diffeomorphisms of the circle, in: Special Functions (Okayama, 1990), ICM-90 Satell. Conf. Proc., Springer, Tokyo, 1991, pp. 234–244.
[34] Mattia Martini, Kolmogorov equations on the space of probability measures associated to the nonlinear filtering equation: the viscosity approach, arXiv:2202.11072, 2022.
[35] Mattia Martini, Kolmogorov equations on spaces of measures associated to nonlinear filtering processes, in: Stochastic Processes and Their Applications, vol. 161, 2023, pp. 385–423.
[36] John Nash, C^1 isometric imbeddings, Ann. Math. (2) 60 (1954) 383–396.
[37] Felix Otto, The geometry of dissipative evolution equations: the porous medium equation, Commun. Partial Differ. Equ. 26 (1–2) (2001) 101–174.
[38] Gabriel Peyré, Marco Cuturi, Computational optimal transport: with applications to data science, Found. Trends Mach. Learn. 11 (5–6) (2019) 355–607.
[39] Huyên Pham, Xiaoli Wei, Bellman equation and viscosity solutions for mean-field stochastic control problem, ESAIM Control Optim. Calc. Var. 24 (1) (2018) 437–461.
[40] R. Tyrrell Rockafellar, Convex Analysis, Princeton University Press, Princeton, 1970.

[41] Michael Röckner, Alexander Schied, Rademacher's theorem on configuration spaces and applications, J. Funct. Anal. 169 (2) (1999) 325–356.
[42] Giuseppe Savaré, Lipschitz approximation, capacity, and capacitary modulus in complete metric-measure spaces, in preparation.
[43] Giuseppe Savaré, Self-improvement of the Bakry-Émery condition and Wasserstein contraction of the heat flow in $\mathrm{RCD}(K,\infty)$ metric measure spaces, Discrete Contin. Dyn. Syst. 34 (4) (2014) 1641–1661.
[44] Giuseppe Savaré, Sobolev spaces in extended metric-measure spaces, in: New Trends on Analysis and Geometry in Metric Spaces, Springer, 2022, pp. 117–276.
[45] Nageswari Shanmugalingam, Newtonian spaces: an extension of Sobolev spaces to metric measure spaces, Rev. Mat. Iberoam. 16 (2) (2000) 243–279.
[46] Giacomo E. Sodini, The general class of Wasserstein Sobolev spaces: density of cylinder functions, reflexivity, uniform convexity and Clarkson's inequalities, arXiv:2212.10955, 2022.
[47] Karl-Theodor Sturm, Entropic measure on multidimensional spaces, in: Seminar on Stochastic Analysis, Random Fields and Applications VI, in: Progr. Probab., vol. 63, Birkhäuser/Springer, Basel AG, Basel, 2011, pp. 261–277.
[48] Cédric Villani, Topics in Optimal Transportation, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003.
[49] Cédric Villani, Optimal Transport. Old and New, Grundlehren der Mathematischen Wissenschaften, vol. 338, Springer-Verlag, Berlin, 2009.
[50] Max-K. von Renesse, Karl-Theodor Sturm, Entropic measure and Wasserstein diffusion, Ann. Probab. 37 (3) (2009) 1114–1191.