Advances in Mathematics 281 (2015) 1145–1177

Contents lists available at ScienceDirect

Advances in Mathematics

www.elsevier.com/locate/aim

Tensorization of Cheeger energies, the space $H^{1,1}$ and the area formula for graphs

Luigi Ambrosio *, Andrea Pinamonti, Gareth Speight

Scuola Normale Superiore, Pisa, Italy

ARTICLE INFO

Article history:
Received 20 July 2014
Received in revised form 18 May 2015
Accepted 7 June 2015
Available online 18 June 2015
Communicated by Camillo De Lellis

Keywords:
Metric measure spaces
Tensorization
Area formula
Sobolev spaces

ABSTRACT

First we study in detail the tensorization properties of weak gradients in metric measure spaces $(X, d, \mathfrak{m})$. Then, we compare potentially different notions of the Sobolev space $H^{1,1}(X, d, \mathfrak{m})$ and of weak gradient with exponent 1. Eventually we apply these results to compare the area functional $\int \sqrt{1 + |\nabla f|_w^2}\, d\mathfrak{m}$ with the perimeter of the subgraph of f, in the same spirit as the classical theory.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

The aim of this paper is the investigation of some fine analytic questions related to the theory of weak gradients in metric measure spaces. One of our motivations has been the study of the area functional

$$\mathcal{A}(f) = \int_X \sqrt{1 + |\nabla f|_w^2}\, d\mathfrak{m}$$

* Corresponding author.
E-mail addresses: luigi.ambrosio@sns.it (L. Ambrosio), andrea.pinamonti@sns.it (A. Pinamonti), gareth.speight@sns.it (G. Speight).

http://dx.doi.org/10.1016/j.aim.2015.06.004
0001-8708/© 2015 Elsevier Inc. All rights reserved.

in metric measure spaces $(X, d, \mathfrak{m})$, where $|\nabla f|_w$ denotes a suitable notion of weak gradient induced by the metric measure structure. In the metric measure setting, the functional $\mathcal{A}$ has been an object of investigation in the recent paper [17], see also [16] for generalizations. In the classical theory, it is well known that $\mathcal{A}$ corresponds to the surface area of the subgraph

$$E_f := \big\{(x,t) \in X \times \mathbb{R} : \ t < f(x)\big\}$$

of f. One of the aims of this paper is to investigate to what extent this correspondence holds also in metric measure spaces. The question makes sense, because by now there is a well established theory of BV functions and sets of finite perimeter in metric measure spaces, initiated in [21] (see also [1,20] for more recent developments). This theory provides, among other things, a canonical definition of surface area for sets of finite perimeter, see Section 2.5 for details.

Heuristically, the representation of $\mathcal{A}(f)$ as surface area of a subgraph seems to be closely related to a tensorization property of the weak gradients in the product metric measure structure, where one of the factors is simply the Euclidean real line endowed with the Lebesgue measure $\mathscr{L}^1$. Specifically, we mean the following: if $(X, d_X, \mathfrak{m}_X)$ and $(Y, d_Y, \mathfrak{m}_Y)$ are the factors, then for the metric measure structure

$$\big(X \times Y, \sqrt{d_X^2 + d_Y^2}, \mathfrak{m}_X \times \mathfrak{m}_Y\big)$$

the weak gradient (for the sake of simplicity we ignore in this introduction its potential dependence on the integrability exponent) should satisfy

$$|\nabla f(x,y)|_w^2 = |\nabla f(\cdot, y)|_w^2(x) + |\nabla f(x,\cdot)|_w^2(y) \qquad \text{for } \mathfrak{m}_X \times \mathfrak{m}_Y\text{-a.e. } (x,y) \in X \times Y. \tag{1.1}$$

There exist many properties that are easily seen to be stable under tensorization: compactness, properness, completeness, separability, length and geodesic properties, doubling, Poincaré inequalities, etc. As a matter of fact, although there are heuristic arguments suggesting that tensorization should always be true for the weak gradients (see Proposition 3.1, dealing with Euclidean spaces endowed with general norms), this question has not been much investigated so far. In [7] the tensorization property has been proved assuming a curvature lower bound on the factors $(X, d_X, \mathfrak{m}_X)$, $(Y, d_Y, \mathfrak{m}_Y)$ and the quadratic structure of the Cheeger energy. In Section 3 we refine the analysis of [7] and we prove three results, all independent of curvature assumptions (with the first and third also independent of doubling and Poincaré assumptions, the first and the second independent of quadraticity assumptions):

(a) The weak gradient $|\nabla f|_w$ in the product structure always coincides with the $L^2(X \times Y, \mathfrak{m}_X \times \mathfrak{m}_Y)$ relaxation of the squared "cartesian" slope

$$|\nabla f|_c^2(x,y) := |\nabla f(\cdot,y)|^2(x) + |\nabla f(x,\cdot)|^2(y),$$

starting from the class of locally Lipschitz functions (here $|\nabla g|$ is the slope, or the local Lipschitz constant, defined in (2.1)), see Theorem 3.2.

(b) Using (a), we show that if the factors are doubling and satisfy a $(1,2)$-Poincaré inequality, then (1.1) holds, see Theorem 3.4. Although we follow a different path for the proof, it might be that this result could also be obtained starting from Proposition 3.1 using Cheeger's differentiable structure and a suitable notion of product of Cheeger's charts. We also prove a more refined result, namely if $f \in L^2(X\times Y, \mathfrak{m}_X\times\mathfrak{m}_Y)$ satisfies $f(\cdot,y)\in H^{1,2}(X,d_X,\mathfrak{m}_X)$ for $\mathfrak{m}_Y$-a.e. $y\in Y$, $f(x,\cdot)\in H^{1,2}(Y,d_Y,\mathfrak{m}_Y)$ for $\mathfrak{m}_X$-a.e. $x\in X$ and

$$\int_X\int_Y |\nabla f(x,\cdot)|_w^2\, d\mathfrak{m}_Y d\mathfrak{m}_Y(x) + \int_Y\int_X |\nabla f(\cdot,y)|_w^2\, d\mathfrak{m}_X d\mathfrak{m}_Y(y) < \infty,$$

then $f\in H^{1,2}(X\times Y, \sqrt{d_X^2+d_Y^2}, \mathfrak{m}_X\times\mathfrak{m}_Y)$.

(c) If the factors are strongly asymptotically Hilbertian, meaning that $\int_X |\nabla f|^2\, d\mathfrak{m}_X$ and $\int_Y |\nabla g|^2\, d\mathfrak{m}_Y$ are quadratic forms on locally Lipschitz functions, then the tensorization property (1.1) holds, see Theorem 3.5.

Section 4 is devoted to the analysis of the Sobolev space $H^{1,1}(X,d,\mathfrak{m})$. We define $H^{1,1}(X,d,\mathfrak{m})$ in the spirit of the BV theory, considering those functions whose total variation $|Df|$ is absolutely continuous w.r.t. $\mathfrak{m}$, and which are $H^{1,1}$ along almost every curve (in the sense of [6]). Then, in Theorem 4.3 we compare this definition with the other ones already considered in the literature. An interesting fact is that, even though under doubling and $(1,1)$-Poincaré the spaces are the same, the associated notions of weak gradient do not coincide (even though they are comparable, see (4.9) and (4.10)), see [16] and Remark 4.4. More precisely, the gradient $|\nabla f|_{*,1}$ singled out by the BV theory can be approximated only in the sense of measures by slopes of Lipschitz functions (see (2.15)) and, by construction, it gives rise to a functional $f\mapsto \int_X |\nabla f|_{*,1}\, d\mathfrak{m}$ which is lower semicontinuous on locally Lipschitz functions. On the other hand, the gradient $|\nabla f|_{w,1}$ singled out by the theory of functions absolutely continuous along Mod_1-a.e. curve enjoys a stronger approximation property (see Proposition 4.5 for the simple proof) but, in general, it fails to give raise to a lower semicontinuous functional $f\mapsto \int |\nabla f|_{w,1}\, d\mathfrak{m}$.

Finally, in Section 5 we prove upper and lower bounds for the perimeter of E_f, even for $f\in BV(X,d,\mathfrak{m})$. For $f\in H^{1,1}(X,d,\mathfrak{m})$, under the doubling and $(1,1)$-Poincaré assumptions we get

$$\int_B \sqrt{1+|\nabla f|_{*,1}^2}\, d\mathfrak{m} \le P(E_f, B\times\mathbb{R}) \le \int_B \sqrt{1+|\nabla f|_{w,1}^2}\, d\mathfrak{m}$$

for any Borel set $B\subset X$, so that equalities hold if and only if the two notions of gradient coincide for the metric measure structure. Besides the case of asymptotically Hilbertian metric measure spaces with a curvature lower bound (see [13], which are the

so-called $RCD(K,\infty)$ spaces introduced in [7]), we are not aware of conditions ensuring the coincidence $\mathfrak{m}$-a.e. of the two gradients.

2. Notation and preliminary results

We assume that all metric measure spaces $(X,d,\mathfrak{m})$ are complete and separable, that $\mathfrak{m}$ is a finite Borel measure and that the support of $\mathfrak{m}$ is the whole space X.

We denote by $\mathscr{L}^d$ the Lebesgue measure in $\mathbb{R}^d$. We say that $f : X \to \mathbb{R}$ is *locally Lipschitz* if all its restrictions to bounded sets are Lipschitz. The *slope* (or the local Lipschitz constant) of $f : X \to \mathbb{R}$ is defined by

$$|\nabla f|(x) = \limsup_{y\to x} \frac{|f(y)-f(x)|}{d(y,x)}. \tag{2.1}$$

2.1. Doubling and Poincaré inequality

We say that $(X,d,\mathfrak{m})$ is doubling if there exists a constant $C>0$ satisfying

$$\mathfrak{m}(B_{2r}(x)) \le C\mathfrak{m}(B_r(x)) \qquad \forall x\in X,\ r>0. \tag{2.2}$$

Given $p\in[1,\infty)$, we say that the $(1,p)$-Poincarè inequality holds for Lipschitz functions if, for all open balls $B_r(x)\subset X$, one has

$$\fint_{B_r(x)} |f-f_{x,r}|\,d\mathfrak{m} \le cr\Big(\fint_{B_{\lambda r}(x)} |\nabla f|^p\,d\mathfrak{m}\Big)^{1/p} \qquad \forall f\in \mathrm{Lip}_{\mathrm{loc}}(X) \tag{2.3}$$

for suitable $c\ge 0$ and $\lambda\ge 1$ independent of $B_r(x)$ (here $\fint_A$ denotes the averaged integral on a Borel set A and $f_{x,r}=\fint_{B_r(x)} f$). In the sequel we say that a constant is structural if it depends only on the doubling constant in (2.2) and the constants c, λ and the exponent p in (2.3).

2.2. Maximal functions and approximate continuity

Given $g\in L^1(X,\mathfrak{m})$ nonnegative, we define

$$M_r g(x) := \sup\left\{ \fint_{B_s(x)} g\,d\mathfrak{m} :\ s\in(0,r)\right\} \qquad x\in X,\ r\in(0,\infty]. \tag{2.4}$$

More generally, if μ is a finite Borel measure, we define

$$M_r\mu(x) := \sup\left\{ \frac{\mu(B_s(x))}{\mathfrak{m}(B_s(x))} :\ s\in(0,r)\right\} \qquad x\in X,\ r\in(0,\infty], \tag{2.5}$$

so that $M_r(g\mathfrak{m}) = M_r g$.

For doubling metric measure spaces $(X, d, \mathfrak{m})$ we will need the weak L^1 estimate for the maximal function; we need it in the stronger form, for M_∞:

$$\mathfrak{m}\big(\{M_\infty \mu > \lambda\}\big) \le c\frac{\mu(X)}{\lambda} \qquad \forall \lambda > 0. \tag{2.6}$$

The constant c in (2.6) depends only on the doubling constant of $(X, d, \mathfrak{m})$. For absolutely continuous measures we will also use the more refined estimate

$$\lim_{\lambda\to\infty} \lambda \mathfrak{m}\big(\{M_\infty g > \lambda\}\big) = 0. \tag{2.7}$$

This asymptotic version follows by (2.6), taking the inclusion

$$\{M_\infty g > 2\lambda\} \subset \{M_\infty (g-\lambda)^+ > \lambda)\}$$

into account.

Using maximal functions, and under the doubling assumption, one can prove the existence of approximate limits of functions $f \in L^1(X, \mathfrak{m})$: the approximate limit $\tilde f(x)$ at x is defined by the property

$$\lim_{r\downarrow 0} \fint_{B_r(x)} |f(y) - \tilde f(x)|\, d\mathfrak{m}(y) = 0.$$

Points where the approximate limit exists are called approximate continuity points. It turns out that the approximate limit exists $\mathfrak{m}$-a.e. in X and that $\tilde f(x) = f(x)$ for $\mathfrak{m}$-a.e. $x \in X$ (notice that the first function is pointwise defined in its domain, while the latter belongs to a Lebesgue equivalence class).

2.3. Pseudo gradients and Cheeger energies

In this section we recall the basic facts of the theory of relaxed gradients. As we will see, even though the initial class of functions is a priori chosen to be the class of locally Lipschitz functions, it is sometimes technically useful to consider objects different from the slope in the relaxation process; for instance in Cheeger's paper [10] upper gradients were used; while in [5, Section 8.3] and [8] the so-called asymptotic Lipschitz constant was used, and proved to be technically useful. See also [14] for a closely related axiomatization.

Definition 2.1 *(Pseudo gradient).* We call a function G on $\mathrm{Lip}_{\mathrm{loc}}(X)$ with values into nonnegative Borel functions a *pseudo gradient* if the following properties hold:

(i) G is positively 1-homogeneous, i.e. $G(tf) = tG(f)$ for all $t \ge 0$;
(ii) G is convex;

(iii) for some constant C,

$$G(\chi u + (1-\chi)v) \le \chi G(u) + (1-\chi)G(v) + C\mathrm{Lip}(\chi)|u-v|$$

for all u, v, $\chi \in \mathrm{Lip}_{\mathrm{loc}}(X)$, $0 \le \chi \le 1$.

The main example of pseudo gradient is the slope in (2.1), other examples are for instance the one-sided slopes

$$|\nabla^+ f|(x) = \limsup_{y\to x} \frac{(f(y)-f(x))^+}{d(y,x)}, \qquad |\nabla^- f|(x) = \limsup_{y\to x} \frac{(f(y)-f(x))^-}{d(y,x)},$$

or the cartesian slope in product spaces, see (3.3) below.

Given a pseudo gradient G, the Cheeger energy $\mathsf{Ch}_G : L^2(X,\mathfrak{m}) \to [0,\infty]$ associated to G is defined by

$$\mathsf{Ch}_G(f) := \inf\left\{ \liminf_{h\to\infty} \int_X G^2(f_h)\, d\mathfrak{m} :\ f_h \in \mathrm{Lip}_{\mathrm{loc}}(X),\ f_h \to f \text{ in } L^2(X,\mathfrak{m}) \right\},$$

with the convention $\inf\emptyset = \infty$. We shall denote by Ch the "canonical" Cheeger energy associated to the slope and define

$$H^{1,2}(X,d,\mathfrak{m}) := \left\{ f \in L^2(X,\mathfrak{m}) :\ \mathsf{Ch}(f) < \infty \right\}.$$

We call $u \in L^2(X,\mathfrak{m})$ a G-relaxed slope of f if there exist Lipschitz functions f_n satisfying $f_n \to f$ strongly in $L^2(X,\mathfrak{m})$ and $G(f_n) \to v$ weakly in $L^2(X,\mathfrak{m})$, with $v \le u$ $\mathfrak{m}$-a.e. in X.

The following results collect the main facts about Cheeger energies and G-relaxed slopes; although the results in [10] and [6] are not stated in terms of pseudo-gradients, their proof extend with no change to this more general framework.

Theorem 2.2. *For any $f \in L^2(X,\mathfrak{m})$ and any pseudo gradient G the following properties hold.*

(i) *The collection of G-relaxed slopes of f is a convex closed set, possibly empty.*
(ii) *If the collection of G-relaxed slope is not empty, its element with minimal norm $|\nabla f|_{*,G}$ satisfies*

$$|\nabla f|_{*,G} \le u \ \mathfrak{m}\text{-a.e. in } X \text{ for any } G\text{-relaxed slope } u. \tag{2.8}$$

Furthermore, $|\nabla f|_{*,G}$ can be obtained as the strong $L^2(X,\mathfrak{m})$ limit of a sequence $G(f_n)$, with f_n locally Lipschitz and $f_n \to f$ in $L^2(X,\mathfrak{m})$.

(iii) $\sqrt{\mathsf{Ch}_G}$ *is a convex lower semicontinuous functional in* $L^2(X,\mathfrak{m})$*, positively 1-homogeneous and even if* G *is even. In addition,* $\mathsf{Ch}_G(f) < \infty$ *if and only if the collection of* G*-relaxed slopes is not empty. In this case,*

$$\mathsf{Ch}_G(f) = \int_X |\nabla f|^2_{*,G}\, d\mathfrak{m}.$$

We call $|\nabla f|_{*,G}$ the minimal relaxed slope and, as we did for Ch, denote by $|\nabla f|_*$, the "canonical" one associated to $G(f) = |\nabla f|$.

Recall that the subdifferential of a convex lower semicontinuous functional $\Phi : H \to (-\infty, +\infty]$ in a Hilbert space is defined, at any point u such that $\Phi(u) < \infty$, by

$$\partial\Phi(u) := \{\xi \in H :\ \Phi(v) \geq \Phi(u) + \langle \xi, v-u\rangle \ \forall v \in H\}.$$

Now, denoting by Δ_G the element with minimal norm in the subdifferential of Ch_G, one can use the inequality

$$|\nabla(f+\epsilon g)|_{*,G} \leq |\nabla f|_{*,G} + \epsilon |\nabla g|_{*,G} \qquad \mathfrak{m}\text{-a.e. in } X, \text{ for all } \epsilon > 0,$$

which is a simple consequence of the convexity and homogeneity of G, to prove the following integration by parts formula, see for instance [6, Proposition 4.15] (the second part of the proposition, dealing with the equality cases, uses also the chain rule [6, Proposition 4.8]).

Proposition 2.3. *For all* $u \in D(\mathsf{Ch}_G)$ *and all* $v \in D(\mathsf{Ch}_G)$ *such that* $\partial\mathsf{Ch}_G(v)$ *is not empty, one has*

$$-\int u\Delta_G v\, d\mathfrak{m} \leq \int |\nabla u|_{*,G} |\nabla v|_{*,G}\, d\mathfrak{m}, \tag{2.9}$$

with equality if $u = \phi(v)$ *with* $\phi : \mathbb{R} \to \mathbb{R}$ *Lipschitz, continuously differentiable and nonincreasing.*

In Section 5 we will need to consider the Cheeger energy and associated weak gradient for a measure finite on bounded sets; these are defined in the same way as for the finite case. Locality of the weak gradients allows us to generalize statements about pointwise behavior also to this case (arguing, for instance, as in [6, Lemma 4.11]). The assumption that $\mathfrak{m}$ is finite is not necessary in the following fundamental result obtained, among other things, in [10].

Theorem 2.4 *(Minimal relaxed slope coincides* $\mathfrak{m}$*-a.e. with the slope). Assume that* $(X,d,\mathfrak{m})$ *is doubling and that* $(1,2)$*-Poincaré holds for locally Lipschitz functions. Then*

$$|\nabla f| = |\nabla f|_* \qquad \mathfrak{m}\text{-a.e. in } X, \text{ for all } f \in \mathrm{Lip}_{\mathrm{loc}}(X) \cap H^{1,2}(X,d,\mathfrak{m}).$$

On $X = \mathbb{R}^d$ endowed with the distance induced by a norm $\|\cdot\|$ and with the Lebesgue measure $\mathfrak{m} = \mathscr{L}^d$, it is easily seen that (for f smooth or locally Lipschitz)

$$\mathsf{Ch}(f) = \int_X \|\nabla f\|_*^2 \, d\mathfrak{m}, \tag{2.10}$$

where $\|\cdot\|_*$ is the dual norm and ∇f is the canonical Euclidean gradient. Hence, Ch is a quadratic form if and only if $\|\cdot\|$ is induced by a scalar product. As in [12], this motivates two possible definition of spaces which are "Hilbertian on small scales".

Definition 2.5 *(Strongly asymptotically Hilbertian m.m.s.).* We say that $(X, d, \mathfrak{m})$ is strongly asymptotically Hilbertian if $f \mapsto \int_X |\nabla f|^2 \, d\mathfrak{m}$ is a quadratic form on $\mathrm{Lip}_{\mathrm{loc}}(X)$.

Definition 2.6 *(Asymptotically Hilbertian m.m.s.).* We say that $(X, d, \mathfrak{m})$ is asymptotically Hilbertian if Ch is a quadratic form on $L^2(X, \mathfrak{m})$.

Proposition 2.7. *Any strongly asymptotically Hilbertian space is asymptotically Hilbertian. If doubling and $(1, 2)$-Poincaré inequality hold, then the two properties are equivalent.*

Proof. The first statement is a simple consequence of the fact that the lower semicontinuous relaxation of a quadratic form is still a quadratic form. The second one follows by Theorem 2.4. □

2.4. Quadratic forms

Let H be a separable Hilbert space and let $Q : H \to [0, \infty]$ be a lower semicontinuous quadratic form. We shall denote by A the associated bilinear form on $D(Q) = \{Q < \infty\}$, namely

$$A(u, v) = \frac{1}{4}\big(Q(u + v) - Q(u - v)\big) \qquad u, \, v \in D(Q).$$

Assuming $D(Q)$ to be dense in H, we shall also denote by L the possibly unbounded operator whose domain $D(L)$ consists of all $u \in D(Q)$ satisfying

$$A(u, v) = \langle w, v\rangle \qquad \forall v \in D(Q) \tag{2.11}$$

for some $w \in H$. Since the density of $D(Q)$ in H ensures that w is uniquely determined by (2.11) we can set $w = Lu$, so that

$$A(u, v) = \langle Lu, v\rangle \qquad u \in D(L), \, v \in D(Q). \tag{2.12}$$

The following classical result is the spectral theorem for compact operators, stated with assumptions at the level of the quadratic form.

Theorem 2.8 *(Spectral theorem). Let $Q : H \to [0,\infty]$ be a lower semicontinuous quadratic form with dense domain. Assume the existence of $c > 0$ satisfying $Q(u) \geq c\|u\|^2$ for all $u \in H$, and that the sublevel sets $\{u \in H : Q(u) \leq M\}$ are compact in H for all $M \geq 0$. Then there exists a complete orthonormal basis of H made of eigenvectors of L.*

Proof. It is simple to check, by the minimization of $v \mapsto \frac{1}{2}Q(v) - \langle f, v\rangle$, that $L : D(L) \to H$ is onto. In addition,

$$c\|u\|^2 \leq Q(u) = A(u,u) = \langle Lu, u\rangle \leq \|u\|\|Lu\| \qquad \forall u \in D(L)$$

shows that L^{-1} is continuous, with $\|L^{-1}\| \leq 1/c$. In order to apply the classical spectral theorem for compact operators to L^{-1} (see for instance [9, Theorem VI.11]), we need only to show that any sequence $(u_n) \subset D(L)$ such that $(\|Lu_n\|)$ is bounded is relatively compact in H. To this aim, notice that the previous inequality yields $Q(u_n) \leq \|L(u_n)\|^2/c$, so that $Q(u_n)$ is bounded and (u_n) is relatively compact in H. □

2.5. BV functions

We recall the definition of BV introduced in [21] for locally compact spaces; this class has been further studied in [1], dropping the local compactness assumption, and it has been characterized by the behavior of the function along curves, see Theorem 2.12 below.

Definition 2.9 *(The space $BV(X,d,\mathfrak{m})$).* Let $f \in L^1(X,\mathfrak{m})$. We say that f belongs to $BV(X,d,\mathfrak{m})$ if there exist locally Lipschitz functions $f_n : X \to \mathbb{R}$ such that $f_n \to f$ in $L^1(X,\mathfrak{m})$ and $\limsup_n \int_X |\nabla f_n|\, d\mathfrak{m} < \infty$.

For $f \in BV(X,d,\mathfrak{m})$ and $A \subset X$ open, we may consider the set function

$$|Df|(A) := \inf\left\{ \liminf_{n\to\infty} \int_A |\nabla f_n|\, d\mathfrak{m} : \ f_n \in \mathrm{Lip}_{\mathrm{loc}}(A), \ \lim_{n\to\infty} \int_A |f_n - f|\, d\mathfrak{m} = 0 \right\}. \tag{2.13}$$

It can be proved that $A \mapsto |Df|(A)$ is the restriction to open sets of a finite Borel measure (it is in this proof that it turns out to be useful to consider *locally* Lipschitz functions in Definition 2.9 and (2.13)), that we still denote by $|Df|$. Notice also the elementary inequality (first proved on open sets, and then extended to Borel sets)

$$|D(f+g)| \leq |Df| + |Dg| \qquad \forall f,\, g \in BV(X,d,\mathfrak{m}). \tag{2.14}$$

By construction, we have the lower semicontinuity property

$$\liminf_{n\to\infty} |Df_n|(X) < \infty \quad \Rightarrow \quad f \in BV(X,d,\mathfrak{m}) \text{ and } |Df|(X) \leq \liminf_{n\to\infty} |Df_n|(X)$$

for $f_n \in BV(X,d,\mathfrak{m})$ convergent to f in $L^1(X,\mathfrak{m})$.

Remark 2.10 *(Sets of finite perimeter in locally finite m.m.s.).* When concerned with characteristic functions, we will apply these concepts also in m.m.s. whose reference measure $\mathfrak{m}$ is finite on bounded sets, i.e. we weaken the finiteness assumption on $\mathfrak{m}$ by requiring that $\mathfrak{m}(B) < \infty$ for any bounded Borel set B (specifically, in Section 5 we will consider the product of a finite m.m.s. with the real line).

We say that $E \subset X$ has finite perimeter in X if the characteristic function χ_E belongs to $BV(Y, d, \mathfrak{m})$ for any closed subset $Y \subset X$ with finite measure and $\sup_n |D\chi_E|(B_n(x_0)) < \infty$ for some (and thus all) $x_0 \in X$. By monotone approximation with open sets with finite measure, still $|D\chi_E|(B)$ is well defined for any open set $B \subset X$ and $B \mapsto |D\chi_E|(B)$ is a positive finite Borel measure in X. Furthermore, the monotonicity of the approximation gives that $E \mapsto |D\chi_E|(A)$ is lower semicontinuous with respect to local convergence in $\mathfrak{m}$-measure (i.e. $L^1_{\rm loc}$ convergence of the characteristic functions) for any open set $A \subset X$.

When dealing with sets of finite perimeter, we will use the traditional notation

$$P(E, B) := |D\chi_E|(B). \quad \blacksquare$$

By a diagonal argument in (2.13), it is clear that there exist f_n locally Lipschitz convergent to f in $L^1(X, \mathfrak{m})$ and satisfying $\limsup_n \int_X |\nabla f_n| \, d\mathfrak{m} \le |Df|(X)$. Since, by the very definition of $|Df|(A)$, it holds

$$\liminf_{n\to\infty} \int_A |\nabla f_n| \, d\mathfrak{m} \ge |Df|(A) \qquad \text{for any open set } A \subset X,$$

a well known criterion for weak convergence of measures gives the approximation property:

$$\forall f \in BV(X, d, \mathfrak{m}) \ \exists f_n \in \mathrm{Lip}_{\rm loc}(X) \quad \text{with} \quad |\nabla f_n|\mathfrak{m} \rightharpoonup |Df| \quad \text{in duality with } C_b(X). \tag{2.15}$$

We can now investigate the implications of doubling and $(1,1)$-Poincaré for Lipschitz functions. From (2.15) we immediately obtain the BV version of the Poincaré inequality in BV, namely

$$\fint_{B_r(x)} |f - f_{x,r}| \, d\mathfrak{m} \le cr|Df|\big(\overline{B}_{\lambda r}(x)\big) \qquad \forall f \in BV(X, d, \mathfrak{m}).$$

Possibly replacing λ by $\lambda' > \lambda$, we will keep using the traditional form

$$\fint_{B_r(x)} |f - f_{x,r}| \, d\mathfrak{m} \le cr|Df|\big(B_{\lambda r}(x)\big) \qquad \forall f \in BV(X, d, \mathfrak{m}), \tag{2.16}$$

stated with open sets. This estimate, when combined with the doubling property of $(X, d, \mathfrak{m})$, leads by standard arguments (see for instance [10,8]) to the following proposition.

Proposition 2.11. *Assume that $(X, d, \mathfrak{m})$ is doubling and that $(1,1)$-Poincaré holds for Lipschitz functions. Then, there exists a structural constant $c > 0$ satisfying*

$$|\tilde{f}(x) - \tilde{f}(y)| \le cd(x,y)\big(M_r|Df|(x) + M_r|Df|(y)\big) \tag{2.17}$$

whenever $f \in BV(X, d, \mathfrak{m})$, x and y are approximate continuity points of f, $r > 0$ and $d(x, y) < r/c$.

2.6. Equivalence of weak gradients and locality

We denote by $C([0,1]; X)$ the space of continuous maps from $[0,1]$ to X endowed with the standard Polish structure and denote by $e_t : C([0,1]; X) \to X$, $t \in [0,1]$, the evaluation maps at time t, namely $e_t(\gamma) = \gamma(t)$. We shall also denote by $(e_t)_\#\boldsymbol{\pi}$ the push-forward probability measure on X induced by $\boldsymbol{\pi} \in \mathscr{P}(C([0,1]; X))$, namely

$$\int_X \phi \, d(e_t)_\#\boldsymbol{\pi} = \int \phi(\gamma(t)) \, d\boldsymbol{\pi}(\gamma)$$

for any $\phi : X \to \mathbb{R}$ bounded Borel, or Borel nonnegative.

We say that $\boldsymbol{\pi} \in \mathscr{P}(C([0,1]; X))$ is a ∞-test plan if it is concentrated on Lipschitz curves, $\mathrm{Lip}(\gamma) \in L^\infty(C([0,1]; X), \boldsymbol{\pi})$ and the non-concentration property

$$(e_t)_\sharp\boldsymbol{\pi} \le C\mathfrak{m} \qquad \forall t \in [0,1] \tag{2.18}$$

holds for some constant $C \ge 0$. We shall denote by $C(\boldsymbol{\pi})$ the least constant satisfying (2.18).

We say that a Borel family $\Gamma \subset C([0,1]; X)$ is 1-negligible if $\boldsymbol{\pi}(\Gamma) = 0$ for any ∞-test plan $\boldsymbol{\pi}$. This notion is weaker than the non-parametric notion of Mod_1-negligible set. Recall that Γ is said to be Mod_1-negligible if for any $\varepsilon > 0$ there exists a Borel function $\rho : X \to [0, \infty]$ with $\int_X \rho \, d\mathfrak{m} < \varepsilon$ and $\int_\gamma \rho \ge 1$ for all $\rho \in \Gamma$. Since

$$\boldsymbol{\pi}(\Gamma) \le \int \int_\gamma \rho \, d\boldsymbol{\pi}(\gamma) \le \|\mathrm{Lip}(\gamma)\|_\infty \int_0^1 \int \rho(\gamma(t)) \, d\boldsymbol{\pi}(\gamma) \, dt \le \|\mathrm{Lip}(\gamma)\|_\infty C(\boldsymbol{\pi}) \int_X \rho \, d\mathfrak{m}$$

for any ∞-test plan $\boldsymbol{\pi}$, we obtain that Mod_1-negligible sets are 1-negligible. See [4] for a much more detailed comparison between notions of negligibility for families of curves, both parametric and non-parametric.

The next theorem is one of the main results of [1].

Theorem 2.12. *Let $f \in L^1(X, d, \mathfrak{m})$. Then $f \in BV(X, d, \mathfrak{m})$ if and only if there exists a finite Borel measure μ on X with the following property: for any ∞-test plan $\boldsymbol{\pi}$ one has:*

(i) $f \circ \gamma \in BV(0,1)$ *for* $\boldsymbol{\pi}$*-a.e.* γ*;*
(ii) $\int \gamma_\sharp |D(f\circ\gamma)|\, d\boldsymbol{\pi}(\gamma) \le C(\boldsymbol{\pi}) \|\mathrm{Lip}(\gamma)\|_\infty \mu$.

For all $f \in BV(X,d,\mathfrak{m})$*, the smallest measure* μ *satisfying (ii) is* $|Df|$.

Again, in (ii) by $\gamma_\sharp|D(f\circ\gamma)|$ we mean the push-forward measure defined by

$$\int_X \phi \, d\gamma_\sharp |D(f\circ\gamma)| = \int_0^1 \phi\circ\gamma \, d|D(f\circ\gamma)|.$$

Notice that (i) and (ii) are invariant in the Lebesgue equivalence class of f: indeed, because of (2.18), $\boldsymbol{\pi}$-a.e. curve γ hits a prescribed $\mathfrak{m}$-negligible set in an $\mathscr{L}^1$-negligible set of times, so that the integral in (ii) does not change if we replace f by $\hat f$, with $\mathfrak{m}(\{f \neq \hat f\}) = 0$.

Remark 2.13. Let us split $|D(f\circ\gamma)|$ into absolutely continuous $|D^a(f\circ\gamma)|$ and singular $|D^s(f\circ\gamma)|$ parts w.r.t. $\mathscr{L}^1$. Notice that without further assumptions on the metric measure structure, while we know that

$$\int \gamma_\sharp |D^a(f\circ\gamma)|\, d\boldsymbol{\pi}(\gamma) \ll \mathfrak{m}$$

(again because $\boldsymbol{\pi}$-a.e. curve γ hits an $\mathfrak{m}$-negligible set in an $\mathscr{L}^1$-negligible set of times), we can't say in general that

$$\int \gamma_\sharp |D^s(f\circ\gamma)|\, d\boldsymbol{\pi}(\gamma)$$

is singular w.r.t. $\mathfrak{m}$. See [1, Example 7.4] for an explicit example of a function $f \in BV(X,d,\mathfrak{m})$ (actually a characteristic function) with $|Df| \le C\mathfrak{m}$, but $|D(f\circ\gamma)|$ singular w.r.t. $\mathscr{L}^1$ for any curve γ.

3. Tensorization of metric measure spaces

In this section, as in Section 6 of [7], we consider two complete and separable m.m.s. $(X,d_X,\mathfrak{m}_X)$, $(Y,d_Y,\mathfrak{m}_Y)$ and their product $Z = X\times Y$, endowed with the product distance d satisfying $d^2 = d_X^2 + d_Y^2$ and the product measure $\mathfrak{m}$. We assume $\mathfrak{m}_X$ and $\mathfrak{m}_Y$ to be finite. We denote by Ch the Cheeger energy in $(Z,d,\mathfrak{m})$ and use the notation $f^x(y) = f(x,y)$, $f^y(x) = f(x,y)$.

In Section 6 of [7], it is proved, under curvature assumptions on the factors, that Ch is quadratic whenever Ch_X and Ch_Y are quadratic. In addition, defining

$$\Lambda := \Big\{ f \in L^2(Z,\mathfrak{m}) : \begin{cases} f^x \in H^{1,2}(Y,d_Y,\mathfrak{m}_Y) \text{ for } \mathfrak{m}_X\text{-a.e. } x \in X \\ f^y \in H^{1,2}(X,d_X,\mathfrak{m}_X) \text{ for } \mathfrak{m}_Y\text{-a.e. } y \in Y \end{cases} \Big\}, \tag{3.1}$$

we have the inclusion $H^{1,2}(Z,d,\mathfrak{m}) \subset \Lambda$ and any $f \in H^{1,2}(Z,d,\mathfrak{m})$ satisfies

$$|\nabla f|_*^2(x,y) = |\nabla f^y|_{*,X}^2(x) + |\nabla f^x|_{*,Y}^2(y) \qquad \text{for } \mathfrak{m}\text{-a.e. } (x,y) \in Z. \tag{3.2}$$

Notice that (3.2) provides immediately that Ch is a quadratic form whenever Ch_X and Ch_Y are quadratic forms, since

$$\mathsf{Ch}(f) = \int_Y \mathsf{Ch}_X(f^y)\, d\mathfrak{m}_Y(y) + \int_X \mathsf{Ch}_Y(f^x)\, d\mathfrak{m}_X(x).$$

Notice also that, as illustrated by Proposition 3.1 below, (3.2) might be true independently of quadraticity assumptions on Ch_X and Ch_Y, just by "duality" with the formula $d^2 = d_X^2 + d_Y^2$ defining d, but the full validity of this dual formula is presently an open problem (more precisely, the general validity of $\le$ in (3.2) is not known, while we will prove that the converse inequality always holds).

Proposition 3.1 *(An easy case of tensorization). Assume that X and Y are Euclidean spaces, with distances d_X and d_Y induced by norms $\|\cdot\|_X$ and $\|\cdot\|_Y$ respectively. If $\mathfrak{m}_X$ and $\mathfrak{m}_Y$ are the corresponding Lebesgue measures, then (3.2) holds.*

Proof. The norm $\|\cdot\|_Z$ in $Z = X \times Y$ corresponding to the distance $d_Z = \sqrt{d_X^2 + d_Y^2}$ obviously satisfies $\|(x,y)\|_Z^2 = \|x\|_X^2 + \|y\|_Y^2$. According to (2.10), we have to prove that

$$\|(x^*,y^*)\|_{*,Z}^2 = \|x^*\|_{*,X}^2 + \|y^*\|_{*,Y}^2 \qquad \forall x^* \in X^*,\ y^* \in Y^*,$$

where $\|\cdot\|_{*,Z}$ denotes the dual norm of $\|\cdot\|_Z$ and $\|\cdot\|_{*,X}$ and $\|\cdot\|_{*,Y}$ denote the dual norms in X and Y respectively. We can estimate

$$\begin{aligned}\langle (x^*,y^*),(x,y)\rangle &= \langle x^*,x\rangle + \langle y^*,y\rangle \\ &\le \|x^*\|_{*,X}\|x\|_X + \|y^*\|_{*,Y}\|y\|_Y \\ &\le \sqrt{\|x^*\|_{*,X}^2 + \|y^*\|_{*,Y}^2}\,\|(x,y)\|_Z.\end{aligned}$$

This proves that $\|(x^*,y^*)\|_{*,Z}^2 \le \|x^*\|_{*,X}^2 + \|y^*\|_{*,Y}^2$. On the other hand, if we choose nonzero vectors x, y such that $\langle x^*,x\rangle = \|x^*\|_{*,X}\|x\|_X$ and $\langle y^*,y\rangle = \|y^*\|_{*,Y}\|y\|_X$ we obtain that the first inequality is an equality. Then, replacing x by ax and y by by with a, $b \ge 0$ in such a way that

$$\|x\|_X = \|x^*\|_{*,X} \quad \text{and} \quad \|y\|_Y = \|y^*\|_{*,Y}$$

we retain the first two equalities and we obtain that also the last inequality is an equality. □

In Section 6 of [7], through a detailed analysis of the Hopf–Lax formula, inequality $\geq$ in (3.2) is proved independently of curvature assumptions, while curvature enters in the proof of the converse one (see [7, Lemma 6.15]).

In this section we will revise carefully the arguments of [7], obtaining some refinements independently of curvature assumptions.

We denote by $|\nabla f|_c : \mathrm{Lip}_{\mathrm{loc}}(Z) \to [0,\infty]$ the convex pseudo gradient

$$|\nabla f|_c(x,y) := \sqrt{|\nabla f^y|^2(x) + |\nabla f^x|^2(y)}, \tag{3.3}$$

corresponding to a "cartesian" slope. Notice that

$$|\nabla f|(x,y) \geq \max\{|\nabla f^x|(y), |\nabla f^y|(x)\} \geq \frac{1}{2}|\nabla f|_c(x,y). \tag{3.4}$$

We denote by Ch_c the Cheeger energy arising from the L^2 relaxation of $\int |\nabla f|_c^2\, d\mathfrak{m}$ restricted to locally Lipschitz functions, namely

$$\mathsf{Ch}_c(f) := \inf\left\{ \liminf_{n\to\infty} \int_Z |\nabla f_n|_c^2\, d\mathfrak{m} : \ f_n \in \mathrm{Lip}_{\mathrm{loc}}(Z),\ \int_Z |f_n - f|^2\, d\mathfrak{m} \to 0 \right\}.$$

As we discussed in Section 2.3, we have a minimal relaxed gradient associated to any $f \in D(\mathsf{Ch}_c)$, that we shall denote by $|\nabla f|_{*,c}$. It is also clear from (3.4) that $\mathsf{Ch}_c \leq 2\mathsf{Ch}$. Even though it seems difficult to establish a pointwise converse inequality

$$|\nabla f|(x,y) \leq C \max\{|\nabla f^x|(y), |\nabla f^y|(x)\} \leq C|\nabla f|_c$$

for some constant C, we are able to refine a bit the scheme of [7] to prove the following result, independently of doubling, quadraticity and Poincaré assumptions.

Theorem 3.2. $\mathsf{Ch}_c = \mathsf{Ch}$ *on* $L^2(Z,\mathfrak{m})$ *and*

$$|\nabla f|_* = |\nabla f|_{*,c} \quad \mathfrak{m}\text{-a.e. in } Z, \text{ for all } f \in H^{1,2}(Z,d,\mathfrak{m}). \tag{3.5}$$

Proof. We provide the proof of the inequality $\mathsf{Ch}_c \leq \mathsf{Ch}$ in the appendix, since it involves tools (Hopf–Lax formula, gradient flows) not directly connected to the rest of the paper. Let us prove the inequality $\mathsf{Ch} \leq \mathsf{Ch}_c$ and the corresponding inequality $\leq$ in (3.5). In the proof we shall use the dual point of view, i.e. we will work with the minimal weak upper gradient, which always coincides $\mathfrak{m}$-a.e. with the minimal relaxed gradient, see [6, Section 6] (this is the Sobolev counterpart of Theorem 2.12). By Lemma 6.14 in [7] (in turn based on the calculus Lemma 4.3.4 of [3]), $|\nabla f|_c$ is an upper gradient for locally Lipschitz functions, namely

$$\left|\frac{d}{dt}(f\circ\gamma)\right| \leq |\nabla f|_c \circ \gamma |\dot\gamma| \quad \mathscr{L}^1\text{-a.e. in } (0,1) \tag{3.6}$$

for all $\gamma \in AC([0,1];Z)$ and $f \in \mathrm{Lip}_{\mathrm{loc}}(Z)$. Recall now that for any $f \in D(\mathsf{Ch}_c)$ we can find locally Lipschitz functions f_n with $f_n \to f$ in $L^2(Z,\mathfrak{m})$ and $|\nabla f_n|_c \to |\nabla f|_{*,c}$ in $L^2(Z,\mathfrak{m})$. Then, the standard argument for the closure of weak upper gradients (even under weak $L^2(Z,\mathfrak{m})$ convergence, see [22, Lemma 4.11] or [6, Theorem 5.14]) provides the inclusion $D(\mathsf{Ch}) \supset D(\mathsf{Ch}_c)$ and the inequality

$$|\nabla f|_* \le |\nabla f|_{*,c} \qquad \mathfrak{m}\text{-a.e. in } Z. \qquad \square \tag{3.7}$$

It will be useful to compare Ch_c with the convex functional (actually a quadratic form, when Ch_X and Ch_Y are quadratic)

$$J(f) := \int_Z |\nabla f^y|^2_{*,X}(x) + |\nabla f^x|^2_{*,Y}(y)\, d\mathfrak{m}(x,y). \tag{3.8}$$

The functional J is well defined on the set Λ in (3.1), and set equal to ∞ on $L^2(Z,\mathfrak{m})\setminus\Lambda$. A standard Fubini argument then proves that J is $L^2(Z,\mathfrak{m})$-lower semicontinuous.

Lemma 3.3. *If $f \in L^2(Z,\mathfrak{m})$ is representable as $f_X f_Y$, with $f_X \in H^{1,2}(X,d_X,\mathfrak{m}_X)$ and $f_Y \in H^{1,2}(Y,d_Y,\mathfrak{m}_Y)$, then $\mathsf{Ch}_c(f) = J(f)$. Furthermore, if $-\xi_X \in \partial\mathsf{Ch}_X(f_X)$ and $-\xi_Y \in \partial\mathsf{Ch}_Y(f_Y)$, then*

$$-(f_Y\xi_X + f_X\xi_Y) \in \partial J\big(f_X f_Y\big). \tag{3.9}$$

Proof. The inequality $\mathsf{Ch}_c(f) \ge J(f)$, even for all functions f, is guaranteed by the definition of Ch_c, the inequality $\int_Z |\nabla f|^2_c\, d\mathfrak{m} \ge J(f)$ for $f \in \mathrm{Lip}_{\mathrm{loc}}(Z)$ and the lower semicontinuity of J. In order to prove the converse inequality $\mathsf{Ch}_c(f) \le J(f)$ on tensor products f, by a simple truncation argument we can assume with no loss of generality that f_X and f_Y are bounded functions. Let $f_X^n \in \mathrm{Lip}_{\mathrm{loc}}(X)$ with $|\nabla f_X^n| \to |\nabla f_X|_{*,X}$ in $L^2(X,\mathfrak{m}_X)$; analogously, let $f_Y^n \in \mathrm{Lip}_{\mathrm{loc}}(Y)$ with $|\nabla f_Y^n| \to |\nabla f_Y|_{*,Y}$ in $L^2(Y,\mathfrak{m}_Y)$. Since f_X and f_Y are bounded, we can also assume that (f_X^n) and (f_Y^n) are uniformly bounded. Since $f_X^n f_Y^n \in \mathrm{Lip}_{\mathrm{loc}}(Z)$ and

$$|\nabla(f_X^n f_Y^n)|^2_c = (f_X^n)^2|\nabla f_Y^n|^2 + (f_Y^n)^2|\nabla f_X^n|^2$$

we can easily conclude.

The verification of (3.9) is straightforward. $\square$

Now we can state the tensorization of doubling metric measure spaces satisfying a $(1,2)$-Poincaré inequality for locally Lipschitz functions. Note that, by a localization argument, the following theorem is also valid for measures finite on bounded sets.

Theorem 3.4. *Assume that $(X,d_X,\mathfrak{m}_X)$ and $(Y,d_Y,\mathfrak{m}_Y)$ are doubling and satisfy a $(1,2)$-Poincaré inequality for locally Lipschitz functions. Then the tensorization prop-*

erty (3.2) of weak gradients holds for all functions $f \in D(\mathsf{Ch})$. In addition, Ch coincides with the functional J defined in (3.8).

Our third result on the tensorization provides the quadratic property of Ch independently of curvature assumptions, but assuming the strong asymptotically Hilbertian property on the factors. As we discussed in Section 2.3, this assumption is equivalent to asymptotic Hilbertianity in the presence of doubling and $(1,2)$-Poincaré inequality for locally Lipschitz functions, which are already covered by Theorem 3.4. Again, by a localization argument, the following theorem also holds for measures finite on bounded sets.

Theorem 3.5. *If Ch_X, Ch_Y are quadratic, then (3.2) holds iff Ch is quadratic. In particular, if $(X, d_X, \mathfrak{m}_X)$ and $(Y, d_Y, \mathfrak{m}_Y)$ are strongly asymptotically Hilbertian, then Ch is a quadratic form and (3.2) holds.*

We devote the rest of the section to the proof of Theorem 3.4 and Theorem 3.5.

Lemma 3.6. *J is a convex and weakly lower semicontinuous functional on $L^2(Z,\mathfrak{m})$. In addition, $\mathsf{Ch}_c \geq J$ on $L^2(Z,\mathfrak{m})$ and $|\nabla f^x|^2_{*,Y}(y) + |\nabla f^y|^2_{*,X}(x) \leq |\nabla f|^2_{*,c}(x,y)$ for $\mathfrak{m}$-a.e. $(x,y) \in Z$.*

Proof. Convexity of J is trivial, so it suffices to show lower semicontinuity of J in the strong topology of $L^2(Z,\mathfrak{m})$. This and the last part of the statement are a simple application of Fubini's theorem and of the fact that any convergent sequence $g_n \to g$ in $L^2(Z,\mathfrak{m})$ can be refined to obtain a subsequence (n_k) satisfying $g^x_{n(k)} \to g^x$ in $L^2(Y,\mathfrak{m}_Y)$ for $\mathfrak{m}_X$-a.e. $x \in X$ and $g^y_{n(k)} \to g^y$ in $L^2(X,\mathfrak{m}_X)$ for $\mathfrak{m}_Y$-a.e. $y \in Y$. □

Proof of Theorem 3.4. Taking Lemma 3.6, (3.5) and the equality $\mathsf{Ch} = \mathsf{Ch}_c$ into account, to prove the tensorization property (3.2) for all functions $f \in D(\mathsf{Ch})$ we need only to show that $\mathsf{Ch}_c \leq J$ on $H^{1,2}(Z,d,\mathfrak{m})$ and that $D(J) \subset H^{1,2}(Z,d,\mathfrak{m})$. To prove the inequality we need to find, for all $f \in H^{1,2}(Z,d,\mathfrak{m})$ a sequence of functions $f_n \in \mathrm{Lip}_{\mathrm{loc}}(Z) \cap L^2(Z,\mathfrak{m})$ satisfying $f_n \to f$ in $L^2(Z,\mathfrak{m})$ and $J(f_n) \to J(f)$. Indeed, since Theorem 2.4 gives

$$\mathsf{Ch}_c(f_n) \leq \int_Z |\nabla f_n|^2_c \, d\mathfrak{m} = J(f_n),$$

the definition of Ch_c provides the result.

Since $(Z,d,\mathfrak{m})$ is a doubling space satisfying the $(1,2)$-Poincaré inequality on Lipschitz functions, in order to build f_n we can use the inequality (see for instance [8, Lemma 46])

$$|\tilde f(z) - \tilde f(z')| \leq C\big((M|\nabla f|^2_*)^{1/2}(z) + (M|\nabla f|^2_*)^{1/2}(z')\big) d(z,z')$$

valid at all approximate continuity of f, and the sets

$$E_n := \left\{ z \in Z : \ \max\{|\tilde f(z)|, (M|\nabla f|_*^2)^{1/2}(z)\} \le n \right\}.$$

By the MacShane lemma we can extend the restriction of $\tilde f$ to E_n to a Lipschitz function f_n with $\mathrm{Lip}(f_n) \le Cn$ and $|f_n| \le n$. Then, a standard argument based on maximal inequalities shows that $f_n \to f$ in $L^2(Z,\mathfrak{m})$, $\mathfrak{m}(Z\setminus E_n) \downarrow 0$ and $|\nabla f_n|^2$ are equi-integrable in $L^1(Z,\mathfrak{m})$. Using the locality of weak gradients in the base spaces, this immediately yields $J(f_n) \to J(f)$ and completes the proof of the inequality $\mathsf{Ch}_c \le J$ on $H^{1,2}(Z,d,\mathfrak{m})$.

Now, let us prove the more delicate inclusion $D(J) \subset H^{1,2}(Z,d,\mathfrak{m})$. In this proof it will be useful to assume that the distances d_X and d_Y are geodesic; this is not restrictive, since the geodesic distances associated to d_X and d_Y are (because of doubling and Poincaré) equivalent to the original distances and even induce the same weak gradients (see for instance [19, Corollary 7.3.17], but note that we need only equivalence of distances). We will prove that any $h \in D(J)$ whose support is contained in a product of balls $\overline{B}_R(x_0) \times \overline{B}_R(y_0)$ belongs to $H^{1,2}(Z,d,\mathfrak{m})$. The general case can be easily achieved by approximation.

Let us proceed now to the proof of the inclusion $D(J) \subset H^{1,2}(Z,d,\mathfrak{m})$. Notice that, by the very definition of J, for any $h \in D(J)$ one has not only $h^x \in H^{1,2}(Y,d_Y,\mathfrak{m}_Y)$ for $\mathfrak{m}_X$-a.e. $x \in X$ and $h^y \in H^{1,2}(X,d_X,\mathfrak{m}_X)$ for $\mathfrak{m}_Y$-a.e. $y \in Y$, but also $\mathsf{Ch}_Y(h^x) \in L^1(X,\mathfrak{m}_X)$ and $\mathsf{Ch}_X(h^y) \in L^1(Y,\mathfrak{m}_Y)$. Thanks to Cheeger's theory (see also Theorem 40 of the recent paper [8] for a different construction based on difference quotients and Γ-convergence which uses only the doubling property of the metric space) we can find L^2-lower semicontinuous quadratic forms $\mathcal{E}_X$ in $L^2(X,\mathfrak{m}_X)$ and $\mathcal{E}_Y$ in $L^2(Y,\mathfrak{m}_Y)$ equivalent to the Cheeger energies, namely

$$c_1\,\mathsf{Ch}_X \le \mathcal{E}_X \le \frac{1}{c_1}\mathsf{Ch}_X, \qquad c_2\,\mathsf{Ch}_Y \le \mathcal{E}_Y \le \frac{1}{c_2}\mathsf{Ch}_Y \tag{3.10}$$

for suitable positive structural constants c_1, c_2. The quadratic forms

$$\mathcal{E}_{X,1}(f) := \mathcal{E}_X(f) + \int_X f^2\,d\mathfrak{m}_X, \qquad \mathcal{E}_{Y,1}(g) := \mathcal{E}_Y(g) + \int_Y g^2\,d\mathfrak{m}_Y$$

induce Hilbertian structures in $H^{1,2}(X,d_X,\mathfrak{m}_X)$ and $H^{1,2}(Y,d_Y,\mathfrak{m}_Y)$ respectively. We also denote by J_1 the quadratic form

$$J_1(h) := \int_X \mathcal{E}_{Y,1}(f^x)\,d\mathfrak{m}_X(x) + \int_Y \mathcal{E}_{X,1}(f^y)\,d\mathfrak{m}_Y(y)$$

and notice that $D(J_1) = D(J)$, because of (3.10).

Fix $x_0 \in X$, $y_0 \in Y$ and $R > 0$. Fix the notation:

$$D(x_0,R) = \{f\colon X \to \mathbb{R}\colon \mathrm{supp}(f) \subset \overline{B}(x_0,R)\},$$
$$D(y_0,R) = \{g\colon Y \to \mathbb{R}\colon \mathrm{supp}(g) \subset \overline{B}(y_0,R)\}.$$

Then, for every $M > 0$, the sets

$$\{f \in D(x_0, R) : \mathcal{E}_{X,1}(f) \leq M\},$$
$$\{g \in D(y_0, R) : \mathcal{E}_{Y,1}(g) \leq M\},$$

are compact in $L^2(X, \mathfrak{m}_X)$ and $L^2(Y, \mathfrak{m}_Y)$ respectively. Indeed, [15, Theorem 8.1] gives this Rellich–Kondrachov type result for domains in which the measure is doubling and also a global Poincaré inequality result holds (see [15, Inequality (46)]). However, it is easy to see that balls in geodesic spaces are John domains (see Section 9.1 in [15] for the definition) and hence [15, Theorem 9.7] provides the required global Poincaré inequality. Let

$$H_X = L^2(X, \mathfrak{m}_X) \cap D(x_0, R),$$
$$H_Y = L^2(Y, \mathfrak{m}_Y) \cap D(y_0, R).$$

Notice that $H^{1,2}(X, d_X, \mathfrak{m}_X) \cap D(x_0, R)$ is dense in H_X with respect to the $L^2(X, m_X)$ norm and similarly for $H^{1,2}(Y, d_Y, \mathfrak{m}_Y) \cap D(y_0, R)$. Hence, by applying Theorem 2.8 to $\mathcal{E}_{X,1}$ in

$$H_X = L^2(X, \mathfrak{m}_X) \cap D(x_0, R)$$

and to $\mathcal{E}_{Y,1}$ in

$$H_Y = L^2(Y, \mathfrak{m}_Y) \cap D(y_0, R),$$

we can find complete orthonormal bases (f_n) of H_X and (g_n) of H_Y made of eigenvectors of the operators L_X and L_Y associated to these quadratic forms. In particular:

(a) the linear semigroup P_t^X associated to $\mathcal{E}_{X,1}$ on H_X leaves all 1-dimensional vector spaces $\mathbb{R}f_n$ invariant, and an analogous property holds for P_t^Y;
(b) the functions f_n are mutually orthogonal for the scalar product induced by $\mathcal{E}_{X,1}$ and an analogous property holds for g_m.

Now, it is easily seen that the vector space E spanned by the tensor products $f_n g_m$, where $f_n \in H^{1,2}(X, d_X, m_X) \cap D(x_0, R)$ and $g_m \in H^{1,2}(Y, d_Y, m_Y) \cap D(y_0, R)$, is dense in the space

$$\{h \in D(J) : \mathrm{supp}(h) \subset \overline{B}(x_0, R) \times \overline{B}(y_0, R)\}$$

w.r.t. the norm induced by J_1. Indeed, a simple functional analytic argument is based on the fact that the linear semigroup P_t induced by the quadratic form J_1 leaves E invariant, thanks to (a) and

$$P_t(f_n g_m) = (P_t^X f_n)(P_t^Y g_m) \qquad t \geq 0,$$

and on the fact that E is dense in $L^2(Z, \mathfrak{m})$ norm; these two facts (see for instance [7, Proposition 4.9]) imply density in the stronger norm induced by J_1.

In order to conclude the proof we use first the convexity and 1-homogeneity of $\sqrt{\mathsf{Ch}}$ and then Lemma 3.3 to get

$$\begin{aligned}\sqrt{\mathsf{Ch}}(h) &\leq \sum_{n,m} |\lambda_{n,m}| \sqrt{\mathsf{Ch}}(f_n g_m) = \sum_{n,m} |\lambda_{n,m}| \sqrt{J}(f_n g_m) \\ &\leq c \sum_{n,m} |\lambda_{n,m}| \sqrt{J_1}(f_n g_m) \leq c \left(\sum_{n,m} \lambda_{n,m}^2 J_1(f_n g_m) \right)^{1/2} \\ &= c \sqrt{J_1}(h)\end{aligned}$$

for any $h \in E$ written as a sum $\sum_{n,m} \lambda_{n,m} f_n g_m$, with only finitely many $\lambda_{n,m}$ nonzero, where in the last equality we used the fact that $(f_n g_m)$ are mutually orthogonal w.r.t. the scalar product induced by J_1 (as a consequence of (b)). Finally we can use the density of E and the lower semicontinuity of Ch to conclude that $\mathsf{Ch} \leq c^2 J_1$ on the set

$$\{h \in D(J) : \mathrm{supp}(h) \subset \overline{B}(x_0, R) \times \overline{B}(y_0, R)\}.$$

By letting $R \to \infty$ we deduce Ch is finite on $D(J)$ so that $D(J) \subset D(\mathsf{Ch})$. □

A simple consequence of (3.9) is that, independently of quadraticity assumptions on Ch_X and Ch_Y, the semigroup P_t associated to J (namely the gradient flow of J) acts on tensor products $f_X f_Y$ as follows

$$P_t(f_X f_Y) = (P_t^X f_X)(P_t^Y f_Y), \tag{3.11}$$

where P_t^X and P_t^Y are the semigroups on the factors. Indeed, (3.9) with $\xi_X = -\Delta P_t^X f_X$ and $\xi_Y = -\Delta_Y P_t^Y f_Y$ and the Leibniz rule give that

$$-\frac{d}{dt}(P_t^X f_X)(P_t^Y f_Y) \in \partial J\big((P_t^X f_X)(P_t^Y f_Y)\big),$$

which is the subdifferential formulation of the gradient flow.

Proof of Theorem 3.5. Assuming that Ch_X, Ch_Y are quadratic, it is clear that (3.2) implies that Ch are quadratic, so let us prove the converse implication from Ch quadratic to (3.2). By Lemma 3.6 and the equality $\mathsf{Ch}_c = \mathsf{Ch}$ provided by Theorem 3.2, we know that the two lower semicontinuous quadratic forms

$$Q_1(f) := \mathsf{Ch}(f), \qquad Q_2(f) := J(f)$$

satisfy $Q_1 \geq Q_2$, hence $Q = Q_1 - Q_2$ is a nonnegative quadratic form which, by Lemma 3.3, vanishes on tensor products. By the inequality $Q(u+v) \leq 2Q(u) + 2Q(v)$, valid for nonnegative quadratic forms, we obtain that Q vanishes on the vector space spanned by tensor products. Now, as in the proof of Theorem 3.4, it is easily seen that the vector space E spanned by the tensor products is dense w.r.t. the norm $\|\cdot\|_{Q_2} := \sqrt{\|\cdot\|_2^2 + Q_2^2}$. Indeed, the linear semigroup P_t induced by the quadratic form Q_2 leaves E invariant (because of (3.11)) and E is dense in $L^2(Z, \mathfrak{m})$ norm; these two facts (see for instance [7, Proposition 4.9]) imply density in the stronger norm induced by $\|\cdot\|_{Q_2}$. Therefore for any $f \in L^2(Z, \mathfrak{m})$ we can find $f_n \in E$ convergent to f in $L^2(Z, \mathfrak{m})$ with $Q_2(f_n) \to Q_2(f)$. Using lower semicontinuity of Q_1 we get

$$Q_1(f) \leq \liminf_{n\to\infty} Q_1(f_n) = \liminf_{n\to\infty} Q_2(f_n) = Q_2(f).$$

Taking into account the inequality $|\nabla f^x|^2_{*,Y} + |\nabla f^y|^2_{*,X} \leq |\nabla f|^2_c(x,y)$ for $\mathfrak{m}$-a.e. $(x,y) \in Z$ provided by the combination of Theorem 3.2 and Lemma 3.6, we obtain (3.2).

To prove the last statement of the theorem, assume that the factors are strongly asymptotically Hilbertian. Notice that

$$f \mapsto \int_Z |\nabla f|^2_c \, d\mathfrak{m} = \int_X \left(\int_Y |\nabla f^x|^2 \, d\mathfrak{m}_Y \right) d\mathfrak{m}_X(x) + \int_Y \left(\int_X |\nabla f^y|^2 \, d\mathfrak{m}_X \right) d\mathfrak{m}_Y(y)$$

is a quadratic form on $\mathrm{Lip}_{\mathrm{loc}}(Z)$. Since the lower semicontinuous relaxation of a quadratic form is still a quadratic form, it follows that $\mathsf{Ch} = \mathsf{Ch}_c$ is a quadratic form. Therefore we can apply the first part of the statement to obtain (3.2). □

4. The Sobolev space $H^{1,1}(X, d, \mathfrak{m})$

In this section we investigate in more detail the possible definitions of the space $H^{1,1}$ described in [1]. We adopt the definition leading to the larger space and study a few structural properties. Then, assuming doubling and $(1,1)$-Poincarè, we compare with more restrictive definitions. The results of this section will be applied in the next section, dealing with $H^{1,1}$ functions in $X \times \mathbb{R}$. In this section we assume that $\mathfrak{m}$ is finite on bounded sets.

Definition 4.1 *(The space $H^{1,1}(X,d,\mathfrak{m})$).* We denote by $H^{1,1}(X,d,\mathfrak{m})$ the subspace of all $f \in BV(X,d,\mathfrak{m})$ satisfying:

(i) $f \circ \gamma \in H^{1,1}(0,1)$ for 1-almost every $\gamma \in C([0,1]; X)$;
(ii) $|Df| \ll \mathfrak{m}$.

By analogy with the Sobolev case, we denote by $|\nabla f|_{*,1}$ the density of $|Df|$ w.r.t. $\mathfrak{m}$.

In view of the example mentioned in Remark 2.13, we know that (ii) would not be sufficient to provide Sobolev regularity of f along 1-almost every curve (not even if $|Df|$ has a bounded density), a natural requirement for Sobolev functions, also in view of the connection with other definitions. In the presence of doubling and $(1,1)$-Poincaré inequality for locally Lipschitz functions, however, (ii) is sufficient, see the proof of the implication from (iii) to (i) in Theorem 4.3 which only uses $|Df| \ll \mathfrak{m}$ (see also [16, Theorem 4.6]).

We use the notation $|\nabla f|_{*,1}$ because, at this level of generality, we expect that weak gradients depend on the integrability exponent, even for Lipschitz functions, see [11] for examples compatible even with the doubling assumption. Notice that the obvious inequality $|Df| \le |\nabla f|\mathfrak{m}$ on locally Lipschitz functions gives $|\nabla f|_{*,1} \le |\nabla f|$ $\mathfrak{m}$-a.e. in X for all $f \in \mathrm{Lip}_{\mathrm{loc}}(X)$, while additional assumptions are needed to reverse this inequality (see (4.5) below).

In the next proposition we are able to prove the locality of $|\nabla f|_{*,1}$. It was difficult for us to prove this fact using only (2.15) (because, unlike the Sobolev case, the convergence of slopes is weak and not strong), therefore we adopt the dual point of view provided by Theorem 2.12.

Proposition 4.2 *(Locality of $|\nabla f|_{*,1}$). Let $f \in H^{1,1}(X,d,\mathfrak{m})$. Then*

$$|\nabla f|_{*,1} = 0 \qquad \mathfrak{m}\text{-a.e. in } \{f = 0\}. \tag{4.1}$$

In particular, if f, g belong to $H^{1,1}(X,d,\mathfrak{m})$, then

$$|\nabla f|_{*,1} = |\nabla g|_{*,1} \qquad \mathfrak{m}\text{-a.e. in } \{f = g\}. \tag{4.2}$$

Proof. It is well known that for a Sobolev function $g : (0,1) \to \mathbb{R}$ the property holds, namely $g' = 0$ $\mathscr{L}^1$-a.e. on $\{g = 0\}$. Given a ∞-test plan $\boldsymbol{\pi}$, from part (i) of Definition 4.1 it follows that

$$(f \circ \gamma)' = 0 \quad \mathscr{L}^1\text{-a.e. on } \gamma^{-1}(\{f = 0\}), \text{ for } \boldsymbol{\pi}\text{-a.e. } \gamma.$$

Since $|D(f \circ \gamma)| = |(f \circ \gamma)'|\mathscr{L}^1$, the definition of push-forward gives $\gamma_\sharp|D(f \circ \gamma)|(\{f = 0\}) = 0$ for $\boldsymbol{\pi}$-a.e. γ. By integration w.r.t. $\boldsymbol{\pi}$, since $\boldsymbol{\pi}$ is arbitrary we get that the measure

$$\mu := \chi_{\{f \neq 0\}}|\nabla f|_{*,1}\mathfrak{m} = \chi_{\{f \neq 0\}}|Df|$$

still satisfies (ii) of Theorem 2.12. Then, the minimality of $|Df| = |\nabla f|_{*,1}\mathfrak{m}$ gives (4.1).

It follows from the inequalities (derived from (2.14)) that

$$-|\nabla(f-g)|_{*,1} + |\nabla g|_{*,1} \le |\nabla f|_{*,1} \le |\nabla g|_{*,1} + |\nabla(f-g)|_{*,1} \qquad \mathfrak{m}\text{-a.e. in } X.$$

Using (4.1) we obtain (4.2). □

If we assume that the $(1,1)$-Poincaré inequality for Lipschitz functions we can specialize (2.16) to $H^{1,1}(X,d,\mathfrak{m})$, obtaining

$$\fint_{B_r(x)} |f - f_{x,r}|\,d\mathfrak{m} \le cr \fint_{B_{\lambda r}(x)} |\nabla f|_{*,1}\,d\mathfrak{m} \qquad \forall f \in H^{1,1}(X,d,\mathfrak{m}). \tag{4.3}$$

Analogously, (2.17) gives

$$|\tilde f(x) - \tilde f(y)| \le cd(x,y)\big(M_r|\nabla f|_{*,1}(x) + M_r|\nabla f|_{*,1}(y)\big). \tag{4.4}$$

When f is locally Lipschitz we can use (4.4) at approximate continuity points of $M_r|\nabla f_{*,1}|$ (see for instance the argument in [8]) to get

$$|\nabla f| \le cM_r(|\nabla f|_{*,1}) \qquad \mathfrak{m}\text{-a.e. in } X, \text{ for all } f \in \mathrm{Lip}_{\mathrm{loc}}(X).$$

Letting $r \downarrow 0$ eventually gives, for some structural constant c,

$$|\nabla f| \le c|\nabla f|_{*,1} \qquad \mathfrak{m}\text{-a.e. in } X, \text{ for all } f \in \mathrm{Lip}_{\mathrm{loc}}(X). \tag{4.5}$$

Now, we can combine the "local" Lipschitz estimate (4.4), with the weak L^1 and asymptotic estimates (2.6), (2.7) to get the following Lusin type approximation result of $H^{1,1}$ functions by Lipschitz functions; this way, we get the equivalence with other, a priori stronger, definitions.

Theorem 4.3 *(Equivalent definitions of $H^{1,1}(X,d,\mathfrak{m})$). Let $f \in L^1(X,\mathfrak{m})$ and let us consider the following conditions:*

(i) *there exist $f_n \in \mathrm{Lip}(X) \cap L^1(X,\mathfrak{m})$ with $f_n \to f$ in $L^1(X,\mathfrak{m})$, $\mathfrak{m}(\{f \ne f_n\}) \to 0$ and $|\nabla f_n|$ equi-integrable;*

(ii) *there exist a Borel function $g : X \to [0,\infty)$ and a representative $\hat f$ of f such that $\int_X g\,d\mathfrak{m} < \infty$ and*

$$|\hat f(\gamma(1)) - \hat f(\gamma(0))| \le \int_\gamma g \qquad \text{for } \mathrm{Mod}_1\text{-a.e. curve } \gamma; \tag{4.6}$$

(iii) $f \in H^{1,1}(X,d,\mathfrak{m})$.

Then (i) $\Rightarrow$ (ii), with g limit point of $|\nabla f_n|$ in the weak $L^1(X,\mathfrak{m})$ topology, and (ii) $\Rightarrow$ (iii), with

$$|\nabla f|_{*,1} \le g \qquad \mathfrak{m}\text{-a.e. in } X. \tag{4.7}$$

If $(X,d,\mathfrak{m})$ is doubling and the $(1,1)$-Poincaré inequality for locally Lipschitz functions holds, then $f \in BV(X,d,\mathfrak{m})$ and $|Df| \ll \mathfrak{m}$ imply (i). In particular (iii) $\Rightarrow$ (i).

Proof. (i) ⇒ (ii). See for instance [22], [18, Theorem 10.5] and the recent monograph [19] for the construction of the representative $\hat f$. The main point is to find Lipschitz functions $\tilde f_n$ convergent to f $\mathfrak m$-a.e. in X and upper gradients g_n of f strongly convergent to g in $L^1(X,\mathfrak m)$, using also the fact that this implies $\int_\gamma g_n \to \int_\gamma g$ for Mod_1-a.e. γ. The functions $\tilde f_n$ are finite convex combinations of f_n, where the coefficients of the convex combinations are chosen in such a way that the corresponding convex combination g_n of $|\nabla f_n|$, which are upper gradients of $\tilde f_n$, are strongly convergent.

(ii) ⇒ (iii). Since families of subcurves of a Mod_1-negligible set of curves are still Mod_1-negligible, the argument in [22] shows that $\hat f$ is absolutely continuous (and not only $H^{1,1}$) along Mod_1-a.e. curve, with

$$\Big|\frac{d}{dt} f\circ\gamma(t)\Big| \le g(\gamma(t))|\dot\gamma(t)| \qquad \text{for } \mathscr{L}^1\text{-a.e. } t\in(0,1),\ \text{for } \mathrm{Mod}_1\text{-a.e. } \gamma. \tag{4.8}$$

Since we already observed that Mod_1-negligible sets are 1-negligible, we obtain (i) of Definition 4.1. In connection with condition (ii) of Definition 4.1, we can use (4.8) to show that the measure $\mu = g\mathfrak m$ satisfies

$$\int \gamma_\sharp|D(f\circ\gamma)|\,d\boldsymbol\pi(\gamma) \le C(\boldsymbol\pi)\|\mathrm{Lip}(\gamma)\|_\infty\mu$$

for any ∞-test plan $\boldsymbol\pi$. The minimality property of $|Df|$ stated in Theorem 2.12 then gives $|Df|\le g\mathfrak m$, so that $|Df|\ll\mathfrak m$.

(iii) ⇒ (i), under the doubling and Poincaré assumptions. We apply (2.17) with $r=\infty$ and denote by E_n the set of approximate continuity points of f where $M_\infty(|\nabla f|_{*,1}+|f|)$ is smaller than n. By the McShane Lipschitz extension theorem, we can extend $\tilde f|_{E_n}$ to a $2cn$-Lipschitz function on X, denoted by f_n, with $|f_n|\le n$. Since the weak L^1 estimate holds, it is then clear that $f_n\in L^1(X,\mathfrak m)$ and $\mathfrak m(\{f\ne f_n\})\le\mathfrak m(X\setminus E_n)\to 0$. In connection with equi-integrability of $|\nabla f_n|$, namely

$$\lim_{z\to\infty}\limsup_{n\to\infty}\int_{\{|\nabla f_n|>z\}}|\nabla f_n|\,d\mathfrak m = 0,$$

it suffices to split the integral on E_n and on $X\setminus E_n$. The former can be estimated uniformly in n from above, thanks to (4.5) and locality, with

$$c\int_{E_n\cap\{|\nabla f_n|_{*,1}>z/c\}}|\nabla f_n|_{*,1}\,d\mathfrak m \le c\int_{\{|\nabla f|_{*,1}>z/c\}}|\nabla f|_{*,1}\,d\mathfrak m$$

which is infinitesimal as $z\to\infty$. The latter can be estimated using (2.7). A similar and simpler argument shows that $f_n\to f$ in $L^1(X,\mathfrak m)$. □

The characterization (ii) in Theorem 4.3 of $H^{1,1}(X,d,\mathfrak m)$ suggests another definition of minimal 1-gradient $|\nabla f|_{w,1}$, namely the smallest function g (up to $\mathfrak m$-negligible sets)

such that, for some representative $\hat{f}$ of f, the inequality (4.6) holds. This is the point of view also adopted, besides the relaxation point of view, in the theory of Sobolev spaces with exponent $p > 1$, see [22,19,5] for a comparison between the definitions. From (4.7) we immediately get

$$|\nabla f|_{*,1} \le |\nabla f|_{w,1} \qquad \mathfrak{m}\text{-a.e. in } X, \tag{4.9}$$

while the inequality (4.5) gives, for some structural constant c,

$$|\nabla f|_{w,1} \le |\nabla f| \le c|\nabla f|_{*,1} \qquad \mathfrak{m}\text{-a.e. in } X. \tag{4.10}$$

Remark 4.4. In general equality does not hold in (4.9), not even under the doubling and Poincaré assumptions: in [16], a metric measure space is built endowing $X = [0,1]$ with the Euclidean distance and the weighted measure $\mathfrak{m} = \omega\mathscr{L}^1 \in \mathscr{P}(X)$, with $\omega \equiv 1$ on a "fat" Cantor set $K \subset (0,1)$ (i.e. a compact totally disconnected set with positive Lebesgue measure) and $\omega = 1/2$, say, on $(0,1) \setminus K$. It is clear that $(X, d, \mathfrak{m})$, being comparable to the standard Euclidean structure, satisfies the doubling and Poincaré assumptions. On the other hand, in [16] Lipschitz functions f_n convergent to the identity function f in $L^1(X, \mathfrak{m})$ are built in such a way that

$$\limsup_{n\to\infty} \int_X |\nabla f_n|\, d\mathfrak{m} < \int_X |\nabla f|\, d\mathfrak{m}.$$

Now, the very definition of $H^{1,1}(X, d, \mathfrak{m})$ gives that $|Df|(X) < \int_X |\nabla f|\, d\mathfrak{m}$, hence $|Df| = |\nabla f|_{*,1}\mathfrak{m}$ gives that $|\nabla f|_{*,1} < |\nabla f|$ in a set with positive $\mathfrak{m}$-measure. Since it is easy to check that $|\nabla f| = |\nabla f|_{w,1} = 1$ $\mathfrak{m}$-a.e. for the identity function, we have a situation where the two notions of weak gradient in the limiting case $p = 1$ differ. The difference of the notions arises basically from the fact that $f \mapsto \int_X |\nabla f|_{*,1}\, d\mathfrak{m}$ has an L^1-lower semicontinuity property built in, while $f \mapsto \int_X |\nabla f|_{w,1}\, d\mathfrak{m}$ a priori does not. ■

In the following proposition we show that $|\nabla f|_{w,1}$ enjoys a stronger approximation property by Lipschitz functions, compared to (2.15). In the proof we will use the identity

$$|\nabla f| = |\nabla f|_{w,1} \qquad \mathfrak{m}\text{-a.e. in } X, \text{ for all } f \in \mathrm{Lip}_{\mathrm{loc}}(X), \tag{4.11}$$

which is proved in [19, Theorem 12.5.1] and it extends Theorem 2.4 to the limiting case $p = 1$. For the sake of completeness we sketch the main ideas used to prove (4.11). Clearly, it suffices to prove that for every $f \in \mathrm{Lip}_{\mathrm{loc}}(X)$ it holds $|\nabla f| \le |\nabla f|_{w,1}$ $\mathfrak{m}$-a.e. in X. By the Vitali–Carathéodory theorem (see [19, Theorem 3.2.4]), it suffices to show that $|\nabla f| \le g$ for any countably valued lower semicontinuous upper gradient $g \in L^\infty_{\mathrm{loc}}(X)$ of f for which there is a positive real number c with $g \ge c$. By [19, Lemma 12.5.9] there is sequence (f_k) of Lipschitz functions, with a corresponding sequence (g_k) of continuous upper gradients, such that (f_k) converges to f in $L^2_{\mathrm{loc}}(X)$ and $\limsup_k g_k \le g$ $\mathfrak{m}$-a.e.

in X. Then (f_k) is bounded in $H^{1,2}_{\mathrm{loc}}(X)$ and we conclude exactly as in the proof of [19, Theorem 12.5.1].

Proposition 4.5. *Assume that $(X,d,\mathfrak{m})$ is doubling and that the $(1,1)$-Poincaré inequality holds for locally Lipschitz functions. Then for all $f \in H^{1,1}(X,d,\mathfrak{m})$ there exist locally Lipschitz functions f_n with $f_n \to f$ and $|\nabla f_n| \to |\nabla f|_{w,1}$ in $L^1(X,\mathfrak{m})$.*

Proof. Using the identity (4.11) it suffices to prove $|\nabla f_n|_{w,1} \to |\nabla f|_{w,1}$. Now, notice that, because of the inequality (4.10), the weak gradient $|\nabla g|_{w,1}$ is local on Borel sets as well as $|\nabla g|_{*,1}$ (but, this property could be proved directly with the definition of $|\nabla g|_{w,1}$). Therefore the sequence (f_n) provided by Theorem 4.3(i) provides the result. □

5. Sets of finite perimeter and the area formula

In the setup of the tensorization Section 3, assume that $Y = \mathbb{R}$ endowed with the standard Euclidean structure and the Lebesgue measure $\mathscr{L}^1$. This is a σ-finite and not finite metric measure space, but we can use a localization argument for weak gradients (in the same spirit of [1, Theorem 1.1] or [6, Lemma 4.11]) and apply all results of that section to this situation. Departing a bit from the notation of Section 3, we will then consider a m.m. space $(X,d,\mathfrak{m})$ and endow $X \times \mathbb{R}$ with the product distance $\tilde{d}$ and the product measure $\tilde{\mathfrak{m}} = \mathfrak{m} \times \mathscr{L}^1$.

If we consider a set of finite perimeter $E \subset X \times \mathbb{R}$ that is the subgraph of a function $f : X \to \mathbb{R}$, namely

$$E_f := \big\{(x,t) \in X \times \mathbb{R} : \ t < f(x)\big\}$$

then it is natural to compare the perimeter of E_f, defined according to the well established metric BV theory in $X \times \mathbb{R}$, to the area of the graph of f. In this context the natural regularity condition on f is $f \in H^{1,1}(X,d,\mathfrak{m})$ (or even $BV(X,d,\mathfrak{m})$); namely we would like to know whether

$$P(E_f, B \times \mathbb{R}) = \int_B \sqrt{1 + |\nabla f|^2_{*,1}}\, d\mathfrak{m} \qquad \text{for all } B \subset X \text{ Borel} \tag{5.1}$$

for $f \in H^{1,1}(X,d,\mathfrak{m})$. We will provide a partial positive answer using the tensorization property of weak gradients; this is not surprising, since locally Lipschitz functions on A are used to define the perimeter $P(E_f, A)$ for $A \subset X \times \mathbb{R}$ open.

Theorem 5.1. *Assume that $(X,d,\mathfrak{m})$ is a doubling metric measure space, and that the $(1,1)$-Poincaré inequality holds for Lipschitz functions. Then:*

(a) *for all* $f \in H^{1,1}(X,d,\mathfrak{m})$ *and all* $B \subset X$ *Borel one has*

$$P(E_f, B\times\mathbb{R}) \le \int_B \sqrt{1+|\nabla f|^2_{w,1}}\,d\mathfrak{m}. \tag{5.2}$$

(b) *for all* $f \in BV(X,d,\mathfrak{m})$, *denoting by* $|Df| = |\nabla f|_{*,1}\mathfrak{m} + |D^s f|$ *the Radon–Nikodym decomposition of* $|Df|$, *one has*

$$P(E_f, B\times\mathbb{R}) \ge \int_B \sqrt{1+|\nabla f|^2_{*,1}}\,d\mathfrak{m} + |D^s f|(B) \qquad \text{for all } B\subset X \text{ Borel.} \tag{5.3}$$

Proof. We first prove the inequality $\le$ in (5.2), for f locally Lipschitz and $B \subset X$ open. Let $\chi_\varepsilon \in C^\infty(\mathbb{R})$ be monotonically convergent to $\chi_{(0,\infty)}$ with $0 \le \chi_\varepsilon \le 1$, $\chi'_\varepsilon \ge 0$ and $\int \chi'_\varepsilon\,dt \le 1$. Then,

$$g_\varepsilon(x,t) := \chi_\varepsilon(f(x)-t)$$

provides a locally Lipschitz approximation of the characteristic function of E_f. Given our assumptions on $(X,d,\mathfrak{m})$, Theorem 2.4 (applied in $(X\times\mathbb{R}, \tilde{d}, \tilde{\mathfrak{m}})$ which is doubling and satisfies $(1,1)$-Poincaré for locally Lipschitz functions), Theorem 3.4 and the chain rule for weak gradients give

$$\begin{aligned}|\nabla g_\varepsilon|(x,t) = |\nabla g_\varepsilon|_*(x,t) &= \sqrt{|\nabla g^x_\varepsilon|^2_{*,\mathbb{R}}(t) + |\nabla g^t_\varepsilon|^2_{*,X}(x)}\\ &= \chi'_\varepsilon(f(x)-t)\sqrt{1+|\nabla f|^2_*(x)} = \chi'_\varepsilon(f(x)-t)\sqrt{1+|\nabla f|^2(x)}\end{aligned}$$

for $\tilde{\mathfrak{m}}$-a.e. $(x,t)\in X\times\mathbb{R}$. If we integrate on $B\times\mathbb{R}$ and use Fubini's theorem, we get

$$\int_{B\times\mathbb{R}} |\nabla g_\varepsilon|\,d\tilde{\mathfrak{m}} \le \int_B \sqrt{1+|\nabla f|^2}\,d\mathfrak{m}.$$

By the definition of $P(E_f, B\times\mathbb{R})$ we obtain the inequality

$$P(E_f, B\times\mathbb{R}) \le \int_B \sqrt{1+|\nabla f|^2}\,d\mathfrak{m}. \tag{5.4}$$

For $f \in H^{1,1}(X,d,\mathfrak{m})$ and B open we use the Lipschitz approximation provided by Proposition 4.5 and the lower semicontinuity of the perimeter in open sets to obtain (5.2). Being an inequality between positive and finite Borel measures, it extends from open to Borel sets.

We now prove the inequality (5.3). Let $g_n \in \mathrm{Lip}_{\mathrm{loc}}(X\times(-N,N))$ be convergent to χ_{E_f} in $L^1(X\times(-N,N),\tilde{\mathfrak{m}})$. By Fubini's theorem, possibly refining the sequence,

we can assume that $g_n^t \to \chi_{E_f}^t = \chi_{\{f(x)>t\}}$ in $L^1(X,\mathfrak{m})$ for $\mathscr{L}^1$-a.e. $t \in (-N,N)$, and that $g_n^x \to \chi_{E_f}^x = \chi_{\{t<f(x)\}}$ in $L^1(-N,N)$ for $\mathfrak{m}$-a.e. $x \in X$. Applying once more Theorem 2.4 in $X \times \mathbb{R}$ and Theorem 3.4 gives $|\nabla g_n|^2(x,t) = |\nabla g_n^x|(t)^2 + |\nabla g_n^t|^2(x)$ $\tilde{\mathfrak{m}}$-a.e. in $X \times (-N,N)$. Fix now a, $b : X \to [0,1]$ continuous with $a^2 + b^2 \le 1$ and notice that

$$\begin{aligned}
&\liminf_{n\to\infty} \int_{X\times(-N,N)} \sqrt{|\nabla g_n^x|^2(t) + |\nabla g_n^t|^2(x)}\, d\tilde{\mathfrak{m}}(x,t)\\
&\quad \ge \liminf_{n\to\infty} \int_{X\times(-N,N)} a(x)|\nabla g_n^x|(t) + b(x)|\nabla g_n^t|(x)\, d\tilde{\mathfrak{m}}(x,t)\\
&\quad \ge \int_X \liminf_{n\to\infty} \int_{(-N,N)} a(x)|\nabla g_n^x|(t)\, dt d\mathfrak{m}(x) + \int_{(-N,N)} \liminf_{n\to\infty} \int_X b(x)|\nabla g_n^t|(x)\, d\mathfrak{m}(x) dt\\
&\quad \ge \int_X a\chi_{f\in(-N,N)}\, d\mathfrak{m} + \int_{(-N,N)} \int_X b\, d|D\chi_{\{f>t\}}|\, dt.
\end{aligned}$$

Now we use the fact that (g_n) is arbitrary to get

$$P(E_f, X\times(-N,N)) \ge \int_X a\chi_{f\in(-N,N)}\, d\mathfrak{m} + \int_{(-N,N)} \int_X b\, d|D\chi_{\{f>t\}}|\, dt.$$

Letting $N \uparrow \infty$ and using the coarea formula

$$|Df| = \int_{-\infty}^{\infty} |D\chi_{\{f>t\}}|\, dt$$

(see for instance [21]) gives

$$P(E_f, X\times\mathbb{R}) \ge \int_X a\, d\mathfrak{m} + \int b\, d|Df| = \int_X (a + b|\nabla f|_{*,1})\, d\mathfrak{m} + \int_X b\, d|D^s f|.$$

Since $|D^s f| \perp \mathfrak{m}$, taking the supremum among all admissible pairs (a,b) we get (5.3) with $B = X$. Repeating the argument with any open set $B \subset X$ the inequality is proved on all open sets and then on all Borel sets. □

In the proof of the statement made in the next remark, the following lemma will be useful.

Lemma 5.2. *Let $u_n \in L^1(X,\mathfrak{m})$ be nonnegative functions satisfying*

$$\limsup_{n\to\infty} \int_X \sqrt{1+u_n^2}\, d\mathfrak{m} \le \int_X \sqrt{1+u^2}\, d\mathfrak{m}$$

for some nonegative $u \in L^1(X,\mathfrak{m})$. If $u_n\mathfrak{m}$ weakly converge to a measure $\mu \geq u\mathfrak{m}$, then $\mu = u\mathfrak{m}$ and $u_n \to u$ in $L^1(X,\mathfrak{m})$.

Proof. Let $\nu = \frac{\nu}{\mathfrak{m}}\mathfrak{m} + \nu^s$ be the Radon–Nikodym decomposition of ν with respect $\mathfrak{m}$. Since the functional

$$\nu \mapsto \int_X \sqrt{1 + \left|\frac{\nu}{\mathfrak{m}}\right|^2}\, d\mathfrak{m} + \nu^s(X)$$

is sequentially weakly lower semicontinuous and the density of μ is, by assumption, larger than u $\mathfrak{m}$-a.e., we obtain

$$\liminf_{n\to\infty} \int_X \sqrt{1+u_n^2}\, d\mathfrak{m} \geq \int_X \sqrt{1 + \left|\frac{\mu}{\mathfrak{m}}\right|^2}\, d\mathfrak{m} + \mu^s(X) \geq \int_X \sqrt{1+u^2}\, d\mathfrak{m} + \mu^s(X),$$

hence $\mu^s(X) = 0$, $\mu = u\mathfrak{m}$ and $\int \sqrt{1+u_n^2}\, d\mathfrak{m}$ converge to $\int \sqrt{1+u^2}\, d\mathfrak{m}$. We can now use the strict convexity of $z \mapsto \sqrt{1+z^2}$ (see for instance Exercise 1.20 of [2]) to conclude. □

By the lower semicontinuous nature of both functionals, we believe that equality always holds in (5.3) at least on $f \in H^{1,1}(X,d,\mathfrak{m})$, i.e.

$$P(E_f, B\times\mathbb{R}) = \int_B \sqrt{1+|\nabla f|^2_{*,1}}\, d\mathfrak{m} \qquad \text{for all } B \subset X \text{ Borel.}$$

In the next remark, instead, we compare with the relaxation point of view of [16].

Remark 5.3. Let us discuss the question of the validity of the equality

$$\begin{aligned} &\inf\left\{ \liminf_{n\to\infty} \int_X \sqrt{1+|\nabla f_n|^2}\, d\mathfrak{m} :\ f_n \in \mathrm{Lip}_{\mathrm{loc}}(X),\ f_n \to f \text{ in } L^1(X,\mathfrak{m}) \right\} \\ &\quad = \int_X \sqrt{1+|\nabla f|^2_{*,1}}\, d\mathfrak{m} \end{aligned} \tag{5.5}$$

for $f \in H^{1,1}(X,d,\mathfrak{m})$. Notice that the inequality $\geq$ follows at once from the lower semicontinuity of the perimeter and (5.4), (5.3) with $B = X$, while the inequality $\leq$ with the apriori larger gradient $|\nabla f|_{w,1}$ instead of $|\nabla f|_{*,1}$ in the right hand side follows immediately by Proposition 4.5. Hence, if $|\nabla f|_{*,1} = |\nabla f|_{w,1}$ $\mathfrak{m}$-a.e. in X, equality holds in (5.5). Conversely, if equality holds in (5.5), a diagonal argument provides $f_n \in \mathrm{Lip}_{\mathrm{loc}}(X)$ such that $f_n \to f$ in $L^1(X,\mathfrak{m})$ and

$$\limsup_{n\to\infty} \int_X \sqrt{1+|\nabla f_n|^2}\, d\mathfrak{m} \leq \int_X \sqrt{1+|\nabla f|^2_{*,1}}\, d\mathfrak{m}.$$

If we denote by μ a weak limit of $|\nabla f_n|$, we obviously have $\mu \geq |Df|$ on open sets A, hence $\mu \geq |\nabla f|_{*,1}\mathfrak{m}$. From Lemma 5.2 we obtain that $\mu \ll \mathfrak{m}$ and that $|\nabla f_n| \to |\nabla f|_{*,1}$ in $L^1(X,\mathfrak{m})$. As in [22], this implies the existence of representatives $\tilde{f}$ of f and g of $|\nabla f|_{*,1}$ such that

$$|\tilde{f}(\gamma(1)) - \tilde{f}(\gamma(0))| \leq \int_\gamma g$$

for Mod_1-a.e. curve γ, hence $|\nabla f|_{*,1} \geq |\nabla f|_{w,1}$ $\mathfrak{m}$-a.e. in X.

Acknowledgments

We would like to thank N. Shanmugalingam for helpful hints on the theory of weak gradients in the critical case $p = 1$ and the reviewer for the careful reading of the paper. The authors acknowledge the support of the grant ERC ADG GeMeThNES (Grant Agreement No. 246923). The second author has been partially supported by the Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM).

Appendix A. Proof of the inequality $|\nabla f|_* \geq |\nabla f|_{*,c}$

In [6], a very detailed analysis of the fine properties of the Hopf–Lax semigroup

$$Q_t g(w) := \inf_{w' \in W} g(w') + \frac{1}{2t} d_W^2(w', w) \tag{A.1}$$

in a metric space (W, d_W) has been made. The analysis is based on the quantities

$$D_g^+(w,t) := \sup \limsup_{n\to\infty} d_W(w, w_n'), \qquad D_g^-(w,t) := \inf \liminf_{n\to\infty} d_W(w, w_n),$$

where the supremum and the infimum run among all minimizing sequences (w_n) in (A.1). These quantities reduce respectively to the maximum and minimum distance from w of minimizers in the locally compact case. Confining for simplicity our discussion to the case of bounded functions $g : W \to \mathbb{R}$, which suffices for our purposes, for all $t > 0$ it has been shown that (see Propositions 3.2, 3.3 and 3.4 of [6]):

(a) D_g^+ and D_g^- are respectively upper and lower semicontinuous in $W \times (0,\infty)$ and, given w, $D_g^+(w,t) = D_g^-(w,t)$ with at most countably many exceptions;
(b) $Q_t g$ is Lipschitz in X, with $\mathrm{Lip}(Q_t g) \leq 2\sqrt{\mathrm{osc}(g)/t}$, where $\mathrm{osc}(f) = \sup f - \inf f$;
(c) $D_g^-(\cdot,t)/t$ is an upper gradient of $Q_t g$;
(d) the following pointwise equality holds:

$$\frac{d^+}{dt} Q_t g(w) + \frac{(D_g^+(w,t))^2}{2t^2} = 0, \tag{A.2}$$

where d^+/dt stands for right derivative (part of the statement is its existence at every point).

Notice that, since $D_g^+(\cdot,t)/t \geq D_g^-(\cdot,t)/t$ is an upper semicontinuous upper gradient of $Q_t g$, it bounds the slope of $Q_t g$ from above. Therefore (A.2) implies the Hamilton–Jacobi subsolution property $\frac{d^+}{dt}Q_t g + |\nabla Q_t g|^2/2 \leq 0$, but the sharper form (A.2) is often essential for the proofs.

More precisely, our proof of the inequality $|\nabla f|_* \geq |\nabla f|_{*,c}$ is based on a refinement of Lemma 6.16, Lemma 6.17 and Proposition 6.18 of [7] and it is split in these steps:

Step 1. *Let $g : Z \to \mathbb{R}$ be a bounded function and set*

$$Q_t g(x,y) := \inf\left\{ g(x',y') + \frac{1}{2t}\big(d_X^2(x,x') + d_Y^2(y,y')\big)\right\}. \tag{A.3}$$

In this step we show that for $\mathscr{L}^1$-a.e. $t > 0$ one has

$$\frac{d^+}{dt}Q_t g + \frac{1}{2}|\nabla Q_t g|_c^2 \leq 0 \qquad \mathfrak{m}\text{-a.e. in } Z. \tag{A.4}$$

In (6.40) of [7], starting from (A.2), an inequality similar to (A.4) for *all* $t > 0$ has been proved, where instead of $|\nabla Q_t g|_c^2$ the sum $|(Q_t g)^x|_{*,Y}^2 + |(Q_t g)^y|_{*,X}^2$ was used. Our refinement (A.4) still relies on (A.2) and the pointwise inequality (see (6.43) of [7])

$$[D_g^+((x,y),t)]^2 \geq [D^-_{L_{t,y}}(x,t)]^2 + [D^-_{R_{t,x}}(y,t)]^2 \tag{A.5}$$

where $L_{t,y}(x') := Q_t^Y g(x',\cdot)(y)$ and $R_{t,x}(y') := Q_t^X g(\cdot,y')(x)$. By minimizing first w.r.t. to one variable and then with respect to the other variable in (A.3) we have the easy identities

$$(Q_t g)^y(x) = Q_t^X(L_{t,y})(x), \qquad (Q_t g)^x(y) = Q_t^Y(R_{t,x})(y). \tag{A.6}$$

By (a) we can apply Fubini's theorem to convert D^- into D^+ in (A.5), thus obtaining

$$[D_g^+((x,y),t)]^2 \geq [D^+_{L_{t,y}}(x,t)]^2 + [D^+_{R_{t,x}}(y,t)]^2 \qquad \text{for } \mathfrak{m}\text{-a.e. } (x,y) \in Z$$

for $\mathscr{L}^1$-a.e. $t > 0$. Since D^+ is an upper semicontinuous upper gradient and upper semicontinuous upper gradients bound, as we already said, the slope from above, we can use (A.6) to get eventually (A.4).

Step 2. (The so-called Kuwada lemma.) *In this step we show that if $h \in L^\infty(Z,\mathfrak{m})$ is a probability density, and h_t is the solution to the gradient flow of Ch_c starting from h, then $\mu_t = h_t\mathfrak{m} \in \mathscr{P}(Z)$, $t \mapsto \mu_t$ is locally absolutely continuous from $[0,\infty)$ to $\mathscr{P}(Z)$ endowed with the quadratic Wasserstein distance W_2. In addition, its metric speed satisfies*

$$|\dot\mu_t|^2 \le \int_{\{h_t>0\}} \frac{|\nabla h_t|^2_{*,c}}{h_t}\,d\mathfrak{m} \qquad \text{for } \mathscr{L}^1\text{-a.e. } t>0. \tag{A.7}$$

The proof of this fact uses the Hamilton–Jacobi subsolution property (A.4) of Step 1 (used in integral form, so its validity for $\mathscr{L}^1$-a.e. $t>0$ is sufficient), as well as the integration by parts formula (2.9) with the Laplacian Δ_G corresponding to the energy Ch_G induced by the pseudo gradient $G(f)=|\nabla f|_c$.

Step 3. *In this step we conclude the proof of the inequality* $|\nabla f|_* \ge |\nabla f|_{*,c}$. Now for any bounded nonnegative $f\in D(\mathsf{Ch})$ with $\int f^2\,d\mathfrak{m}=1$ we can find, arguing exactly as in Proposition 6.18 of [7], a sequence of locally Lipschitz functions f_n convergent to f in $L^2(X,\mathfrak{m})$ satisfying

$$\limsup_{n\to\infty}\int_Z |\nabla f_n|^2_{*,c}\,d\mathfrak{m} \le \mathsf{Ch}(f) \tag{A.8}$$

(the only difference is that the gradient flow of Ch_c has to be used, as in the previous steps). Therefore $f\in D(\mathsf{Ch}_c)$ and $\mathsf{Ch}_c(f)\le \mathsf{Ch}(f)$ which, in combination with (3.7), gives the result. Using invariance under addition by constants and homogeneity, we extend the result to all bounded functions f. Eventually a truncation argument and the locality of weak gradients provide the result for general f. The preliminary reduction to nonnegative and normalized f's is necessary in view of Step 2, because we use in the construction of f_n the estimate (A.7) on metric derivative for a $\mathscr{P}(Z)$-valued map.

We briefly sketch, for the reader's convenience the argument leading to (A.8), referring to [6, Theorem 6.2] and [7, Proposition 6.18] for more details. By homogeneity and invariance under addition of constants we can assume, besides $\int f^2\,d\mathfrak{m}=1$, that $c^{-1}\ge f\ge c>0$ $\mathfrak{m}$-almost everywhere in Z. We consider the gradient flow (h_t) of Ch_c with initial datum $h:=f^2$, setting $\mu_t=h_t\mathfrak{m}$. The maximum principle yields $c^{-2}\ge h_t\ge c^2$ and a standard argument based on chain rule and integration by parts (2.9) yields the energy dissipation identity

$$\frac{d}{dt}\int h_t\log h_t\,d\mathfrak{m} = -\int_{\{h_t>0\}}\frac{|\nabla h_t|^2_{*,c}}{h_t}\,d\mathfrak{m} \qquad \text{for } \mathscr{L}^1\text{-a.e. } t>0. \tag{A.9}$$

Let $g=h^{-1}|\nabla h|_*$; by the chain rule we know that $\log h$ is Sobolev along almost every curve and we can use the same argument of [6, Theorem 6.2] to get

$$\int \big(h\log h - h_t\log h_t\big)\,d\mathfrak{m} \le \int \log h(h-h_t)\,d\mathfrak{m} \le \Big(\int_0^t\int g^2 h_s\,d\mathfrak{m}\,ds\Big)^{1/2}\Big(\int_0^t|\dot\mu_s|^2\,ds\Big)^{1/2}.$$

Now, inequality (A.7) gives

$$\int \left(h\log h - h_t \log h_t\right) d\mathfrak{m} \le \frac{1}{2}\int_0^t \int g^2 h_s \, d\mathfrak{m}\, ds + \frac{1}{2}\int_0^t |\dot{\mu}_s|^2 \, ds$$
$$\le \frac{1}{2}\int_0^t \int g^2 h_s \, d\mathfrak{m}\, ds + \frac{1}{2}\int_0^t \int_{\{h_s>0\}} \frac{|\nabla h_s|^2_{*,c}}{h_s} \, d\mathfrak{m}\, ds.$$

Recalling the entropy dissipation formula (A.9) we obtain

$$\int_0^t \int_{\{h_s>0\}} \frac{|\nabla h_s|^2_{*,c}}{h_s} \, d\mathfrak{m}\, ds \le \int_0^t \int g^2 h_s \, d\mathfrak{m}\, ds.$$

Now, the chain rule and the identity $g = 2f^{-1}|\nabla f|_*$ give $\int_0^t \mathsf{Ch}_c(\sqrt{h_s})\, ds \le \int_0^t \int |\nabla f|^2_* f^{-2} h_s \, d\mathfrak{m}\, ds$, so that dividing by t and passing to the limit as $t \downarrow 0$ we get

$$\limsup_{t\downarrow 0} \frac{1}{t}\int_0^t \mathsf{Ch}_c(\sqrt{h_s})\, ds \le \int_X |\nabla f|^2_* \, d\mathfrak{m}.$$

Therefore there exists $s_i \downarrow 0$ with $\limsup_i \mathsf{Ch}_c(\sqrt{h_{s_i}}) \le \mathsf{Ch}(f)$. Since $\sqrt{h_{s_i}}$ are equibounded and converge strongly to f in $L^2(Z,\mathfrak{m})$ as $s \downarrow 0$, a diagonal argument provides (A.8).

References

[1] L. Ambrosio, S. Di Marino, Equivalent definitions of BV space and of total variation on metric measure spaces, J. Funct. Anal. 266 (2014) 4150–4188.
[2] L. Ambrosio, N. Fusco, D. Pallara, Functions of Bounded Variation and Free Discontinuity Problems, Oxford University Press, 2000.
[3] L. Ambrosio, N. Gigli, G. Savaré, Gradient Flows in Metric Spaces and in the Space of Probability Measures, second ed., Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2008.
[4] L. Ambrosio, S. Di Marino, G. Savaré, On the duality between p-modulus and probability measures, to appear on JEMS, arXiv:1311.1381, 2013.
[5] L. Ambrosio, N. Gigli, G. Savaré, Density of Lipschitz functions and equivalence of weak gradients in metric measure spaces, Rev. Mat. Iberoam. 29 (2013) 969–996.
[6] L. Ambrosio, N. Gigli, G. Savaré, Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below, Invent. Math. 195 (2014) 289–391.
[7] L. Ambrosio, N. Gigli, G. Savaré, Metric measure spaces with Riemannian Ricci curvature bounded from below, Duke Math. J. 163 (2014) 1405–1490.
[8] L. Ambrosio, M. Colombo, S. Di Marino, Sobolev spaces in metric measure spaces: reflexivity and lower semicontinuity of slope, Adv. Stud. Pure Math. (2015), in press, arXiv:1212.3779, 2012.
[9] H. Breziz, Analyse Fonctionelle, Masson, Paris, 1983.
[10] J. Cheeger, Differentiability of Lipschitz functions on metric measure spaces, Geom. Funct. Anal. 9 (1999) 428–517.
[11] S. Di Marino, G. Speight, The p weak gradient depends on p, Proc. Amer. Math. Soc. (2015), http://dx.doi.org/10.1090/S0002-9939-2015-12641-X, in press.
[12] N. Gigli, On the Differential Structure of Metric Measure Spaces and Applications, Memoirs AMS, vol. 236 (1113), 2015.

[13] N. Gigli, B.-X. Han, Unifications of weak gradients on metric measure spaces, Calc. Var. Partial Differential Equations 53 (1–2) (May 2015) 149–177.
[14] V. Gol'dshtein, M. Troyanov, Axiomatic theory of Sobolev spaces, Expo. Math. 19 (2001) 289–336.
[15] P. Hajlasz, P. Koskela, Sobolev Met Poincaré, Memoirs AMS, vol. 688, 2000.
[16] H. Hakkarainen, J. Kinnunen, P. Lahti, P. Lehtelä, Relaxation and integral representation for functionals of linear growth on metric measure spaces, arXiv:1401.5717, 2014.
[17] H. Hakkarainen, J. Kinnunen, P. Lahti, Regularity of minimizers of the area functional in metric spaces, Adv. Calc. Var. 8 (1) (2015) 55–68.
[18] J. Heinonen, Nonsmooth calculus, Bull. Amer. Math. Soc. (N.S.) 44 (2007) 163–232.
[19] J. Heinonen, P. Koskela, N. Shanmugalingham, J. Tyson, Sobolev Spaces on Metric Measure Spaces: An Approach Based on Upper Gradients, Mathematical Monographs, vol. 27, Cambridge University Press, New, ISBN 9781107092341, 2015.
[20] R. Korte, P. Lahti, N. Shanmugalingam, Semmes family of curves and a characterization of functions of bounded variation in terms of curves, Calc. Var. Partial Differential Equations (2015), http://dx.doi.org/10.1007/s00526-015-0829-y, in press.
[21] M. Miranda Jr, Functions of bounded variation on "good" metric spaces, J. Math. Pures Appl. 82 (2003) 975–1004.
[22] N. Shanmugalingham, Newtonian spaces: an extension of Sobolev spaces to metric measure spaces, Rev. Mat. Iberoam. 16 (2000) 243–279.