GENERALIZED BAKRY–ÉMERY CURVATURE CONDITION AND EQUIVALENT ENTROPIC INEQUALITIES IN GROUPS

GIORGIO STEFANI

ABSTRACT. We study a generalization of the Bakry–Émery pointwise gradient estimate for the heat semigroup and its equivalence with some entropic inequalities along the heat flow and Wasserstein geodesics for metric-measure spaces with a suitable group structure. Our main result applies to Carnot groups of any step and to the $\mathbb{SU}(2)$ group.

1. INTRODUCTION

1.1. **The Riemannian framework.** Let $(\mathbb{M}, \mathrm{g})$ be a (complete and connected) N-dimensional smooth Riemannian manifold with Laplace–Beltrami operator Δ. The celebrated *Bochner formula* states that

$$\frac{1}{2} \Delta |\nabla f|_{\mathrm{g}}^2 = \langle \nabla \Delta f, \nabla f \rangle_{\mathrm{g}} + ||\mathrm{Hess} f||_2^2 + \mathrm{Ric}(\nabla f, \nabla f) \tag{1.1}$$

for all $f \in \mathrm{C}^\infty(\mathbb{M})$. Defining

$$\begin{aligned} \Gamma(f, g) &= \frac{1}{2} \Big(\Delta(fg) - f \, \Delta g - g \, \Delta f \Big) = \langle \nabla f, \nabla g \rangle_{\mathrm{g}}, \\ \Gamma_2(f, g) &= \frac{1}{2} \big(\Delta \Gamma(f, g) - \Gamma(f, \Delta g) - \Gamma(\Delta f, g) \big) \end{aligned} \tag{1.2}$$

for all $f, g \in \mathrm{C}^\infty(\mathbb{M})$, we can rewrite (1.1) as

$$\frac{1}{2} \Delta \Gamma(f) = \Gamma(\Delta f, f) + ||\mathrm{Hess} f||_2^2 + \mathrm{Ric}(\nabla f, \nabla f),$$

so that

$$\Gamma_2(f) = ||\mathrm{Hess} f||_2^2 + \mathrm{Ric}(\nabla f, \nabla f)$$

for all $f \in \mathrm{C}^\infty(\mathbb{M})$. Here and in the following, we write $\Gamma(f) = \Gamma(f, f)$ and $\Gamma_2(f) = \Gamma_2(f, f)$ for simplicity. Using Cauchy–Schwartz inequality, we can estimate

$$||\mathrm{Hess} f||_2^2 \geq \frac{1}{N} (\Delta f)^2,$$

thus the *geometric* information

$$\mathrm{Ric} \geq K \text{ for some } K \in \mathbb{R} \tag{1.3}$$

implies the *analytical* information

$$\Gamma_2(f) \geq \frac{1}{N}\,(\Delta f)^2 + K\,\Gamma(f) \tag{1.4}$$

for all $f \in \mathrm{C}^\infty(\mathbb{M})$. Nowadays, (1.4) is the so-called and well-known *Bakry–Émery curvature-dimension inequality* $\mathsf{CD}(K, N)$. Remarkably, it is also possible to prove the converse implication, see [22, Proposition 6.2]: if a Riemannian manifold $\mathbb{M}$ satisfies $\mathsf{CD}(K, N)$ for some $K \in \mathbb{R}$ and $N \in (0, +\infty)$, then $\dim \mathbb{M} \leq N$ and $\mathrm{Ric} \geq K$.

Let us now drop the role of the dimension of $\mathbb{M}$ (which formally corresponds to the choice $N = +\infty$ in (1.4)) and focus on the lower bound on the Ricci tensor encoded by the $\mathsf{CD}(K, \infty)$ condition. After the works of Bakry–Émery [27], Otto–Villani [123], Cordero-Erausquin–McCann–Schmuckenschläger [64] and von Renesse–Sturm [127], the analytical condition (1.4) for $N = +\infty$ on a Riemannian manifold can be equivalently formulated in other three ways (at least, see [127] for other equivalent statements): via the *pointwise gradient estimate* for the heat flow, via the Wasserstein *contractivity property* of the dual heat flow and via the K-*convexity* of the entropy along geodesics in the Wasserstein space.

The *heat kernel* $\mathsf{p}_t(x, y)$ of the Riemannian manifold $(\mathbb{M}, \mathrm{g})$ is the fundamental solution of the heat differential operator $\partial_t - \Delta$. The function $\mathsf{p}_t(x, y)$ is smooth in $(t, x, y) \in (0, +\infty) \times \mathbb{M} \times \mathbb{M}$, symmetric in (x, y) and naturally defines the associated *heat semigroup* $\mathsf{P}_t \colon \mathrm{C}_c^\infty(\mathbb{M}) \to \mathrm{C}^\infty(\mathbb{M})$ as

$$\mathsf{P}_t f(x) = \int_{\mathbb{M}} f(y)\,\mathsf{p}_t(x, y)\,\mathrm{dVol}_{\mathrm{g}}(y), \quad x \in \mathbb{M},$$

for all $f \in \mathrm{C}_c^\infty(\mathbb{M})$. Inequality (1.4) describes the behavior of the commutation between the gradient ∇ and the heat semigroup $(\mathsf{P}_t)_{t>0}$. More precisely, the $\mathsf{CD}(K, \infty)$ condition is equivalent to the *Bakry–Émery pointwise gradient estimate*

$$\Gamma(\mathsf{P}_t f) \leq e^{-2Kt}\,\mathsf{P}_t \Gamma(f) \tag{1.5}$$

for all $t > 0$ and $f \in \mathrm{C}_c^\infty(\mathbb{M})$.

The *dual heat semigroup* $\mathsf{H}_t \colon \mathscr{P}(\mathbb{M}) \to \mathscr{P}(\mathbb{M})$ is nothing but the extension of the heat semigroup to the space $\mathscr{P}(\mathbb{M})$ of probability measures on $\mathbb{M}$ and can be defined by setting

$$\int_{\mathbb{M}} f\,\mathrm{d}\mathsf{H}_t\mu = \int_{\mathbb{M}} \mathsf{P}_t f\,\mathrm{d}\mu$$

for all $f \in \mathrm{C}_c^\infty(\mathbb{M})$, whenever $\mu \in \mathscr{P}(\mathbb{M})$. The subset of $\mathscr{P}(\mathbb{M})$ of probability measures with finite second moment

$$\mathscr{P}_2(\mathbb{M}) = \left\{ \mu \in \mathscr{P}(\mathbb{M}) : \int_{\mathbb{M}} \mathsf{d}_{\mathrm{g}}^2(x, x_0)\,\mathrm{d}\mu(x) < +\infty \text{ for some } x_0 \in \mathbb{M} \right\}$$

endowed with the 2-*Wasserstein distance* W_2, given by

$$\frac{1}{2}\,W_2^2(\mu, \nu) = \sup\left\{ \int_{\mathbb{M}} \varphi\,\mathrm{d}\mu + \int_{\mathbb{M}} \psi\,\mathrm{d}\nu : \varphi(x) + \psi(y) \leq \frac{1}{2}\,\mathsf{d}_{\mathrm{g}}^2(x, y) \text{ for all } x, y \in \mathbb{M} \right\}$$

for all $\mu, \nu \in \mathscr{P}_2(X)$, is a complete and separable geodesic space. The lower bound (1.3) on the Ricci tensor can be equivalently stated as a *contractivity property* of the dual heat semigroup with respect to the 2-Wasserstein distance, in the sense that

$$W_2(\mathsf{H}_t\mu, \mathsf{H}_t\nu) \leq e^{-Kt}\,W_2(\mu, \nu) \tag{1.6}$$

for all $t > 0$ and $\mu, \nu \in \mathscr{P}_2(X)$.

The (Boltzmann) *entropy* with respect to the volume measure $\mathrm{Vol}_{\mathrm{g}}$ is defined as

$$\mathsf{Ent}(\mu) = \begin{cases} \int_{\mathbb{M}} f \log f \, \mathrm{dVol}_{\mathrm{g}} & \text{if } \mu = f \, \mathrm{Vol}_{\mathrm{g}}, \\ +\infty & \text{otherwise,} \end{cases}$$

whenever $\mu \in \mathscr{P}_2(X)$. Note that, by the Bishop Volume Comparison Theorem, it actually holds that $\mathsf{Ent}(\mu) > -\infty$ for all $\mu \in \mathscr{P}_2(\mathbb{M})$, see [71, Lemma 4.1] for the proof. The $\mathsf{CD}(K, \infty)$ condition can be equivalently reformulated as a convexity property of the entropy along all (constant speed, as usual) geodesics joining two measures in $\mathscr{P}_2(\mathbb{M})$. More precisely, if $[0,1] \in s \mapsto \mu_s \in \mathscr{P}_2(\mathbb{M})$ is a geodesic joining $\mu_0, \mu_1 \in \mathscr{P}_2(\mathbb{M})$, then the following *displacement K-convexity inequality*

$$\mathsf{Ent}(\mu_s) \leq (1-s) \, \mathsf{Ent}(\mu_0) + s \, \mathsf{Ent}(\mu_1) - \frac{K}{2} \, s(1-s) \, W_2^2(\mu_0, \mu_1) \qquad (1.7)$$

holds for all $s \in [0,1]$.

1.2. **The non-smooth framework: $\mathsf{CD}(K, \infty)$ spaces.** The peculiar feature of inequality (1.7) is that it can be stated uniquely in terms of the distance and the volume measure, no matter they come from the underlying smooth structure of the Riemannian manifold, and thus can be considered as a metric-measure definition of the lower bound on the Ricci tensor.

This observation has led Lott–Villani and Sturm in their groundbreaking works [110, 137, 138] to study the properties of very general metric-measure spaces $(X, \mathsf{d}, \mathfrak{m})$ satisfying the displacement K-convexity for some $K \in \mathbb{R}$. Besides the many powerful consequences successively derived from their ideas, see [81, 109, 126] and the monograph [141] for example, a key feature of the Lott–Sturm–Villani approach is that the displacement K-convexity of the entropy actually provides a metric-measure definition of the $\mathsf{CD}(K, \infty)$ condition that is stable under (metric-measure) Gromov–Hausdorff convergence.

As pointed out by Gigli [81], starting from the $\mathsf{CD}(K, \infty)$ condition, it is possible to prove that the *metric gradient flow* (see the monograph [12]) $(\mathsf{S}_t)_{t>0}$ of the entropy functional in $(\mathscr{P}_2(X), W_2)$ is an evolution semigroup on the convex subset of $\mathscr{P}_2(X)$ given by probability measures with finite entropy. However, since also Finsler geometries (as in the flat case of $\mathbb{R}^N$ endowed with a non-Euclidean norm) can satisfy the $\mathsf{CD}(K, \infty)$ condition, the semigroup $(\mathsf{S}_t)_{t>0}$ can be non-linear in such a general setting. Nevertheless, $(\mathsf{S}_t)_{t>0}$ can be extended to a continuous semigroup of contractions in $\mathrm{L}^2(X, \mathfrak{m})$ (and actually in any L^p-space) which can be also characterized as the gradient flow in $\mathrm{L}^2(X, \mathfrak{m})$ of the convex and 2-homogeneous functional

$$\mathsf{Ch}(f) = \inf\left\{ \liminf_{n \to +\infty} \frac{1}{2} \int_X |\mathrm{D}f_n|^2 \, \mathrm{d}\mathfrak{m} : f_n \in \mathrm{Lip}_b(X), \ f_n \to f \text{ in } \mathrm{L}^2(X, \mathfrak{m}) \right\}, \qquad (1.8)$$

the *Cheeger energy* of $f \in \mathrm{L}^2(X, \mathfrak{m})$, see the celebrated work [62]. Here

$$|\mathrm{D}f|(x) = \limsup_{y \to x} \frac{|f(y) - f(x)|}{\mathsf{d}(y, x)} \qquad (1.9)$$

is the *slope* at $x \in X$ of the bounded Lipschitz function $f \in \mathrm{Lip}_b(X)$. Since the slope (1.9) plays the same role of the absolute value of the gradient in the smooth framework, it is natural to consider the gradient flow of the Cheeger energy as a metric-measure definition of the heat flow $(\mathsf{P}_t)_{t>0}$ in the non-smooth context. The identification of the entropic semigroup and the heat flow has been proved in Euclidean spaces in [100]

by Jordan–Kinderleher–Otto (see also [122]) and then extended to Riemannian manifolds [71, 141], Hilbert spaces [19], Finsler spaces [120], Alexandrov spaces [84] and eventually to $\mathsf{CD}(K, \infty)$ spaces in the fundamental work [14] of Ambrosio–Gigli–Savaré. We refer the reader also to the works of Kuwada [97, 98] for their key role in the understanding of the equivalence between the gradient estimate (1.5) and the W_2-contraction inequality (1.6) in the non-smooth framework.

Having a metric-measure notion of heat flow $(\mathsf{P}_t)_{t>0}$ at hand, it is then natural to see if the displacement K-convexity is still equivalent to suitable analogues of the Bakry–Émery inequality (1.5) and the W_2-contractivity property (1.6) in this abstract setting. Building upon the non-smooth Calculus developed in [14], Ambrosio–Gigli–Savaré in [15, 16] and Ambrosio–Gigli–Mondino–Rajala in [11] proved this equivalence under the additional assumption that the heat flow $(\mathsf{P}_t)_{t>0}$ is linear or, equivalently, that the Cheeger energy (1.8) is a Dirichlet (and thus *quadratic*) form on $\mathrm{L}^2(X, \mathfrak{m})$, in order to naturally rule out Finsler geometries. For this reason, such metric-measure spaces, forming a smaller family of $\mathsf{CD}(K, \infty)$ spaces remarkably still stable under Gromov–Hausdorff convergence, are said to have *Riemannian* Ricci curvature bounded from below by $K \in \mathbb{R}$, or *infinitesimally Hilbertian* $\mathsf{CD}(K, \infty)$ spaces, or $\mathsf{RCD}(K, \infty)$ spaces for short. We refer the reader to [121] and to [84, 147] for strictly related results in Finsler and Alexandrov spaces respectively.

One of the most important results of [11, 15] is that $\mathsf{RCD}(K, \infty)$ spaces can be equivalently characterized as those metric-measure spaces for which the gradient flow $(\mathsf{S}_t)_{t>0}$ of the entropy in the Wasserstein space $(\mathscr{P}_2(X), W_2)$ satisfies the following *Evolution Variational Inequality* with parameter $K \in \mathbb{R}$, EVI_K for short,

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{W_2^2(\mathsf{S}_t\mu, \nu)}{2} + \frac{K}{2} W_2^2(\mathsf{S}_t\mu, \nu) + \mathsf{Ent}(\mathsf{S}_t\mu) \leq \mathsf{Ent}(\nu) \quad \text{for a.e. } t > 0 \qquad (1.10)$$

whenever $\mu, \nu \in \mathscr{P}_2(X)$. Thus, in the infinitesimally Hilbertian case (and so in the particular case of smooth Riemannian manifolds), inequality (1.10) provides an alternative equivalent metric-measure formulation of lower bound on the Ricci curvature.

The above analysis has been extended also to the finite dimensional case $N \in (0, +\infty)$, where however the equivalent metric-measure formulations of the lower bound on the Ricci curvature and the upper bound on the dimension become more involved. We refer the reader to the seminal works of Erbar–Kuwada–Sturm [72] and Kuwada [99], and to the more recent developments obtained in [53, 54, 80]. The theory of $\mathsf{CD}(K, N)$ spaces has been extended also to the case of *negative* dimension $N \in (-\infty, 0)$, see [119]. Finally, we refer the reader to the recent work [139] by Sturm for a more general approach to curvature in the metric-measure setting.

1.3. The sub-Riemannian framework.

Although the class of $\mathsf{CD}(K, N)$ spaces is very broad, a large and widely-studied family of spaces is left out, the sub-Riemannian manifolds. For an introduction on the subject, we refer the reader to the papers [104, 133, 134] and to the monographs [1, 55, 116].

A *sub-Riemannian manifold* is a triple $(\mathbb{M}, \mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{H}})$, where $\mathbb{M}$ is a (connected) smooth manifold, $\mathcal{H} \subset T\mathbb{M}$ is a sub-bundle of the tangent bundle $T\mathbb{M}$ and $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ is a smoothly varying positive definite quadratic form on $\mathcal{H}$. Typically, the sub-bundle $\mathcal{H}$ is assumed to be *bracket generating* and *equiregular*, that is, at each point $x \in \mathbb{M}$ the directions in $\mathcal{H}_x$ together with all their Lie brackets generate the full tangent space $T_x\mathbb{M}$, and the dimensions of the intermediate sub-bundles of commutators obtained at each step do not depend on the choice of $x \in \mathbb{M}$.

A sub-Riemannian manifold $(\mathbb{M}, \mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{H}})$ naturally carries a metric notion, the so-called *Carnot–Carathéodory* (CC for short) *distance*, defined as

$$\mathsf{d}_{\mathrm{cc}}(x, y) = \inf\left\{\int_0^1 |\dot\gamma_t|_{\mathcal{H}}^2 \, \mathrm{d}t : \gamma\colon [0, 1] \to \mathbb{M} \text{ is horizontal, } \gamma_0 = x, \ \gamma_1 = y\right\} \qquad (1.11)$$

for all $x, y \in \mathbb{M}$. A Lipschitz curve $\gamma\colon [0, 1] \to \mathbb{M}$ is *horizontal* if $\dot\gamma_t \in \mathcal{H}_{\gamma_t}$ for a.e. $t \in [0, 1]$. The bracket generating assumption on $\mathcal{H}$ ensures that the function in (1.11) is a finite distance — this is the celebrated Chow–Rashevskii Theorem, see [1, Theorem 3.31] for example.

A sub-Riemannian manifold $(\mathbb{M}, \mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{H}})$ can be endowed with the Hausdorff measure associated to d_{cc}. However, differently from the Riemannian case, the Hausdorff measure does not coincide in general with the volume induced by the distribution $\mathcal{H}$, the *Popp measure* $\mu_{\mathcal{H}}$. We refer the reader to [1, Chapter 20] for the precise definitions and constructions.

The $\mathsf{CD}(K, N)$ condition fails for the metric-measure space $(\mathbb{M}, \mathsf{d}_{\mathrm{cc}}, \mu_{\mathcal{H}})$ and, for this reason, sub-Riemannian manifolds are said to have Ricci curvature unbounded from below. The first result in this direction was obtained by Juillet in [101] for the *Heisenberg group* $\mathbb{H}^N$, building upon some results on optimal transportation in $\mathbb{H}^N$ established by Ambrosio–Rigot in the seminal paper [18]. Later, by exploiting a result of [40], Ambrosio and the author in [20] showed that any non-commutative *Carnot group* is not a $\mathsf{CD}(K, \infty)$ space. Carnot groups, of which $\mathbb{R}^N$ and $\mathbb{H}^N$ are the simplest examples, are nilpotent Lie groups that, in some sense, capture the local infinitesimal behavior of sub-Riemannian manifolds. Precisely, by a famous result of Mitchell [115], Carnot groups are the *tangent metric cones* to sub-Riemannian manifolds, see [104, Section 4.1] and the references therein. Recently, among other non-embedding results, Huang–Sun [95] proved that equiregular sub-Riemannian manifolds do not satisfy the $\mathsf{CD}(K, N)$ condition, and Juillet [103] treated the case of rank-varying distributions.

Despite their intrinsic wild nature of non-CD spaces, in the last fifteen years sub-Riemanninan manifolds have become an active and promising research topic for the study of Optimal Transport, heat and entropy flows and generalized curvature notions beyond the well-established Riemannian and CD frameworks.

After the pioneering work [18] of Ambrosio–Rigot, the well-posedness of optimal transportation was studied in Heisenberg and H-type groups [68, 76, 128], for non-holonomic distributions [4, 96] and in more general sub-Riemannian manifolds [21, 77].

The identification between the heat and the entropy semigroups was established first in the Heisenberg groups by Juillet in [102] and then in all Carnot groups by Ambrosio and the author in [20].

Several notions of *curvature* in sub-Riemannian manifolds have been introduced in recent years, following either the *Lagrangian* or the *Eulerian* approach.

The *Lagrangian* point of view has its roots in the study of Jacobi vector fields initiated in the fundamental works of Agrachev–Li–Zelenko [6, 7, 146] (ALZ for short) and later developed in [3, 32, 33, 86, 87, 117]. Besides the numerous applications inspired by some classical results of Riemannian Geometry, see [2, 5, 31, 32, 106, 130] for example, a deep and powerful byproduct of the Lagrangian approach — in the original spirit of Cordero-Erausquin–McCann–Schmuckenschläger [64] — is a precise control of the *distorsion coefficients* in the *sub-Riemannian interpolation inequalities*. These results were obtained for the first time by Balogh–Kristály–Sipos in Heisenberg and corank-1 Carnot groups [29, 30] via direct methods based on the special structure of these spaces. The link

with the ALZ theory of sub-Riemannian Jacobi fields was made manifest shortly after by Barilari–Rizzi in the more general context of *ideal structures* [35, 36].

Sub-Riemannian interpolation inequalities have two interesting consequences: the *Measure Contraction Property*, MCP for short, and the *distorted* displacement convexity of the entropy along Wasserstein geodesics.

The MCP(K, N) condition, introduced for the first time by Ohta [118], keeps track of the distortion of the volume of a set when it is transported to a Dirac delta. Although for a Riemannian manifold the MCP(K, N) and the CD(K, N) conditions are equivalent (with N the *topological* dimension of the manifold), the MCP(K, N) condition is in general weaker than the CD(K, N) condition. The first result in this direction was obtained by Juillet [101] for the Heisenberg group $\mathbb{H}^N$ (see also [60]). The same property was then proved for other Carnot groups [34, 129] and later established for more general sub-Riemannian manifolds in [35, 36].

Concerning the applications of the MCP(K, N) condition and the interpolation inequalities in the sub-Riemannian framework, we refer the reader to the work [114], where the author obtains L^p-Poincaré and log-Sobolev inequalities on domains via a quasi-convex relaxation of the CD(K, N) condition, namely, the *quasi curvature-dimension condition* QCD(Q, K, N), where $Q \in [1, +\infty)$ is an additional 'slack' parameter.

The entropy $\mathsf{Ent}_{\mathcal{H}}$ with respect to the Popp measure $\mu_{\mathcal{H}}$ of the sub-Riemannian manifolds $(\mathbb{M}, \mathsf{d}_{cc}, \mu_{\mathcal{H}})$ considered in [29, 30, 35, 36] satisfies a *distorted* displacement convexity inequality in the following sense: if $[0, 1] \in s \mapsto \mu_s \in \mathscr{P}_2(\mathbb{M})$ is the geodesic connecting two probability measures $\mu_0, \mu_1 \in \mathscr{P}_2(\mathbb{M})$ with compact support, then

$$\mathsf{Ent}_{\mathcal{H}}(\mu_s) \leq (1 - s)\,\mathsf{Ent}_{\mathcal{H}}(\mu_0) + s\,\mathsf{Ent}_{\mathcal{H}}(\mu_1) + w(s) \tag{1.12}$$

for all $s \in [0, 1]$, where $w\colon [0, 1] \to [0, +\infty)$ is a function, concave and such that $w(0) = w(1) = 1$, depending only on the lower bounds on the distorsion coefficients of $(\mathbb{M}, \mathsf{d}_{cc}, \mu_{\mathcal{H}})$ and compensating the lack of K-convexity of the function $s \mapsto \mathsf{Ent}_{\mathcal{H}}(\mu_s)$.

Although not strictly related to the present work, for the sake of completeness we warn the reader that there are other lines of research in the Lagrangian direction for the definition of curvature in the sub-Riemannian context besides the ALZ approach. We refer the interested reader to [48, 49, 93] for generalizations of the notion of connection and to [50] for the so-called *Solov'ev method*.

The *Eulerian* point of view arises from the fundamental work of Baudoin–Garofalo [44] (BG for short) and relies on a clever adaptation of the Bakry-Émery curvature-dimension inequality (1.4) to sub-Riemannian manifolds with *transverse symmetries*. Roughly said, the tangent space of the sub-Riemannian manifolds considered in [44] splits into the aforementioned subspace of *horizontal* directions $\mathcal{H}$ and a subspace of *vertical* directions $\mathcal{V}$. To this splitting, it is possible to associate two Γ-operators, the usual horizontal one $\Gamma^{\mathcal{H}}$ associated to the CC distance and the horizontal Laplacian $\Delta_{\mathcal{H}}$, and a new vertical one $\Gamma^{\mathcal{V}}$ which satisfies the commutation property

$$\Gamma^{\mathcal{H}}(f, \Gamma^{\mathcal{V}}(f)) = \Gamma^{\mathcal{V}}(f, \Gamma^{\mathcal{H}}(f)) \tag{1.13}$$

for all $f \in \mathrm{C}^{\infty}(\mathbb{M})$. Property (1.13) is typical of step 2 distributions $\mathcal{H}$, where $\mathcal{V} = [\mathcal{H}, \mathcal{H}]$ and $[\mathcal{V}, \mathcal{H}] = 0$. Defining

$$\Gamma_2^{\mathcal{H}}(f, g) = \frac{1}{2}\Big(\Delta_{\mathcal{H}}\Gamma^{\mathcal{H}}(f, g) - \Gamma^{\mathcal{H}}(f, \Delta_{\mathcal{H}}g) - \Gamma^{\mathcal{H}}(\Delta_{\mathcal{H}}f, g)\Big),$$

$$\Gamma_2^{\mathcal{V}}(f, g) = \frac{1}{2}\Big(\Delta_{\mathcal{H}}\Gamma^{\mathcal{V}}(f, g) - \Gamma^{\mathcal{V}}(f, \Delta_{\mathcal{H}}g) - \Gamma^{\mathcal{V}}(\Delta_{\mathcal{H}}f, g)\Big),$$

for all $f, g \in \mathrm{C}^\infty(\mathbb{M})$, as in the Riemannian case (1.2), the *generalized BG curvature-dimension inequality*, $\mathsf{CD}(K_\mathcal{H}, K_\mathcal{V}, \kappa, N)$ for short, with $K_\mathcal{H} \in \mathbb{R}$, $K_\mathcal{V} > 0$, $\kappa \geq 0$ and $N \in (0, +\infty]$, amounts to say that

$$\Gamma_2^\mathcal{H}(f) + \varepsilon\, \Gamma_2^\mathcal{V}(f) \geq \frac{1}{N}\, (\Delta_\mathcal{H} f)^2 + \left(K_\mathcal{H} - \frac{\kappa}{\varepsilon} \right) \Gamma^\mathcal{H}(f) + K_\mathcal{V}\, \Gamma^\mathcal{V}(f) \qquad (1.14)$$

holds for all $f \in \mathrm{C}^\infty(\mathbb{M})$ and all $\varepsilon > 0$. In (1.14), the parameter $K_\mathcal{H} \in \mathbb{R}$ plays the role of the lower bound on the 'generalized Ricci tensor' and, if $\kappa > 0$, then $\left(K_\mathcal{H} - \frac{\kappa}{\varepsilon} \right) \to -\infty$ as $\varepsilon \to 0^+$, coherently with the non-CD nature of sub-Riemannian spaces. The usual $\mathsf{CD}(K, N)$ condition is thus recovered when $\Gamma^\mathcal{V} = 0$ and $\kappa = 0$, with $K = K_\mathcal{H}$.

The BG theory has been developed in several directions, see the numerous applications obtained in [37, 39, 41–43, 46, 47] and the generalizations made in [88, 89]. We refer the reader also to the recent works [74, 75] where condition (1.13) is dropped. A simple but interesting consequence of (1.14) is the following pointwise gradient bound for the heat flow $(\mathsf{P}_t)_{t>0}$ associated to the horizontal Laplacian $\Delta_\mathcal{H}$, in analogy with (1.5): there exists $\alpha \in \mathbb{R}$ such that

$$\Gamma^\mathcal{H}(\mathsf{P}_t f) + \Gamma^\mathcal{V}(\mathsf{P}_t f) \leq e^{-2\alpha t} \left(\mathsf{P}_t \Gamma^\mathcal{H}(f) + \mathsf{P}_t \Gamma^\mathcal{V}(f) \right) \qquad (1.15)$$

for all $f \in \mathrm{C}_c^\infty(\mathbb{M})$ and $t \geq 0$, see [44, Corollary 4.6].

Pointwise gradient bounds for the heat flow $(\mathsf{P}_t)_{t>0}$ associated to the horizontal Laplacian $\Delta_\mathcal{H}$, similar to (1.15) but closer to the Riemannian one (1.5), were proved for the first time by Driver–Melcher for the Heisenberg groups [69] and later generalized to all Carnot groups by Melcher [112] (see also [90] for a different proof). Baudoin–Bonnefont obtained similar inequalities for the $\mathbb{SU}(2)$ group in [38]. Stronger inequalities have been proved for the Heisenberg groups [26, 107], H-type groups [94] and the Grushin plane [145] with different techniques, and very recently Baudoin–Kelleher treated the case of metric graphs via the theory of differential forms on Dirichlet spaces [45] (concerning Dirichlet spaces, we also refer the reader to [65] for a strictly related, although slightly weaker, pointwise gradient bound).

In all the spaces quoted above, the heat flow $(\mathsf{P}_t)_{t>0}$ satisfies an inequality of the form

$$\Gamma(\mathsf{P}_t f) \leq C e^{-2Kt} \mathsf{P}_t \Gamma(f) \qquad (1.16)$$

for all $f \in \mathrm{C}_c^\infty(\mathbb{M})$ and $t \geq 0$, for some constants $C \geq 1$ and $K \in \mathbb{R}$ (with $K = 0$ for Carnot groups, coherently with their homogeneous nature). Since (1.16) reduces to (1.5) when $C = 1$, by analogy with the CD framework the (optimal) parameter $K \in \mathbb{R}$ in (1.16) can be thought as a lower bound on the 'generalized Ricci tensor'.

Accordingly to this interpretation, thanks to the celebrated work [97] of Kuwada, the pointwise gradient bound (1.16) is equivalent to the following contractivity property of the dual heat flow $(\mathsf{H}_t)_{t>0}$ with respect to the 2-Wasserstein distance: if $\mu, \nu \in \mathscr{P}_2(\mathbb{M})$, then

$$W_2(\mathsf{H}_t \mu, \mathsf{H}_t \nu) \leq \sqrt{C}\, e^{-Kt}\, W_2(\mu, \nu) \qquad (1.17)$$

for all $t \geq 0$. Contractivity properties like (1.17) have been established also for the Markovian diffusion semigroup associated to the operator $L = \Delta + Z$ on a (connected and complete) Riemannian manifold $(\mathbb{M}, \mathrm{g})$, where Δ is the usual Laplace–Beltrami operator and Z is a C^1-regular vector field, see [63, 144].

In view of the equivalence between (1.16) and (1.17), and in analogy with the CD framework, the study of (1.16), or equivalently of (1.17), belongs to the Eulerian side of the approach to the definition of curvature in the sub-Riemannian setting.

1.4. **Sub-Riemannian groups as weak** RCD **spaces.** At the present moment, no link is known between the Lagrangian and the Eulerian approach presented above, in the sense that no relation has been shown between the distorted convexity of the entropy (1.12) and the inequalities (1.14) and (1.16) satisfied by the Γ operator and the heat flow $(P_t)_{t>0}$, in the same manner of the CD framework.

The main contribution of this work is to make a partial step towards the connection between the Langrangian and the Eulerian approach in the sub-Riemannian context. We show that the pointwise gradient bound (1.16) is equivalent to a *heated* version of the displacement convexity inequality (1.7) and an *almost-integrated* form of the Evolution Variational Inequality (1.10) in the context of metric-measure spaces with a group structure, extending to this non-CD setting the dimension-free results obtained by Ambrosio–Gigli–Savaré [13, 15] and Ambrosio–Gigli–Mondino–Rajala [11]. Since these inequalities naturally embed the corresponding ones of the CD framework, our main equivalence result can be seen as an attempt to understand the problem of the *grande unification synthétique* proposed by Villani [142] for the special case of metric-measure groups. The present work was also motivated by some questions raised by Balogh–Kristály–Sipos in [29, Section 5].

Let us give a sketch of our idea. We start by assuming that, in a metric-measure space $(X, \mathsf{d}, \mathfrak{m})$, the metric heat flow $(P_t)_{t>0}$, i.e., the gradient flow of the Cheeger energy associated to the distance d (recall (1.8) for the definition) is linear and satisfies the pointwise gradient bound

$$\Gamma(P_t f) \leq \mathsf{c}^2(t)\, P_t \Gamma(f) \quad \mathfrak{m}\text{-a.e. in } X \tag{1.18}$$

for all $t \geq 0$ and all $f \in \mathrm{Dom}(\mathsf{Ch})$, for some function $\mathsf{c} \colon [0, +\infty) \to (0, +\infty)$ locally positively bounded from above and below.

In this non-smooth context, we precisely have $\Gamma(f) = |\nabla f|_w^2$ for all $f \in \mathrm{Dom}(\mathsf{Ch})$, where $|\nabla f|_w \in \mathrm{L}^2(X, \mathfrak{m})$ is the so-called *minimal relaxed gradient of f* in the sense of [14, Definition 4.2] and represents the Cheeger energy (1.8) as

$$\mathsf{Ch}(f) = \frac{1}{2} \int_X |\nabla f|_w^2 \, \mathrm{d}\mathfrak{m}$$

for all $f \in \mathrm{Dom}(\mathsf{Ch})$. However, to avoid technicalities, in what follows we simply consider X as a sub-Riemannian manifold and Γ as the squared modulus of the gradient, $\Gamma(f) = |\nabla f|^2$.

We can think of the function c in (1.18) as the *curvature function* of the space $(X, \mathsf{d}, \mathfrak{m})$ replacing the function $t \mapsto e^{-Kt}$ of the standard pointwise gradient (1.5). Actually, thanks to the Fekete Lemma for sub-additive functions, the *optimal* curvature function c in (1.18) does satisfy $\mathsf{c}(t) \leq C e^{-Kt}$ for all $t \geq 0$ for some $C \geq 1$ and $K \in \mathbb{R}$, as for the pointwise gradient bound (1.16), provided that $\limsup_{t \to 0^+} \mathsf{c}(t) < +\infty$. The (optimal) constant $K \in \mathbb{R}$ plays the role of the lower bound on the 'generalized Ricci tensor' in this situation. We call inequality (1.18) the *weak Bakry–Émery curvature condition* with respect to the curvature function c, $\mathsf{BE}_w(\mathsf{c}, \infty)$ for short.

In this general framework, the equivalence between the pointwise gradient bound (1.18) and the W_2-contractivity property of the dual heat flow $(H_t)_{t>0}$,

$$W_2(H_t \mu, H_t \nu) \leq \mathsf{c}(t)\, W_2(\mu, \nu) \tag{1.19}$$

for all $\mu, \nu \in \mathscr{P}_2(X)$ and $t \geq 0$, has already been addressed by Ambrosio–Gigli–Savaré in [13, Section 3.2] adapting the original idea of Kuwada [97]. Actually, in [13] only the implication (1.18) $\Rightarrow$ (1.19) is proved in detail, while the other implication (1.19) $\Rightarrow$ (1.18)

— of no need for the scopes of [13] — is only stated with a sketch of its proof. However, the line suggested in [13] for the proof of this implication is not completely correct, see Remark 3.20 below for the technical details. Our first task is thus to amend the strategy of [13] and to give a self-contained and complete proof of the equivalence between (1.18) and (1.19).

Having the correspondence between the pointwise gradient bound (1.18) and the W_2-contractivity property (1.19) of the dual heat semigroup at hand, we can focus on the proof of the *almost-integrated* form of the EVI. The following (formal) computations are a sketch of the *action estimates* performed in [14, Section 4.3] for the dimension-free case $N = +\infty$. Actually, our approach takes advantage of the more general point of view assumed in [72, Section 4.2] for the finite dimensional case $N < +\infty$. For the presentation, we also took inspiration from [25, Section 6].

Let $s \mapsto \mu_s = f_s \mathfrak{m}$, $s \in [0, 1]$, be a curve in the 2-Wasserstein space joining two measures $\mu_0, \mu_1 \in \mathscr{P}_2(X)$ and let us define a new curve $s \mapsto \tilde{\mu}_s = \tilde{f}_s \mathfrak{m}$, $s \in [0, 1]$, by setting

$$\tilde{\mu}_s = \mathsf{H}_{\eta(s)} \mu_{\vartheta(s)}, \quad \text{so that } \tilde{f}_s = \mathsf{P}_{\eta(s)} f_{\vartheta(s)}, \quad \text{for all } s \in [0, 1],$$

where $\eta \in \mathrm{C}^2([0, 1]; [0, +\infty))$ and $\vartheta \in \mathrm{C}^1([0, 1]; [0, 1])$ with $\vartheta(0) = 0$ and $\vartheta(1) = 1$. At least formally, we can compute

$$\frac{\mathrm{d}}{\mathrm{d}s} \tilde{f}_s = \dot{\eta}(s) \, \Delta \mathsf{P}_{\eta(s)} f_{\vartheta(s)} + \dot{\vartheta}(s) \, \mathsf{P}_{\eta(s)} \dot{f}_{\vartheta(s)}$$

for $s \in (0, 1)$, where Δ is the infinitesimal generator of the heat flow, the (metric-measure) Laplacian operator.

On the one hand, integrating by parts, we get

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s} \mathsf{Ent}_{\mathfrak{m}}(\tilde{\mu}_s) &= \frac{\mathrm{d}}{\mathrm{d}s} \int_X \tilde{f}_s \log \tilde{f}_s \, \mathrm{d}\mathfrak{m} \\
&= \int_X (1 + \log \tilde{f}_s) \frac{\mathrm{d}}{\mathrm{d}s} \tilde{f}_s \, \mathrm{d}\mathfrak{m} \\
&= -\dot{\eta}(s) \int_X p'(\tilde{f}_s) \, \Gamma(\tilde{f}_s) \, \mathrm{d}\mathfrak{m} + \dot{\vartheta}(s) \int_X p(\tilde{f}_s) \, \mathsf{P}_{\eta(s)} \dot{f}_{\vartheta(s)} \, \mathrm{d}\mathfrak{m}
\end{aligned}$$

for $s \in (0, 1)$, where $p(r) = 1 + \log r$ for all $r > 0$. Observing that $p'(r) = r(p'(r))^2$ for all $r > 0$, by the *chain rule* $\Gamma(\varphi(f)) = (\varphi'(f))^2 \, \Gamma(f)$ valid for all $\varphi \colon \mathbb{R} \to \mathbb{R}$ sufficiently smooth and all $f \in \mathrm{Dom}(\mathsf{Ch})$, we can write

$$\frac{\mathrm{d}}{\mathrm{d}s} \mathsf{Ent}_{\mathfrak{m}}(\tilde{\mu}_s) = -\dot{\eta}(s) \int_X \Gamma(g_s) \, \mathrm{d}\tilde{\mu}_s + \dot{\vartheta}(s) \int_X \dot{f}_{\vartheta(s)} \, \mathsf{P}_{\eta(s)} g_s \, \mathrm{d}\mathfrak{m} \tag{1.20}$$

for $s \in (0, 1)$, where we have set $g_s = p(\tilde{f}_s)$ for brevity.

On the other hand, by Kantorovich duality, we can write

$$\frac{1}{2} W_2^2(\mu, \nu) = \sup \left\{ \int_X Q_1 \varphi \, \mathrm{d}\mu - \int_X \varphi \, \mathrm{d}\nu : \varphi \in \mathrm{Lip}(X) \text{ with bounded support} \right\}, \tag{1.21}$$

where

$$Q_s \varphi(x) = \inf_{y \in X} \varphi(y) + \frac{\mathsf{d}^2(y, x)}{2s}, \quad \text{for } x \in X \text{ and } s > 0,$$

is the *Hopf–Lax infimum-convolution semigroup*. Recalling that $\varphi_s = Q_s\varphi$ solves the Hamilton–Jacobi equation $\partial_s\varphi_s + \frac{1}{2}|\nabla\varphi_s|^2 = 0$, again integrating by parts we can compute

$$\frac{\mathrm{d}}{\mathrm{d}s}\int_X \varphi_s\,\tilde{f}_s\,\mathrm{d}\mathfrak{m} = \int_X \partial_s\varphi_s\,\mathrm{d}\tilde{\mu}_s + \int_X \varphi_s\,\frac{\mathrm{d}}{\mathrm{d}s}\tilde{f}_s\,\mathrm{d}\mathfrak{m}$$

$$= -\frac{1}{2}\int_X \Gamma(\varphi_s)\,\mathrm{d}\tilde{\mu}_s - \dot{\eta}(s)\int_X \Gamma(\varphi_s,\tilde{f}_s)\,\mathrm{d}\mathfrak{m} + \dot{\vartheta}(s)\int_X \dot{f}_{\vartheta(s)}\,\mathsf{P}_{\eta(s)}\varphi_s\,\mathrm{d}\mathfrak{m} \tag{1.22}$$

for $s \in (0,1)$. We can combine (1.20) and (1.22) to get

$$\frac{\mathrm{d}}{\mathrm{d}s}\int_X \varphi_s\,\tilde{f}_s\,\mathrm{d}\mathfrak{m} + \dot{\eta}(s)\frac{\mathrm{d}}{\mathrm{d}s}\,\mathsf{Ent}_{\mathfrak{m}}(\tilde{\mu}_s) \leq -\frac{1}{2}\int_X \left(\Gamma(\varphi_s) + \dot{\eta}(s)^2\,\Gamma(g_s)\right)\mathrm{d}\tilde{\mu}_s$$

$$- \dot{\eta}(s)\int_X \Gamma(\varphi_s,\tilde{f}_s)\,\mathrm{d}\mathfrak{m} \tag{1.23}$$

$$+ \dot{\vartheta}(s)\int_X \dot{f}_{\vartheta(s)}\,\mathsf{P}_{\eta(s)}(\varphi_s + \dot{\eta}(s)\,g_s)\,\mathrm{d}\mathfrak{m}$$

for $s \in (0,1)$, forgetting the term $-\frac{\dot{\eta}(s)^2}{2}\int_X \Gamma(g_s)\,\mathrm{d}\tilde{\mu}_s \leq 0$ in (1.20). Now

$$\Gamma(\varphi_s + \dot{\eta}(s)\,g_s) = \Gamma(\varphi_s) + 2\,\dot{\eta}(s)\,\Gamma(\varphi_s,g_s) + \dot{\eta}(s)^2\,\Gamma(g_s)$$

and, by the chain rule,

$$\Gamma(\varphi_s,g_s) = \Gamma(\varphi_s,p(\tilde{f}_s)) = p'(\tilde{f}_s)\,\Gamma(\varphi_s,\tilde{f}_s).$$

Since $r\,p'(r) = 1$, we have

$$\int_X \Gamma(\varphi_s,g_s)\,\mathrm{d}\tilde{\mu}_s = \int_X \tilde{f}_s\,p'(\tilde{f}_s)\,\Gamma(\varphi_s,\tilde{f}_s)\,\mathrm{d}\mathfrak{m} = \int_X \Gamma(\varphi_s,\tilde{f}_s)\,\mathrm{d}\mathfrak{m},$$

and thus (1.23) simplifies to

$$\frac{\mathrm{d}}{\mathrm{d}s}\int_X \varphi_s\,\tilde{f}_s\,\mathrm{d}\mathfrak{m} + \dot{\eta}(s)\frac{\mathrm{d}}{\mathrm{d}s}\,\mathsf{Ent}_{\mathfrak{m}}(\tilde{\mu}_s) \leq -\frac{1}{2}\int_X \Gamma(\varphi_s + \dot{\eta}(s)\,g_s)\,\mathrm{d}\tilde{\mu}_s$$

$$+ \dot{\vartheta}(s)\int_X \dot{f}_{\vartheta(s)}\,\mathsf{P}_{\eta(s)}(\varphi_s + \dot{\eta}(s)\,g_s)\,\mathrm{d}\mathfrak{m} \tag{1.24}$$

for $s \in (0,1)$. At this point, the crucial information we need to know about the chosen curve $s \mapsto \mu_s = f_s\mathfrak{m}$ is that

$$\int_X \dot{f}_s\,\psi\,\mathrm{d}\mathfrak{m} \leq |\dot{\mu}_s|\left(\int_X \Gamma(\psi)\,\mathrm{d}\mu_s\right)^{\frac{1}{2}} \tag{1.25}$$

for all sufficiently 'nice' functions $\psi \in \mathrm{Dom}(\mathsf{Ch})$, where $|\dot{\mu}_s| = \lim_{h\to 0}\frac{W_2(\mu_{s+h},\mu_s)}{h}$, $s \in (0,1)$, is the *metric velocity* of the curve $s \mapsto \mu_s$ with respect to the 2-Wasserstein distance.

With (1.25) at disposal, we may choose $\psi = \mathsf{P}_{\eta(s)}(\varphi_s + \dot{\eta}(s)\, g_s)$ and estimate

$$\dot{\vartheta}(s) \int_X \dot{f}_{\vartheta(s)}\, \mathsf{P}_{\eta(s)}(\varphi_s + \dot{\eta}(s)\, g_s)\, \mathrm{d}\mathfrak{m} = \int_X \left(\frac{\mathsf{d}}{\mathrm{d}s}\, f_{\vartheta(s)}\right) \mathsf{P}_{\eta(s)}(\varphi_s + \dot{\eta}(s)\, g_s)\, \mathrm{d}\mathfrak{m}$$

$$\leq |\dot{\vartheta}(s)|\, |\dot{\mu}_{\vartheta(s)}| \left(\int_X \Gamma(\mathsf{P}_{\eta(s)}(\varphi_s + \dot{\eta}(s)\, g_s))\, \mathrm{d}\mu_s\right)^{\frac{1}{2}}$$

$$\leq \frac{\mathsf{c}^2(\eta(s))}{2}\, \dot{\vartheta}(s)^2\, |\dot{\mu}_{\vartheta(s)}|^2 + \frac{\mathsf{c}^{-2}(\eta(s))}{2} \int_X \Gamma(\mathsf{P}_{\eta(s)}(\varphi_s + \dot{\eta}(s)\, g_s))\, \mathrm{d}\mu_s$$

$$\leq \frac{\mathsf{c}^2(\eta(s))}{2}\, \dot{\vartheta}(s)^2\, |\dot{\mu}_{\vartheta(s)}|^2 + \frac{1}{2} \int_X \Gamma(\varphi_s + \dot{\eta}(s)\, g_s)\, \mathrm{d}\tilde{\mu}_s$$

$$\tag{1.26}$$

by Young inequality, in virtue of (1.18). By combining (1.24) with (1.26), we conclude that

$$\frac{\mathrm{d}}{\mathrm{d}s} \int_X \varphi_s\, \tilde{f}_s\, \mathrm{d}\mathfrak{m} + \dot{\eta}(s)\, \frac{\mathrm{d}}{\mathrm{d}s}\, \mathsf{Ent}_{\mathfrak{m}}(\tilde{\mu}_s) \leq \frac{\mathsf{c}^2(\eta(s))}{2}\, \dot{\vartheta}(s)^2\, |\dot{\mu}_{\vartheta(s)}|^2$$

for $s \in (0,1)$. If we choose $\dot{\vartheta}(s) = \mathsf{c}^{-2}(\eta(s))$, then we can integrate in $s \in (0,1)$ so that, by Kantorovich duality (1.21), we finally get

$$\frac{1}{2}\, W_2^2(\mathsf{H}_{\eta(1)}\mu_1, \mathsf{H}_{\eta(0)}\mu_0) - \frac{1}{2\,\mathrm{R}(\eta)}\, W_2^2(\mu_1, \mu_0) + \dot{\eta}(1)\, \mathsf{Ent}_{\mathfrak{m}}(\mathsf{H}_{\eta(1)}\mu_1)$$

$$\leq \dot{\eta}(0)\, \mathsf{Ent}_{\mathfrak{m}}(\mathsf{H}_{\eta(0)}\mu_0) + \int_0^1 \ddot{\eta}(s)\, \mathsf{Ent}_{\mathfrak{m}}(\mathsf{H}_{\eta(s)}\mu_{\vartheta(s)})\, \mathrm{d}s,$$

$$\tag{1.27}$$

where $\mathrm{R}(\eta) = \int_0^1 \mathsf{c}^{-2}(\eta(s))\, \mathrm{d}s$.

Note that (1.27) is actually equivalent to the pointwise gradient bound (1.18). Indeed, the choice of the constant function $\eta(s) = t$ for all $s \in [0,1]$ immediately gives (1.19). Moreover, if $\mathsf{c}(t) = e^{-Kt}$ for some $K \in \mathbb{R}$, then we recover (1.10) by choosing $\eta(s) = st$ for all $s \in [0,1]$. Indeed, in this case, we obtain

$$\frac{1}{2}\, W_2^2(\mathsf{H}_t\mu_1, \mu_0) - \frac{Kt}{e^{2Kt} - 1}\, W_2^2(\mu_1, \mu_0) + t\Big(\mathsf{Ent}_{\mathfrak{m}}(\mathsf{H}_t\mu_1) - \mathsf{Ent}_{\mathfrak{m}}(\mu_0)\Big) \leq 0$$

and (1.10) follows (for $t = 0$, which is enough thanks to the semigroup property) by observing that $\frac{Kt}{e^{2Kt}-1} = \frac{1}{2}(1 - Kt + o(t))$ as $t \to 0^+$. For this reason, and adopting the same terminology of [67, Proposition 3.1], we may think of (1.27) as an *almost-integrated* form of the EVI (1.10).

Since we have no information about the behavior of the function $s \mapsto \mathsf{Ent}_{\mathfrak{m}}(\mathsf{H}_{\eta(s)}\mu_{\vartheta(s)})$, to simplify (1.27) it is convenient to choose $\eta(s) = (1-s)t_0 + st_1$ for $s \in [0,1]$, where $0 \leq t_0 \leq t_1$ are fixed. With this choice, inequality (1.27) reduces to

$$W_2^2(\mathsf{H}_{t_1}\mu_1, \mathsf{H}_{t_0}\mu_0) + 2(t_1 - t_0)\Big(\mathsf{Ent}_{\mathfrak{m}}(\mathsf{H}_{t_1}\mu_1) - \mathsf{Ent}_{\mathfrak{m}}(\mathsf{H}_{t_0}\mu_0)\Big) \leq \mathsf{A}[t_0; t_1]^{-1}\, W_2^2(\mu_1, \mu_0),$$

$$\tag{1.28}$$

where $\mathsf{A}[t_0; t_1] = \int_{t_0}^{t_1} \mathsf{c}^{-2}(s)\, \mathrm{d}s$. In analogy with (1.10), we call the above inequality (1.28) the *weak Evolution Variational Inequality* with respect to the curvature function c, $\mathsf{EVI}_w(\mathsf{c})$ for short.

Arguing exactly as in the proof of [67, Theorem 3.2], from (1.28) we deduce that, if $s \mapsto \mu_s$ is a 2-Wasserstein (constant speed) geodesic, then

$$\mathsf{Ent}_{\mathfrak{m}}(\mathsf{H}_{t+h}\mu_s) \leq (1-s)\,\mathsf{Ent}_m(\mathsf{H}_t\mu_0) + s\,\mathsf{Ent}_{\mathfrak{m}}(\mathsf{H}_t\mu_1)$$
$$+ \frac{s(1-s)}{2h}\left(\mathsf{A}[t;t+h]^{-1}\,W_2^2(\mu_1,\mu_0) - W_2^2(\mathsf{H}_t\mu_1,\mathsf{H}_t\mu_0)\right) \tag{1.29}$$

for all $t \geq 0$ and $h > 0$. Note that (1.29) is still equivalent to the pointwise gradient bound (1.18) since, by multiplying both of its sides by $h > 0$ and then letting $h \to 0^+$, we again recover (1.19). Moreover, if we choose $t = 0$ in (1.29), then we obtain

$$\mathsf{Ent}_{\mathfrak{m}}(\mathsf{H}_h\mu_s) \leq (1-s)\,\mathsf{Ent}_m(\mu_0) + s\,\mathsf{Ent}_{\mathfrak{m}}(\mu_1) + \frac{\mathsf{B}[h]}{2}\,s(1-s)\,W_2^2(\mu_1,\mu_0) \tag{1.30}$$

for all $h > 0$, where $\mathsf{B}[h] = \dfrac{\mathsf{A}[0,h]^{-1} - 1}{h}$. In particular, if $\mathsf{c}(t) = e^{-Kt}$ then $\mathsf{B}[h] = -K + o(1)$ as $h \to 0^+$, so that we immediately recover the displacement K-convexity (1.7). For this reason, we may think of (1.29) as a *heated* version of the displacement convexity of the entropy and we call it the (dimension-free) *weak Riemannian Curvature-Dimension condition* with respect to the curvature function c, $\mathsf{RCD}_w(\mathsf{c},\infty)$ for short.

Inequality (1.30) is very close to the distorted convexity inequality (1.12) but, at the same time, it reflects the idea behind the generalized Baudoin–Garofalo CD condition (1.14), in the sense that $\mathsf{B}[h] \to +\infty$ as $h \to 0^+$ in the non-CD setting, coherently with the fact that sub-Riemannian spaces have Ricci curvature unbounded from below.

The explosion of the right-hand side of (1.30) as $h \to 0^+$ can be interpreted also in the light of the *singularity problem* of 2-Wasserstein geodesics, still open for a general sub-Riemannian manifold: if $\mu_0 \ll \mathfrak{m}$, then does any 2-Wasserstein geodesic $s \mapsto \mu_s$, $s \in [0,1]$, joining $\mu_0, \mu_1 \in \mathscr{P}_2(X)$ still satisfy $\mu_s \ll \mathfrak{m}$? This problem was posed for the first time by Ambrosio–Rigot [18] for the Heisenberg group and positively answered by Figalli–Juillet [76]. Later Figalli–Rifford [77] gave an affirmative answer also for more general sub-Riemannian manifolds (see also [35]). As pointed out by Cavalletti–Mondino [61], the answer is still positive if the ambient metric-measure space satisfies the $\mathsf{MCP}(K,N)$ condition and is *essentially non-branching*, a condition roughly saying that *branching* geodesics, i.e., geodesics splitting at intermediate times, are not too many. Note that some sub-Riemannian spaces do have branching geodesics, see [113]. Thus, in this sense, considering the heated version of the 2-Wasserstein geodesic in (1.29) can be seen as a way to bypass its possible singularity.

All in all, apart from technicalities, if we can construct sufficiently good 2-Wasserstein curves $s \mapsto \mu_s = f_s\mathfrak{m}$ satisfying (1.25), then, under the linearity of the heat flow, for the metric-measure space $(X,\mathsf{d},\mathfrak{m})$ we have the equivalences

$$\mathsf{BE}_w(\mathsf{c},\infty) \iff \mathsf{EVI}_w(\mathsf{c}) \iff \mathsf{RCD}_w(\mathsf{c},\infty).$$

Therefore, in analogy with the CD setting, we may call such a metric-measure space $(X,\mathsf{d},\mathfrak{m})$ a *weak RCD-space*.

Property (1.25) can be obtained from a celebrated result of Lisini [108], so that the central problem we need to face for the construction of the curve $s \mapsto \mu_s$ is the absolute continuity property $\mu_s \ll \mathfrak{m}$. Due to the aforementioned singularity problem in this general framework, we cannot choose $s \mapsto \mu_s$ to be a geodesic. Nevertheless, we may choose $s \mapsto \mu_s$ to be a suitable *regularization* $s \mapsto \mu_s^\varepsilon$, for all $\varepsilon > 0$, of a geodesic (or of any other probability curve realizing the 2-Wasserstein distance between μ_0 and μ_1 up to a smaller and smaller error). However, since we need the Lisini inequality (1.25),

the regularized curve $s \mapsto \mu_s^\varepsilon$ has to have 2-Wasserstein metric velocity controlled by the velocity of the original curve.

In [13, 72], the regularized procedure takes advantage of the smoothing property of the heat flow and, precisely, leads to the choice $\mu_s^\varepsilon = \mathsf{H}_\varepsilon \mu_s$. Indeed, on the one hand, the pointwise gradient bound (1.5) implies the *instantaneous diffusion property* $\mathsf{H}_t \mu \ll \mathfrak{m}$ for all $t > 0$ and $\mu \in \mathscr{P}_2(X)$. On the other hand, the W_2-contractivity property (1.6) immediately gives $|\dot{\mu}_s^\varepsilon| \leq e^{-K\varepsilon} |\dot{\mu}_s|$.

Under the weaker $\mathsf{BE}(\mathsf{c}, \infty)$ condition (1.18), it is still possible to prove the instantaneous diffusion property. However, for the choice $\mu_s^\varepsilon = \mathsf{H}_\varepsilon \mu_s$, the W_2-contractivity property (1.19) only gives the weaker estimate $|\dot{\mu}_s^\varepsilon| \leq \mathsf{c}(\varepsilon) |\dot{\mu}_s|$, which is of no use unless $\lim_{t \to 0^+} \mathsf{c}(t) = 1$, a property the curvature function does not satisfy for the sub-Riemannian manifolds under consideration.

Since we cannot rely on the sole properties of the heat flow, it is at this point that we assume that the ambient space X has a group operation left-compatible with the metric-measure structure and exploit the property of convolution. In fact, under this additional assumption, we may choose the regularized curve as the left-convoluted curve $\mu_s^\varepsilon = (\varrho_\varepsilon \star \mu_s)\mathfrak{m}$, where $(\varrho_\varepsilon)_{\varepsilon > 0} \subset \mathrm{L}^1(X, \mathfrak{m})$ is a suitable family of convolution kernels. Since the group operation is left-compatible with the metric, it is not difficult to prove that the 2-Wasserstein metric velocity of the left-convoluted curve does not increase, i.e. $|\dot{\mu}_s^\varepsilon| \leq |\dot{\mu}_s|$ (a property not expected for right-convoluted curves, see Remark 5.14 below). We can thus perform all the above computations on the left-convoluted curve $s \mapsto \mu_s^\varepsilon$ and obtain the desired entropic inequalities by passing to the limit as $\varepsilon \to 0^+$ at the end.

Although the present work is focused only on (metric-measure) groups, we believe that our results may be valid also for other non-group spaces, such as the Grushin plane [145] and metric graphs [45], where the regularization of 2-Wasserstein curves could possibly be performed by exploiting the particular structure of the underlying space. Another interesting problem is whether some sub-Riemmanian manifolds (possibly, with a group structure) may satisfy a more precise form of the pointwise gradient bound (1.16) also taking into account a dimensional parameter $N \in (0, +\infty)$. Finally, from a purely metric-measure point of view, on the one side we do not know if the weak RCD condition (1.29) may imply (a weaker form of) the MCP condition and, on the other side, if the weak EVI (1.28) may be useful for the definition of a weaker notion of metric gradient flow and/or for proving stability properties of weak RCD spaces under Gromov–Hausdorff convergence. We will hopefully come again over these and related topics in a future work.

1.5. **Organization of the paper.** The paper is organized as follows.

In Section 2 we recall all the known definitions and results in the metric-measure setting we will use in the sequel, in order to keep the paper the most self-contained as possible. Almost all the theorems are stated without proofs, but we give the reader precise references to the existing literature where to find the needed technical details.

In Section 3 we introduce the $\mathsf{BE}_w(\mathsf{c}, \infty)$ condition in the metric-measure framework and study its consequences, such as Poincaré inequalities and the definition and the properties of the dual and the pointwise version of the heat semigroup. The main result of this part is the equivalence between the $\mathsf{BE}_w(\mathsf{c}, \infty)$ condition and the weak W_2-contractivity property of the dual heat flow, the so-called Kuwada duality.

In Section 4 we deal with the Fisher information and the $\mathrm{L}\log\mathrm{L}$-regularization property. The results, which will be frequently used in the remaining part of the paper, are

technical and provide a generalization of the known theory to the context of the $\mathsf{BE}_w(\mathsf{c}, \infty)$ condition.

Finally, in Section 5, we prove our main equivalence result for metric-measure groups, see Theorem 5.16 for the precise statement. In the last part of this section, we show how this theorem applies to Carnot groups and to the $\mathbb{SU}(2)$ group. Also, we briefly compare the heated displacement convexity (1.30) we obtain with the distorted displacement convexity (1.12) in the case of 2-Wasserstein geodesics induced by right-translation optimal transport maps.

2. Preliminaries

In this section, we recall the main technical tools we will use throughout the paper. For a more detailed exposition of the results presented below, we refer the reader to [8, 10–17, 22, 24, 57, 62, 82, 83, 140, 141]. At the end of this section, we summarize the main assumptions we will use in the rest of the present work. For the reader's convenience, we will try to keep the paper the most self-contained as possible.

2.1. **AC curves.** Let (X, d) be a metric space. Let $I \subset \mathbb{R}$ be a closed interval and let $p \in [1, +\infty]$. We say that a curve $\gamma \colon I \to X$ belongs to $\mathrm{AC}^p(I; X)$ if

$$\mathsf{d}(\gamma_s, \gamma_t) \leq \int_s^t g(r)\,\mathrm{d}r \qquad s, t \in I, \ s < t, \tag{2.1}$$

for some $g \in \mathrm{L}^p(I)$. The space $\mathrm{AC}^p_{\mathrm{loc}}(I; X)$ is defined analogously. The exponent $p = 1$ corresponds to *absolutely continuous curves* and is simply denoted by $\mathrm{AC}(I; X)$. It turns out that, if $\gamma \in \mathrm{AC}^p(I; X)$, then there is a minimal function $g \in \mathrm{L}^p(I)$ satisfying (2.1), called *metric derivative* of the curve γ, which is given by

$$|\dot{\gamma}_t| = \lim_{s \to t} \frac{\mathsf{d}(\gamma_s, \gamma_t)}{|s - t|} \qquad \text{for } \mathscr{L}^1\text{-a.e. } t \in I,$$

see [12, Theorem 1.1.2] for the simple proof. We thus say that an absolutely continuous curve γ has *constant speed* if $t \mapsto |\dot{\gamma}_t|$ is (equivalent to) a constant.

We say that (X, d) is a *length* (metric) space if for all $x, y \in X$ we have

$$\mathsf{d}(x, y) = \inf\left\{ \int_0^1 |\dot{\gamma}_t|\,\mathrm{d}t : \gamma \in \mathrm{AC}([0, 1]; X), \ \gamma_0 = x, \ \gamma_1 = y \right\}.$$

In addition, we call (X, d) a *geodesic* metric space if for every $x, y \in X$ there exists a curve $\gamma \colon [0, 1] \to X$ such that $\gamma_0 = x$, $\gamma_1 = y$ and

$$\mathsf{d}(\gamma_s, \gamma_t) = |s - t|\,\mathsf{d}(\gamma_0, \gamma_1) \qquad \forall s, t \in [0, 1].$$

In this case, we say that the curve $\gamma \colon [0, 1] \to X$ is a (*constant*) *unit-speed geodesic* and we write $s \mapsto \gamma_s \in \mathrm{Geo}([0, 1]; X)$.

2.2. **Slopes.** Let (X, d) be a metric space. Let $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ and let $f \colon X \to \overline{\mathbb{R}}$ be a function. We define the *effective domain* of f as

$$\mathrm{Dom}(f) = \{x \in X : f(x) \in \mathbb{R}\}.$$

Given $x \in \mathrm{Dom}(f)$, we define the *slope* and the *asymptotic Lipschitz constant* of f at x by

$$|\mathrm{D}f|(x) = \limsup_{y \to x} \frac{|f(y) - f(x)|}{\mathsf{d}(x, y)}, \qquad |\mathrm{D}^* f|(x) = \limsup_{\substack{y, z \to x \\ y \neq z}} \frac{|f(y) - f(z)|}{\mathsf{d}(y, z)} \tag{2.2}$$

The *descending slope* and the *ascending slope* of f at x are respectively given by

$$|\mathrm{D}^- f|(x) = \limsup_{y \to x} \frac{[f(y) - f(x)]^-}{\mathsf{d}(x,y)}, \qquad |\mathrm{D}^+ f|(x) = \limsup_{y \to x} \frac{[f(y) - f(x)]^+}{\mathsf{d}(x,y)}.$$

Here and in the following, a^+ and a^- denote the positive and negative part of $a \in \mathbb{R}$ respectively. When $x \in \mathrm{Dom}(f)$ is an isolated point of X, we set $|\mathrm{D}f|(x) = |\mathrm{D}^* f|(x) = |\mathrm{D}^- f|(x) = |\mathrm{D}^+ f|(x) = 0$. By convention, we set $|\mathrm{D}f|(x) = |\mathrm{D}^* f|(x) = |\mathrm{D}^- f|(x) = |\mathrm{D}^+ f|(x) = +\infty$ for all $x \in X \setminus \mathrm{Dom}(f)$. Clearly, $|\mathrm{D}f| \le |\mathrm{D}^* f|$ on X and the asymptotic Lipschitz constant $|\mathrm{D}^* f|: X \to [0, +\infty]$ is an upper semicontinuous function. Note that the slopes of a Borel function $f: X \to \overline{\mathbb{R}}$ are *universally* measurable, see [14, Lemma 2.6].

According to [62] (see also [14, Section 2.3]), we say that a function $g: X \to [0, +\infty]$ is an *upper gradient* of $f: X \to \overline{\mathbb{R}}$ if, for any curve $\gamma \in \mathrm{AC}([0,1]; (\mathrm{Dom}(f), \mathsf{d}))$, $s \mapsto g(\gamma_s) |\dot{\gamma}_s|$ is measurable in $[0,1]$ (with the convention $0 \cdot \infty = 0$) and

$$|f(\gamma_1) - f(\gamma_0)| \le \int_0^1 g(\gamma_s) |\dot{\gamma}_s| \, \mathrm{d}s. \tag{2.3}$$

If $f \in \mathrm{Lip}(X)$, then $|\mathrm{D}f|$, $|\mathrm{D}^* f|$, $|\mathrm{D}^- f|$ and $|\mathrm{D}^+ f|$ are upper gradients of f, see [14, Remark 2.8]. In addition, if (X, d) is a length space, then

$$|\mathrm{D}^* f|(x) = \limsup_{y \to x} |\mathrm{D}f|(y), \qquad \mathrm{Lip}(f) = \sup_{x \in X} |\mathrm{D}f|(x) = \sup_{x \in X} |\mathrm{D}^* f|(x),$$

see [16, Section 3.1]. In particular, $|\mathrm{D}^* f|$ is the upper semicontinuous envelope of $|\mathrm{D}f|$.

2.3. Hopf–Lax semigroup.

Let (X, d) be a length space. For all $s > 0$, the *Hopf–Lax semigroup* $Q_s: \mathrm{C}_b(X) \to \mathrm{C}_b(X)$ is given by

$$Q_s f(x) = \inf_{y \in X} f(y) + \frac{\mathsf{d}^2(y,x)}{2s} \quad \text{for all } x \in X \text{ and } f \in \mathrm{C}_b(X). \tag{2.4}$$

By convention, we set $Q_0 f = f$ for all $f \in \mathrm{C}_b(X)$. If $f \in \mathrm{C}_b(X)$, then

$$\frac{\mathrm{d}^+}{\mathrm{d}s} Q_s f(x) + \frac{1}{2} |\mathrm{D}Q_s f|^2(x) = 0 \tag{2.5}$$

for all $s > 0$ and $x \in X$, see [14, Theorem 3.6]. If $f \in \mathrm{Lip}_b(X)$, then we also have

$$\mathrm{Lip}(Q_s f) \le 2 \, \mathrm{Lip}(f) \quad \text{and} \quad \mathrm{Lip}(Q_. f(x)) \le 2 \, \mathrm{Lip}(f)^2 \tag{2.6}$$

for all $s \ge 0$ and $x \in X$, see the discussion in [16, Section 3.1]. In addition, by [14, Propositon 3.2 and Theorem 3.6], for all $s > 0$ the slope $x \mapsto |\mathrm{D}Q_s f|(x)$ is upper semicontinuous, so that

$$|\mathrm{D}Q_s f|(x) = |\mathrm{D}^* Q_s f|(x) \tag{2.7}$$

for all $s > 0$ and $x \in X$.

2.4. Wasserstein space.

Let (X, d) be a complete and separable (*Polish*, for short) metric space. We denote by $\mathscr{P}(X)$ the set of probability Borel measures on X. Given $p \in [1, +\infty)$, the *p-Wasserstein (extended) distance* between $\mu, \nu \in \mathscr{P}(X)$ is given by

$$W_p^p(\mu, \nu) = \inf\left\{ \int_{X \times X} \mathsf{d}^p(x,y) \, \mathrm{d}\pi : \pi \in \mathsf{Plan}(\mu, \nu) \right\} \in [0, +\infty], \tag{2.8}$$

where

$$\mathsf{Plan}(\mu, \nu) = \{ \pi \in \mathscr{P}(X \times X) : (p_1)_\sharp \pi = \mu, \ (p_2)_\sharp \pi = \nu \}. \tag{2.9}$$

Here $p_i: X \times X \to X$, $i = 1, 2$, denote the the canonical projections on the components. As usual, if $\mu \in \mathscr{P}(X)$ and $T: X \to Y$ is a μ-measurable map with values in the topological

space Y, the *push-forward measure* $T_\sharp(\mu) \in \mathscr{P}(Y)$ is defined by $T_\sharp(\mu)(B) = \mu(T^{-1}(B))$ for every Borel set $B \subset Y$. The set $\mathsf{Plan}(\mu, \nu)$ introduced in (2.9) is call the set of *admissible plans* or *couplings* for the pair (μ, ν). Since the metric space (X, d) is complete and separable, there exist optimal couplings where the infimum in (2.8) is achieved.

The function W_p is a finite distance on the so-called *p-Wasserstein space* $(\mathscr{P}_p(X), W_p)$, where

$$\mathscr{P}_p(X) = \left\{ \mu \in \mathscr{P}(X) : \int_X \mathsf{d}^p(x, x_0)\, \mathrm{d}\mu(x) < +\infty \text{ for some, and thus any, } x_0 \in X \right\}.$$

The space $(\mathscr{P}_p(X), W_p)$ is complete and separable. If (X, d) is geodesic, then $(\mathscr{P}_p(X), W_p)$ is geodesic as well. Moreover, $\mu_n \xrightarrow{W_p} \mu$ as $n \to +\infty$ if and only if $\mu_n \rightharpoonup \mu$ as $n \to +\infty$ and

$$\lim_{n \to +\infty} \int_X \mathsf{d}^p(x, x_0)\, \mathrm{d}\mu_n(x) = \int_X \mathsf{d}^p(x, x_0)\, \mathrm{d}\mu(x) \quad \text{for some } x_0 \in X.$$

As usual, we write $\mu_n \rightharpoonup \mu$ as $n \to +\infty$, and we say that μ_n *weakly converges to μ* as $n \to +\infty$, if we have

$$\lim_{n \to +\infty} \int_X \varphi\, \mathrm{d}\mu_n = \int_X \varphi\, \mathrm{d}\mu \quad \text{for all } \varphi \in \mathrm{C}_b(X).$$

Given $p \in [1, +\infty)$, the p-Wasserstein distance can be equivalently obtained via the *Kantorovich duality formula*

$$\frac{1}{p} W_p^p(\mu, \nu) = \sup\left\{ \int_X \varphi^c\, \mathrm{d}\mu - \int_X \varphi\, \mathrm{d}\nu : \varphi \in \mathrm{Lip}_b(X) \right\} \in [0, +\infty] \qquad (2.10)$$

for all $\mu, \nu \in \mathscr{P}(X)$, where

$$\varphi^c(x) = \inf_{y \in X} \varphi(y) + \frac{\mathsf{d}^p(y, x)}{p}, \quad \text{for all } x \in X, \qquad (2.11)$$

is the *c-conjugate* of $\varphi \in \mathrm{Lip}_b(X)$ with respect to the cost function $c = \mathsf{d}^p/p$. In particular, if $p = 1$ then (2.11) immediately gives $\varphi^c = \varphi$ and thus we can rewrite (2.10) as

$$W_1(\mu, \nu) = \sup\left\{ \int_X \varphi\, \mathrm{d}(\mu - \nu) : \varphi \in \mathrm{Lip}(X) \text{ with } \mathrm{Lip}(\varphi) \le 1 \text{ and bounded support} \right\} \tag{2.12}$$

for all $\mu, \nu \in \mathscr{P}_1(X)$, the so-called *Kantorovich–Rubinstein formula*, see [141, Particular Case 5.16]. If instead $p = 2$, then by (2.4) we have $\varphi^c = Q_1\varphi$ and thus we can rewrite (2.10) as

$$\frac{1}{2} W_2^2(\mu, \nu) = \sup\left\{ \int_X Q_1\varphi\, \mathrm{d}\mu - \int_X \varphi\, \mathrm{d}\nu : \varphi \in \mathrm{Lip}(X) \text{ with bounded support} \right\} \quad (2.13)$$

for all $\mu, \nu \in \mathscr{P}(X)$. Note that the integral expressions appearing in the right-hand sides of (2.12) and (2.13) are invariant by adding constants to φ, so that we can additionally assume $\varphi \ge 0$ without changing the suprema.

For an account on Kantorovich duality, we refer the reader to [140, Section 1.1.2] and [141, Theorem 5.10] (see also the discussion in [97, Remark 3.6]).

Finally, given a non-negative Borel reference measure $\mathfrak{m}$ on X, finite on bounded sets and such that $\mathrm{supp}(\mathfrak{m}) = X$, for $p \in [1, +\infty)$ we let

$$\mathscr{P}^{\mathrm{ac}}(X) = \{\mu \in \mathscr{P}(X) : \mu \ll \mathfrak{m}\}, \quad \mathscr{P}_p^{\mathrm{ac}}(X) = \{\mu \in \mathscr{P}^{\mathrm{ac}}(X) : \mu \in \mathscr{P}_p(X)\}.$$

Thanks to [141, Theorem 6.18], $\mathscr{P}_p^{\mathrm{ac}}(X)$ is a W_p-dense subset of $\mathscr{P}_p(X)$.

For a proof of the above results and as well as for an agile introduction to the Wasserstein distance, we refer the reader to [10, Section 3], [140, Chapter 1] and [141, Chapters 4–6].

2.5. **Entropy.** Let $(X, \mathsf{d}, \mathfrak{m})$ be a metric-measure space, i.e. (X, d) is a Polish metric space and

$$\mathfrak{m} \text{ is a non-negative Borel-regular measure, finite} \atop \text{on bounded sets and such that supp } \mathfrak{m} = X. \tag{2.14}$$

Note that, in particular, $\mathfrak{m}$ is a Radon measure on X, see [91, Proposition 3.3.44]. In addition, assume that

$$\exists x_0 \in X \ \exists A, B > 0 \text{ such that } \mathfrak{m}\left(B_r(x_0)\right) \le A \exp(Br^2) \text{ for all } r > 0. \tag{2.15}$$

The functional $\mathsf{Ent}_{\mathfrak{m}} \colon \mathscr{P}_2(X) \to (-\infty, +\infty]$ given by

$$\mathsf{Ent}_{\mathfrak{m}}(\mu) = \begin{cases} \int_X f \log f \, \mathrm{d}\mathfrak{m} & \text{if } \mu = f\mathfrak{m} \in \mathscr{P}_2(X), \\ +\infty & \text{otherwise,} \end{cases} \tag{2.16}$$

is called the (*relative*) *entropy* of $\mu \in \mathscr{P}_2(X)$. According to our definition, $\mu \in \mathrm{Dom}(\mathsf{Ent}_{\mathfrak{m}})$ implies that $\mu \in \mathscr{P}_2(X)$ and that the effective domain $\mathrm{Dom}(\mathsf{Ent}_{\mathfrak{m}})$ is convex.

As pointed out in [14, Section 7.1], the growth condition (2.15) guarantees that in fact $\mathsf{Ent}_{\mathfrak{m}}(\mu) > -\infty$ for all $\mu \in \mathscr{P}_2(X)$, see [71, Lemma 4.1]. Hence, if $\mu = f\mathfrak{m} \in \mathscr{P}_2(X)$ for some $f \in \mathrm{L}^p(X, \mathfrak{m})$ with $p \in (1, +\infty]$, then $f |\log f| \in \mathrm{L}^1(X, \mathfrak{m})$ and $\mu \in \mathrm{Dom}(\mathsf{Ent}_{\mathfrak{m}})$.

When $\mathfrak{m} \in \mathscr{P}(X)$, the entropy functional $\mathsf{Ent}_{\mathfrak{m}}$ naturally extends to $\mathscr{P}(X)$, is lower semicontinuous with respect to the weak convergence in $\mathscr{P}(X)$ and positive by Jensen inequality. In addition, if $F \colon X \to Y$ is a Borel map, then

$$\mathsf{Ent}_{F_\sharp \mathfrak{m}}(F_\sharp \mu) \le \mathsf{Ent}_{\mathfrak{m}}(\mu) \quad \text{for all } \mu \in \mathscr{P}(X), \tag{2.17}$$

with equality if F is injective, see [12, Lemma 9.4.5].

When $\mathfrak{m}(X) = +\infty$, if we set $\mathfrak{n} = e^{-c \, \mathsf{d}(\cdot, x_0)^2} \mathfrak{m}$, where $x_0 \in X$ is as in (2.15) and $c > 0$ is chosen so that $\mathfrak{n}(X) < +\infty$ (the existence of such $c > 0$ is ensured by (2.15)), then we obtain the useful formula

$$\mathsf{Ent}_{\mathfrak{m}}(\mu) = \mathsf{Ent}_{\mathfrak{n}}(\mu) - c \int_X \mathsf{d}^2(x, x_0) \, \mathrm{d}\mu \quad \text{for all } \mu \in \mathscr{P}_2(X). \tag{2.18}$$

This shows that $\mathsf{Ent}_{\mathfrak{m}}$ is lower semicontinuous in $(\mathscr{P}_2(X), W_2)$.

2.6. **Cheeger energy.** Let $(X, \mathsf{d}, \mathfrak{m})$ be a metric-measure space with (X, d) a Polish metric space and $\mathfrak{m}$ as in (2.14). The functional $\mathsf{Ch} \colon \mathrm{L}^2(X, \mathfrak{m}) \to [0, +\infty]$ given by

$$\mathsf{Ch}(f) = \inf \left\{ \liminf_n \int_X |\mathrm{D}f_n|^2 \, \mathrm{d}\mathfrak{m} : f_n \to f \text{ in } \mathrm{L}^2(X, \mathfrak{m}), \ f_n \in \mathrm{Lip}(X) \right\}, \tag{2.19}$$

for all $f \in \mathrm{L}^2(X, \mathfrak{m})$, is called *Cheeger energy*. Here $|\mathrm{D}f|$ denotes the slope of $f \in \mathrm{Lip}(X)$ as defined in (2.2). We let

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

be the *Sobolev space* naturally associated to Ch endowed with the norm given by

$$\|f\|_{\mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})}^2 = \|f\|_{\mathrm{L}^2(X, \mathfrak{m})}^2 + 2\, \mathsf{Ch}(f). \tag{2.20}$$

The space $(\mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}), \| \cdot \|_{\mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})})$ is a separable Banach space but can fail to be a Hilbert space in general, see [14, Remark 4.6].

2.7. **Minimal weak gradient.** Let $(X, \mathsf{d}, \mathfrak{m})$ be a metric-measure space with (X, d) a Polish metric space and $\mathfrak{m}$ as in (2.14). If $f \in \mathrm{L}^2(X, \mathfrak{m})$, then

$$\mathsf{Grad}(f) = \left\{ G \in \mathrm{L}^2(X, \mathfrak{m}) : \begin{bmatrix} \exists f_n \in \mathrm{Lip}_b(X, \mathfrak{m}) \text{ such that } f_n \to f \text{ and} \\ |\mathrm{D}f_n| \rightharpoonup G \text{ in } \mathrm{L}^2(X, \mathfrak{m}) \text{ as } n \to +\infty \end{bmatrix} \right\} \tag{2.21}$$

is a convex set, possibly empty (see [14, Definition 4.2] or [131, Section 4.1]). If $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$, then it is possible to show that $\mathsf{Grad}(f) \neq \emptyset$ and thus, by the reflexivity of $\mathrm{L}^2(X, \mathfrak{m})$, $\mathsf{Grad}(f)$ has a unique element of minimal L^2-norm, the *minimal weak (upper or relaxed) gradient* of f, $|\mathrm{D}f|_w \in \mathrm{L}^2(X, \mathfrak{m})$, that is also minimal with respect to the order structure, i.e.

$$G \in \mathsf{Grad}(f) \implies |\mathrm{D}f|_w \leq G \text{ } \mathfrak{m}\text{-a.e. in } X. \tag{2.22}$$

Thanks to [14, Theorems 6.2 and 6.3] (see also [14, Remark 4.7]), if $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$, then the minimal weak gradient $|\mathrm{D}f|_w \in \mathrm{L}^2(X, \mathfrak{m})$ provides an integral representation of the Cheeger energy, so that

$$\mathsf{Ch}(f) = \frac{1}{2} \int_X |\mathrm{D}f|_w^2 \, \mathrm{d}\mathfrak{m} \quad \text{for all } f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}).$$

The minimal weak gradient is a *local* operator, i.e.

$$f, g \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}) \implies |\mathrm{D}f|_w = |\mathrm{D}g|_w \text{ } \mathfrak{m}\text{-a.e. on } \{f - g = c\}, \tag{2.23}$$

for all $c \in \mathbb{R}$, obeys a *Leibniz-rule estimate*, in the sense that if $f, g \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}) \cap \mathrm{L}^\infty(X, \mathfrak{m})$ then $fg \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ with

$$|\mathrm{D}(fg)|_w \leq |f| \, |\mathrm{D}g|_w + |\mathrm{D}f|_w \, |g|, \tag{2.24}$$

and satisfies the following *chain rule*

$$f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}) \implies \varphi(f) \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}) \text{ with } |\mathrm{D}\varphi(f)|_w \leq |\varphi'(f)| \, |\mathrm{D}f|_w \tag{2.25}$$

for any Lipschitz function $\varphi \colon I \to \mathbb{R}$ defined on an interval $I \subset \mathbb{R}$ containing the image of f (with $0 \in I$ and $\varphi(0) = 0$ if $\mathfrak{m}(X) = +\infty$), with $|\mathrm{D}\varphi(f)|_w = \varphi'(f) \, |\mathrm{D}f|_w$ if φ is non-decreasing. Also, if $f, g \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$, then $f \wedge g, f \vee g \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ with

$$|\mathrm{D}(f \wedge g)|_w = \begin{cases} |\mathrm{D}f|_w & \mathfrak{m}\text{-a.e. on } \{f \leq g\} \\ |\mathrm{D}g|_w & \mathfrak{m}\text{-a.e. on } \{f \geq g\} \end{cases} \tag{2.26}$$

and

$$|\mathrm{D}(f \vee g)|_w = \begin{cases} |\mathrm{D}f|_w & \mathfrak{m}\text{-a.e. on } \{f \geq g\} \\ |\mathrm{D}g|_w & \mathfrak{m}\text{-a.e. on } \{f \leq g\}, \end{cases} \tag{2.27}$$

see [14, Lemma 2.5 and Proposition 4.8(e)] and their proofs.

In addition, by [14, Theorem 6.3 and Lemma 4.3(c)], bounded Lipschitz functions are dense in energy in $\mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$, i.e.

$$f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}) \implies \begin{bmatrix} \exists f_n \in \mathrm{Lip}_b(X) \cap \mathrm{L}^2(X, \mathfrak{m}) \text{ such that } f_n \to f \\ \text{and } |\mathrm{D}f_n| \to |\mathrm{D}f|_w \text{ in } \mathrm{L}^2(X, \mathfrak{m}) \text{ as } n \to +\infty. \end{bmatrix} \tag{2.28}$$

In particular, we have that

$$f \in \mathrm{Lip}_b(X, \mathsf{d}) \cap \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}) \implies |\mathrm{D}f|_w \leq |\mathrm{D}f| \text{ } \mathfrak{m}\text{-a.e. in } X, \tag{2.29}$$

see also [14, Remark 5.5]. As observed in [13, Section 8.3], the approximation (2.28) can be enforced by replacing slopes with asymptotic Lipschitz constants (recall (2.2) for the definition), so that

$$f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}) \implies \begin{bmatrix} \exists f_n \in \mathrm{Lip}_b(X) \cap \mathrm{L}^2(X, \mathfrak{m}) \text{ such that } f_n \to f \\ \text{and } |\mathrm{D}^* f_n| \to |\mathrm{D}f|_w \text{ in } \mathrm{L}^2(X, \mathfrak{m}) \text{ as } n \to +\infty, \end{bmatrix} \qquad (2.30)$$

see [83, Theorem 2.8] and [8, Section 4] for a more detailed discussion.

2.8. **Heat semigroup.** Let $(X, \mathsf{d}, \mathfrak{m})$ be a metric-measure space with (X, d) a Polish metric space and $\mathfrak{m}$ as in (2.14). By [14, Theorem 4.5], Ch is convex, lower semicontinuous and 2-homogeneous. The effective domain of the Cheeger energy, which we denote by $\mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$, is dense in $\mathrm{L}^2(X, \mathfrak{m})$. Thus, by the Hilbertian theory of gradient flows, see [12, 57] for the general theory and [10, Theorem 3.1] for a plain exposition of the main results, for each given $f \in \mathrm{L}^2(X, \mathfrak{m})$ there exists a curve

$$t \mapsto f_t = \mathsf{P}_t f \in \mathrm{AC}_{\mathrm{loc}}((0, +\infty); \mathrm{L}^2(X, \mathfrak{m})), \qquad (2.31)$$

called *heat semigroup*, such that

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} f_t \in -\partial^- \mathsf{Ch}(f_t) & \text{for } \mathscr{L}^1\text{-a.e. } t \in (0, +\infty) \\ \lim_{t \to 0^+} f_t = f & \text{in } \mathrm{L}^2(X, \mathfrak{m}). \end{cases} \qquad (2.32)$$

Here and in the following, $\partial^- \mathsf{Ch}(f) \subset \mathrm{L}^2(X, \mathfrak{m})$ denotes the subdifferential of Ch at $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ and is defined by

$$\ell \in \partial^- \mathsf{Ch}(f) \iff \mathsf{Ch}(g) \geq \mathsf{Ch}(f) + \int_X \ell \, (g - f) \, \mathrm{d}\mathfrak{m} \text{ for all } g \in \mathrm{L}^2(X, \mathfrak{m}).$$

The heat flow (2.31) is uniquely determined by (2.32), is 1-homogeneous, i.e.

$$f \in \mathrm{L}^2(X, \mathfrak{m}), \ \lambda \in \mathbb{R} \implies \mathsf{P}_t(\lambda f) = \lambda \mathsf{P}_t f \text{ for all } t \geq 0,$$

and defines a strongly continuous semigroup of contractions in $\mathrm{L}^2(X, \mathfrak{m})$, meaning that

$$\|f_t\|_{\mathrm{L}^2(X, \mathfrak{m})} \leq \|f\|_{\mathrm{L}^2(X, \mathfrak{m})} \text{ for all } t > 0 \text{ and } f \in \mathrm{L}^2(X, \mathfrak{m}). \qquad (2.33)$$

By [14, Theorem 4.16(a)], the heat semigroup preserves one-side essential bounds (*maximum principle*). Precisely, for $C \in \mathbb{R}$ it holds

$$f \leq C \ (\text{resp. } f \geq C) \implies f_t \leq C \ (\text{resp. } f_t \geq C) \text{ for all } t \geq 0 \qquad (2.34)$$

and, moreover,

$$f \leq g + C \implies f_t \leq g_t + C \text{ for all } t \geq 0, \qquad (2.35)$$

whenever $f, g \in \mathrm{L}^2(X, \mathfrak{m})$. By [14, Theorem 4.16(b)], the heat semigroup satisfies the contraction property

$$\|f_t - g_t\|_{\mathrm{L}^p(X, \mathfrak{m})} \leq \|f - g\|_{\mathrm{L}^p(X, \mathfrak{m})} \quad \text{for all } f, g \in \mathrm{L}^2(X, \mathfrak{m}) \cap \mathrm{L}^p(X, \mathfrak{m}), \qquad (2.36)$$

whenever $p \in [1, +\infty]$. Since $\mathrm{L}^p(X, \mathfrak{m}) \cap \mathrm{L}^2(X, \mathfrak{m})$ is L^p-dense in $\mathrm{L}^p(X, \mathfrak{m})$ for all $p \in [1, +\infty)$, we can uniquely extend the heat semigroup to a strongly continuous semigroup of contractions in $\mathrm{L}^p(X, \mathfrak{m})$, $p \in [1, +\infty)$, for which we retain the same notation. The heat semigroup can thus be extended to a weakly*-continuous semigroup of contractions in $\mathrm{L}^\infty(X, \mathfrak{m})$ by duality, i.e.

$$\int_X \varphi \, \mathsf{P}_t f \, \mathrm{d}\mathfrak{m} = \int_X f \, \mathsf{P}_t \varphi \, \mathrm{d}\mathfrak{m} \quad \text{for every } f \in \mathrm{L}^1(X, \mathfrak{m}) \text{ and } \varphi \in \mathrm{L}^\infty(X, \mathfrak{m}). \qquad (2.37)$$

By [14, Theorem 4.20], thanks to (2.36), if $\mathfrak{m}$ satisfies the growth condition (2.15), then the heat semigroup satisfies the *mass preservation property*

$$\int_X f_t \, d\mathfrak{m} = \int_X f \, d\mathfrak{m} \quad \text{for all } t \geq 0 \text{ and } f \in L^1(X, \mathfrak{m}). \tag{2.38}$$

The heat semigroup is regularizing as stated in Lemma 2.1 below. This result is well known to experts, but we quickly prove it here for the reader's convenience.

Lemma 2.1 (Heat flow regularization)**.** *Let $f \in L^2(X, \mathfrak{m})$. Then*

$$t \mapsto \mathsf{Ch}(f_t) \in \mathrm{AC}_{\mathrm{loc}}((0, +\infty); [0, +\infty)) \tag{2.39}$$

with

$$\mathsf{Ch}(f_t) \leq \inf\left\{ \mathsf{Ch}(g) + \frac{1}{2t} \int_X |f - g|^2 \, d\mathfrak{m} : g \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}) \right\} \tag{2.40}$$

for all $t > 0$, and

$$t \mapsto |\mathrm{D}f_t|_w \in \mathrm{C}((0, +\infty); L^2(X, \mathfrak{m})). \tag{2.41}$$

Moreover, if $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$, then the continuity of the maps in (2.39) and (2.41) extends to $t = 0$.

Proof. We divide the proof in two steps.

Step 1: regularity of Cheeger energy along the heat flow. The $\mathrm{AC}_{\mathrm{loc}}$-regularity of the Chegeer energy along the heat flow in (2.39) and the inf-formula in (2.40) follow from the theory of Hilbertian gradient flows, see [10, Theorem 3.1]. If moreover $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$, then $\mathsf{Ch}(f_t) \leq \mathsf{Ch}(f)$ for all $t > 0$ by (2.40), so that $\mathsf{Ch}(f_t) \to \mathsf{Ch}(f)$ as $t \to 0^+$ by the lower semicontinuity of the Cheeger energy.

Step 2: regularity of the minimal weak gradient along the heat flow. Fix $t > 0$ and let $t_n > 0$, $n \in \mathbb{N}$, be such that $t_n \to t$ as $n \to +\infty$. By (2.39), the sequence $(|\mathrm{D}f_{t_n}|_w)_{n \in \mathbb{N}}$ is bounded in $L^2(X, \mathfrak{m})$. We can thus find a subsequence $(|\mathrm{D}f_{t_{n_k}}|_w)_{k \in \mathbb{N}}$ and a function $G \in L^2(X, \mathfrak{m})$ such that $|\mathrm{D}f_{t_k}|_w \rightharpoonup G$ in $L^2(X, \mathfrak{m})$ as $k \to +\infty$. By the weak lower semicontinuity of the L^2-norm and again by (2.39), we must have $\|G\|_{L^2(X,\mathfrak{m})} \leq \||\mathrm{D}f_t|_w\|_{L^2(X,\mathfrak{m})}$. By definition of minimal weak gradient and [14, Lemma 4.3(b)], we must also have that $|\mathrm{D}f_t|_w \leq G$ $\mathfrak{m}$-a.e. in X. Hence $G = |\mathrm{D}f_t|_w$ $\mathfrak{m}$-a.e. in X and thus $|\mathrm{D}f_{t_{n_k}}|_w \to |\mathrm{D}f_t|_w$ in $L^2(X, \mathfrak{m})$ as $k \to +\infty$ by the uniform convexity of $L^2(X, \mathfrak{m})$ (see [58, Proposition 3.32] for example). Hence (2.41) readily follows. If moreover $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$, then we can take $t = 0$ in the above argument and use the continuity in $t = 0$ of the map in (2.39) to extend the L^2-continuity of the minimal weak gradient along the heat flow to $t = 0$. $\square$

2.9. **Metric-measure Laplacian.** Let $(X, \mathsf{d}, \mathfrak{m})$ be a metric-measure space with (X, d) a Polish metric space and $\mathfrak{m}$ as in (2.14). If $f \in L^2(X, \mathfrak{m})$ and $\partial^- \mathsf{Ch}(f) \neq \emptyset$, then the element of minimal L^2-norm in $-\partial^- \mathsf{Ch}(f)$ is called the (*metric-measure*) *Laplacian* of the function f and is denoted by $\Delta_{\mathsf{d},\mathfrak{m}} f$, see [14, Definition 4.13]. The effective domain $\mathrm{Dom}(\Delta_{\mathsf{d},\mathfrak{m}})$ of the Laplacian is a L^2-dense subset of $\mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ (and thus, in particular, a L^2-dense subset of $L^2(X, \mathfrak{m})$), see [57, Proposition 2.11]. Note that the operator $\Delta_{\mathsf{d},\mathfrak{m}}$ is not linear in general, but is 1-homogeneous, in the sense that

$$f \in \mathrm{Dom}(\Delta_{\mathsf{d},\mathfrak{m}}), \ \lambda \in \mathbb{R} \implies \lambda f \in \mathrm{Dom}(\Delta_{\mathsf{d},\mathfrak{m}}) \text{ with } \Delta_{\mathsf{d},\mathfrak{m}}(\lambda f) = \lambda \Delta_{\mathsf{d},\mathfrak{m}} f.$$

By the regularizing properties of gradient flows in Hilbert spaces (see [10, Theorem 3.1]), for every $t > 0$ the right time-derivative $\frac{d^+}{dt} f_t$ exists and it is actually the element with minimal L^2-norm in $-\partial^- \mathsf{Ch}(f)$, so that $f_t \in \mathrm{Dom}(\Delta_{\mathsf{d},\mathfrak{m}})$ for all $t > 0$ and we can rewrite (2.32)

as

$$\begin{cases} \dfrac{d^+}{dt} f_t = \Delta_{\mathsf{d},\mathsf{m}} f_t & \text{for every } t \in (0, +\infty) \\[2mm] \lim_{t \to 0^+} f_t = f & \text{in } L^2(X, \mathfrak{m}) \end{cases}$$

whenever $f \in L^2(X, \mathfrak{m})$ is given. Moreover, by the integration-by-part formula provided by [14, Proposition 4.15], it holds that

$$\frac{d^+}{dt} \mathsf{Ch}(f_t) = -\|\Delta_{\mathsf{d},\mathsf{m}} f_t\|^2_{L^2(X,\mathfrak{m})} \text{ for all } t > 0,$$

whenever $f \in L^2(x, \mathfrak{m})$.

2.10. **Quadratic Cheeger energy.** Let $(X, \mathsf{d}, \mathfrak{m})$ be a metric-measure space with (X, d) a Polish metric space and $\mathfrak{m}$ as in (2.14) and (2.15). As in [15, Section 4.3], we say that the Cheeger energy is *quadratic* if it satisfies the *parallelogram identity*

$$\mathsf{Ch}(f + g) + \mathsf{Ch}(f - g) = 2\mathsf{Ch}(f) + 2\mathsf{Ch}(g) \quad \text{for all } f, g \in W^{1,2}(X, \mathsf{d}, \mathfrak{m}). \tag{2.42}$$

In this case, Ch is a quadratic form on $W^{1,2}(X, \mathsf{d}, \mathfrak{m})$, the functional $\mathcal{E} \colon W^{1,2}(X, \mathsf{d}, \mathfrak{m}) \to [0, +\infty)$ defined by the formula

$$\mathcal{E}(f, g) = \mathsf{Ch}(f + g) - \mathsf{Ch}(f) - \mathsf{Ch}(g) \quad \text{for all } f, g \in W^{1,2}(X, \mathsf{d}, \mathfrak{m}) \tag{2.43}$$

is a symmetric bilinear form on $W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ and $(W^{1,2}(X, \mathsf{d}, \mathfrak{m}), \|\cdot\|_{W^{1,2}(X,\mathsf{d},\mathfrak{m})})$ is a Hilbert space, see [83, Proposition 4.22]. In particular, thanks to (2.28), the set $\mathrm{Lip}_b(X) \cap W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ is $W^{1,2}$-dense in $W^{1,2}(X, \mathsf{d}, \mathfrak{m})$, see [83, Corollary 2.9].

For simplicity, we set $\mathcal{E}(f) = \mathcal{E}(f, f)$ for all $f \in W^{1,2}(X, \mathsf{d}, \mathfrak{m})$. The chain rule (2.25) for the minimal weak gradient proves that $\mathcal{E}$ is *Markovian*, i.e.

$$\mathcal{E}(\varphi \circ f) \leq \mathcal{E}(f) \text{ for all } f \in W^{1,2}(X, \mathfrak{m}),$$

whenever $\varphi \colon \mathbb{R} \to \mathbb{R}$ is 1-Lipschitz with $\varphi(0) = 0$. Since Ch is lower semicontinuous, the form $\mathcal{E}$ is also closed. Thus, thanks to the density of $W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ in $L^2(X, \mathfrak{m})$, we can extend the form $\mathcal{E}$ given in (2.43) to a symmetric bilinear form on $L^2(X, \mathfrak{m})$, for which we retain the same notation. By the locality property of the minimal weak gradient (2.23), the form $\mathcal{E}$ is *strongly local*, meaning that

$$f, g \in W^{1,2}(X, \mathsf{d}, \mathfrak{m}), \ (f + c)g = 0 \ \mathfrak{m}\text{-a.e. in } X \implies \mathcal{E}(f, g) = 0.$$

By [83, Proposition 4.24], the Laplacian $\Delta_{\mathsf{d},\mathsf{m}}$ coincides with the generator of $\mathcal{E}$ and hence satisfies the *integration-by-part formula*

$$\mathcal{E}(f, g) = -\int_X g \, \Delta_{\mathsf{d},\mathsf{m}} f \, d\mathfrak{m} \quad \text{for all } f \in \mathrm{Dom}(\Delta_{\mathsf{d},\mathsf{m}}), \ g \in W^{1,2}(X, \mathsf{d}, \mathfrak{m}). \tag{2.44}$$

Thus, since $\mathcal{E}$ is symmetric, the Laplacian $\Delta_{\mathsf{d},\mathsf{m}}$ is a self-adjoint operator in $L^2(X, \mathfrak{m})$. In addition, by [83, Proposition 4.23], the Laplacian is a linear operator.

Therefore, if the Cheeger energy is quadratic, the heat semigroup $(\mathsf{P}_t)_{t \geq 0}$ is a linear *analytic Markov semigroup* in $L^2(X, \mathfrak{m})$, is a self-adjoint operator in $L^2(X, \mathfrak{m})$ and the map (2.31) is the unique C^1 map with values in $\mathrm{Dom}(\Delta_{\mathsf{d},\mathsf{m}})$ satisfying

$$\begin{cases} \dfrac{d}{dt} f_t = \Delta_{\mathsf{d},\mathsf{m}} f_t & \text{for } t \in (0, +\infty), \\[2mm] \lim_{t \to 0^+} f_t = f & \text{in } L^2(X, \mathfrak{m}). \end{cases}$$

Because of this, $\Delta_{\mathsf{d,m}}$ can be equivalently characterized in terms of the strong convergence

$$\frac{\mathsf{P}_t f - f}{t} \to \Delta_{\mathsf{d,m}} f \quad \text{in } \mathrm{L}^2(X, \mathfrak{m}) \text{ as } t \to 0^+. \tag{2.45}$$

By [83, Proposition 4.21] (see also [15, Theorem 4.18]), if the Cheeger energy is quadratic, then the parallelogram identity (2.42) can be localized at the level of minimal weak gradient, in the sense that

$$|\mathrm{D}(f + g)|_w^2 + |\mathrm{D}(f - g)|_w^2 = 2|\mathrm{D}f|_w^2 + 2|\mathrm{D}g|_w^2 \quad \mathfrak{m}\text{-a.e. in } X$$

for all $f, g \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$. Thus, the naturally associated Γ operator, given by

$$\Gamma(f, g) = |\mathrm{D}(f + g)|_w^2 - |\mathrm{D}f|_w^2 - |\mathrm{D}g|_w^2 \quad \text{for all } f, g \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}),$$

defines a strongly-continuous, symmetric and bilinear map from $\mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ to $\mathrm{L}^1(X, \mathfrak{m})$ which represents the form $\mathcal{E}$, i.e.

$$\mathcal{E}(f, g) = \int_X \Gamma(f, g) \, \mathrm{d}\mathfrak{m} \quad \text{for all } f, g \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}).$$

The Γ operator satisfies the pointwise estimate

$$\Gamma(f, g) \le |\mathrm{D}f|_w \, |\mathrm{D}g|_w \quad \text{for all } f, g \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}), \tag{2.46}$$

the chain rule

$$\Gamma(\varphi(f), g) = \varphi'(f) \, \Gamma(f, g) \quad \text{for all } f, g \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}), \tag{2.47}$$

whenever $\varphi \in \mathrm{Lip}(\mathbb{R})$ with $\varphi(0) = 0$, and the Leibiniz rule

$$\Gamma(fg, h) = g \, \Gamma(f, h) + f \, \Gamma(g, h) \quad \text{for all } f, g, h \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}), \; f, g \in \mathrm{L}^\infty(x, \mathfrak{m}),$$

see the discussion in [83, Chapters 3 and 4]. The operator

$$\Gamma(f) = |\mathrm{D}f|_w^2, \quad \text{defined for } f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}),$$

is therefore the *carré du champ* associated to $\mathcal{E}$ and obeys the rules of Γ-*Calculus*. With these notations, the Laplacian satisfies the following *chain rule*, see [83, Proposition 4.28]: if $f \in \mathrm{Dom}(\Delta_{\mathsf{d,m}}) \cap \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ and $\varphi \in \mathrm{C}^2(\mathbb{R})$ with $\varphi(0) = 0$, then $\varphi(f) \in \mathrm{Dom}(\Delta_{\mathsf{d,m}})$ with

$$\Delta_{\mathsf{d,m}}(\varphi \circ f) = \varphi'(f) \, \Delta_{\mathsf{d,m}} f + \varphi''(f) \, \Gamma(f). \tag{2.48}$$

For an account on Γ-Calculus in the present and related frameworks, we refer the reader to [16, 17, 22, 82, 83] and to the monograph [24].

2.11. **Main assumptions and length property.** We conclude this section summarizing the main assumptions we are going to use throughout this paper. We assume that $(X, \mathsf{d}, \mathfrak{m})$ is a metric-measure space satisfying the following properties:

(P.1) (X, d) is a complete and separable metric space;

(P.2) $\mathfrak{m}$ is a non-negative Borel-regular measure on X, finite on bounded sets and such that $\mathrm{supp}\, \mathfrak{m} = X$;

(P.3) there exist $x_0 \in X$ and $A, B > 0$ such that $\mathfrak{m}(B_r(x_0)) \le A \exp(Br^2)$ for all $r > 0$;

(P.4) the Cheeger energy Ch is quadratic, i.e. $\mathsf{Ch}(f + g) + \mathsf{Ch}(f - g) = 2\mathsf{Ch}(f) + 2\mathsf{Ch}(g)$ for all $f, g \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$;

(P.5) if $L \in [0, +\infty)$ and $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ satisfies $|\mathrm{D}f|_w \le L$ $\mathfrak{m}$-a.e. in X, then $f = \tilde{f}$ $\mathfrak{m}$-a.e. in X for some $\tilde{f} \in \mathrm{Lip}(X)$ with $\mathrm{Lip}(\tilde{f}) \le L$.

We say that a metric-measure space $(X, \mathsf{d}, \mathfrak{m})$ is *admissible* if it satisfies the properties (P.1)-(P.5) listed above.

Let us briefly comment on these assumptions. As we have already seen, assumptions (P.1)–(P.4) ensure that $(X, \mathsf{d}, \mathfrak{m})$ satisfies all the properties we have recalled in this section. The additional assumption (P.5), instead, allows to identify the metric-measure structure of $(X, \mathsf{d}, \mathfrak{m})$ with the energetic-measure structure of $(X, \mathscr{B}, \mathcal{E}, \mathfrak{m})$, where $\mathscr{B}$ is the Borel σ-algebra generate by the topology of open d-balls, thus making the metric-measure and the energetic-measure approaches equivalent. More precisely, once the Dirichlet–Cheeger energy $\mathcal{E}$ is available, it is relevant to check if the function

$$\mathsf{d}_{\mathcal{E}}(x, y) = \sup\Big\{ |f(x) - f(y)| : f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}) \cap \mathrm{C}(X) \text{ with } \Gamma(f) \leq 1 \text{ } \mathfrak{m}\text{-a.e.} \Big\},$$

defined for all $x, y \in X$, actually coincides with the starting distance function d. The function $\mathsf{d}_{\mathcal{E}}$ is known as the Biroli–Mosco distance, see [51] and also [59, 132, 135, 136]. From (2.29), we immediately get that $\mathsf{d} \leq \mathsf{d}_{\mathcal{E}}$. The opposite inequality follows if any function $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}) \cap \mathrm{C}(X)$ with $\Gamma(f) \leq 1$ $\mathfrak{m}$-a.e. in X is 1-Lipschitz, which is precisely (P.5).

In $\mathsf{RCD}(K, \infty)$ spaces, assumption (P.5) is an important consequence of the $\mathsf{BE}(K, \infty)$ condition. Namely, since $\Gamma(\mathsf{P}_t f) \leq e^{-Kt} \mathsf{P}_t \Gamma(f)$ for all $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$, the 1-Lipschitz regularity of $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}) \cap \mathrm{C}(X)$ can be obtained passing to the limit as $t \to 0^+$ from the e^{-Kt}-Lipschitz regularity of the (more regular) function $\mathsf{P}_t f$, see [16, Remark 3.8] and the proof of (v) $\Rightarrow$ (ii) in [16, Theorem 3.17].

In the more general situation in which the function $t \mapsto e^{-Kt}$ is replaced by a function $t \mapsto \mathsf{c}(t)$ such that $\lim_{t \to 0^+} \mathsf{c}(t) > 1$ (as it happens in sub-Riemannian spaces), then property (P.5) cannot be inferred from the contractivity property of the heat semigroup.

One useful consequence of the identification $\mathsf{d}_{\mathcal{E}} = \mathsf{d}$ provided by (P.5) that will be employed several times in the sequel is the length property of the metric space (X, d), see [16, Theorems 3.10] (and also [132]).

Proposition 2.2 (Length property). *If $(X, \mathsf{d}, \mathfrak{m})$ satisfies properties (P.1), (P.2), (P.4) and (P.5) (and thus, in particular, if it is admissible), then (X, d) is a length space.*

3. Weak Bakry–Émery curvature condition and Kuwada duality

In this section, we introduce and study a generalization of the *Bakry–Émery curvature condition* for Sobolev functions and its equivalence with the Wasserstein contraction property of the dual heat semigroup. The presentation of the results will be close in spirit to that of [16, 97, 98]. For the reader's ease, we adopt the notation of [16].

If not otherwise stated, from now on we assume that $(X, \mathsf{d}, \mathfrak{m})$ is an admissible metric-measure space as in Section 2.11.

3.1. **Semigroup mollification.** We begin this section by recalling an useful technical tool that we will use in the following. Let $\kappa \in \mathrm{C}_c^{\infty}((0, +\infty))$ be such that

$$\kappa \geq 0 \quad \text{and} \quad \int_0^{+\infty} \kappa(r) \, \mathrm{d}r = 1. \tag{3.1}$$

Let $p \in [1, +\infty]$. For every $f \in \mathrm{L}^p(X, \mathfrak{m})$, let us set

$$\mathfrak{h}^{\varepsilon} f = \frac{1}{\varepsilon} \int_0^{+\infty} \mathsf{P}_r f \, \kappa\left(\frac{r}{\varepsilon}\right) \mathrm{d}r, \quad \text{for all } \varepsilon > 0, \tag{3.2}$$

be the *semigroup mollification operator*, where the integral is intended in the Bochner sense if $p < +\infty$ and by duality with any function in $\mathrm{L}^1(X, \mathfrak{m})$ if $p = +\infty$.

Obviously, by the semigroup property, we have $\mathsf{P}_t(\mathfrak{h}^\varepsilon f) = \mathfrak{h}^\varepsilon(\mathsf{P}_t f)$ for all $t, \varepsilon > 0$. Since, by a simple change of variable,

$$\mathfrak{h}^\varepsilon f = \int_0^{+\infty} \mathsf{P}_{\varepsilon r} f \, \kappa(r) \, \mathrm{d}r, \quad \text{for all } \varepsilon > 0,$$

by (3.1) we immediately deduce that $\mathfrak{h}^\varepsilon f$ converges to f as $\varepsilon \to 0^+$ strongly in $\mathrm{L}^p(X, \mathfrak{m})$ if $p < +\infty$ and weakly* in $\mathrm{L}^\infty(X, \mathfrak{m})$. The semigroup mollification operator satisfies the following natural $\mathrm{W}^{1,2}$-approximation property.

Lemma 3.1 ($\mathrm{W}^{1,2}$-approximation via $(\mathfrak{h}^\varepsilon)_{\varepsilon>0}$). *If $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$, then $\mathfrak{h}^\varepsilon f \to f$ in $\mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ as $\varepsilon \to 0^+$.*

Proof. From the definition in (2.21), it follows that

$$\Gamma(\mathfrak{h}^\varepsilon f - f) \le \int_0^{+\infty} \Gamma(\mathsf{P}_{\varepsilon r} f - f) \, \kappa(r) \, \mathrm{d}r \quad \mathfrak{m}\text{-a.e. in } X,$$

so that the conclusion follows by Lemma 2.1 and the Dominated Convergence Theorem. $\square$

For the reader's convenience, we briefly prove the following regularity result for the Laplacian of the semigroup mollification operator.

Lemma 3.2 (Laplacian of $(\mathfrak{h}^\varepsilon)_{\varepsilon>0}$). *Let $p \in [1, +\infty]$. If $f \in \mathrm{L}^2(X, \mathfrak{m}) \cap \mathrm{L}^p(X, \mathfrak{m})$, then*

$$-\Delta_{\mathsf{d},\mathfrak{m}}(\mathfrak{h}^\varepsilon f) = \frac{1}{\varepsilon^2} \int_0^{+\infty} \mathsf{P}_r f \, \kappa'\left(\frac{r}{\varepsilon}\right) \mathrm{d}r \in \mathrm{L}^2(X, \mathfrak{m}) \cap \mathrm{L}^p(X, \mathfrak{m})$$

for all $\varepsilon > 0$.

Proof. We argue as in the proof of [125, Theorem 2.7]. Without loss of generality, we can assume $\varepsilon = 1$. If $t > 0$, then

$$\frac{\mathsf{P}_t(\mathfrak{h}^1 f) - \mathfrak{h}^1 f}{t} = \int_0^{+\infty} \frac{\mathsf{P}_{r+t} f - \mathsf{P}_r f}{t} \, \kappa(r) \, \mathrm{d}r = \int_0^{+\infty} \mathsf{P}_r f \, \frac{\kappa(r-t) - \kappa(r)}{t} \, \mathrm{d}r.$$

As $t \to 0^+$, the integrand in the last term converges to $-\mathsf{P}_r f \kappa'(r)$ uniformly for $r \in [0, +\infty)$. Thus, in virtue of (2.45), we get that $\mathfrak{h}^1 f \in \mathrm{Dom}(\Delta_{\mathsf{d},\mathfrak{m}})$ and

$$\Delta_{\mathsf{d},\mathfrak{m}}(\mathfrak{h}^1 f) = \lim_{t \to 0^+} \frac{\mathsf{P}_t(\mathfrak{h}^1 f) - \mathfrak{h}^1 f}{t} = -\int_0^{+\infty} \mathsf{P}_r f \, \kappa'(r) \, \mathrm{d}r$$

and the conclusion follows. $\square$

3.2. **A differentiation formula.** We now prove Lemma 3.3 below. This result was proved for the first time in [16, Lemma 2.1] to provide a very general formulation, in the weak sense and with minimal requirements on the regularity of the functions involved, of the simple differentiation formula

$$\frac{\partial}{\partial s} \mathsf{P}_s(\mathsf{P}_{t-s} f)^2 = \mathsf{P}_s \left\{ \Delta(\mathsf{P}_{t-s} f)^2 - 2(\mathsf{P}_{t-s} f)(\Delta \mathsf{P}_{t-s} f) \right\} = 2\mathsf{P}_s |\nabla \mathsf{P}_{t-s} f|_{\mathrm{g}}^2 \qquad (3.3)$$

valid for all $f \in \mathrm{C}^\infty(\mathbb{M})$ on a Riemannian manifold $(\mathbb{M}, \mathrm{g})$, see [23, 24, 28, 143] for an account.

Note that the differentiation formula (3.3), as well as Lemma 3.3 below, does not require any information about the curvature of the ambient space. For the reader's convenience

and in order to keep this work the most self-contained as possible, we provide a proof of this result in our setting.

Lemma 3.3 (Differentiation formula). *Let $f \in \mathrm{L}^2(X, \mathfrak{m})$ and $\varphi \in \mathrm{L}^2(X, \mathsf{d}, \mathfrak{m}) \cap \mathrm{L}^\infty(X, \mathfrak{m})$. If $t > 0$, then*

$$s \mapsto \mathsf{A}_t[f; \varphi](s) = \frac{1}{2} \int_X (\mathsf{P}_{t-s}f)^2 \, \mathsf{P}_s\varphi \, \mathrm{d}\mathfrak{m} \in \mathrm{C}([0, t]) \cap \mathrm{C}^1([0, t)),$$

$$s \mapsto \mathsf{B}_t[f; \varphi](s) = \int_X \Gamma(\mathsf{P}_{t-s}f) \, \mathsf{P}_s\varphi \, \mathrm{d}\mathfrak{m} \in \mathrm{C}([0, t)) \tag{3.4}$$

and

$$\frac{\partial}{\partial s}\mathsf{A}_t[f; \varphi](s) = \mathsf{B}_t[f; \varphi](s) \quad \text{for all } s \in [0, t). \tag{3.5}$$

The regularity of the functions A *and* B *in* (3.4) *and the differentiation formula in* (3.5) *extend to* $s = t$ *if* $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$.

Proof. We divide the proof in four steps.

Step 1: continuity of A. The function $s \mapsto \mathsf{P}_{t-s}f$, $s \in [0, t]$, is strongly continuous in $\mathrm{L}^2(X, \mathfrak{m})$ by the definition of the heat flow. Thanks to the L^1-contraction property of the heat semigroup, by a simple approximation argument we easily get that $s \mapsto \mathsf{P}_s\varphi$, $s \in [0, t]$, is weakly* continuous in $\mathrm{L}^\infty(X, \mathfrak{m})$. This prove that $s \mapsto \mathsf{A}_t[f; \varphi](s) \in \mathrm{C}([0, t])$.

Step 2: continuity of B. Since the function $s \mapsto \Gamma(\mathsf{P}_{t-s}f)$, $s \in [0, t)$, is strongly continuous in $\mathrm{L}^1(X, \mathfrak{m})$ by (2.41), and since the function $s \mapsto \mathsf{P}_s\varphi$, $s \in [0, t]$, is weakly* continuous in $\mathrm{L}^\infty(X, \mathfrak{m})$ by Step 1, we easily deduce that $s \mapsto \mathsf{B}_t[f; \varphi](s) \in \mathrm{C}([0, t))$.

Step 3: proof of the differentiation formula (3.5). Let us first assume that

$$f \in \mathrm{L}^2(X, \mathfrak{m}) \cap \mathrm{L}^\infty(X, \mathfrak{m}), \ \varphi \in \mathrm{L}^\infty(X, \mathfrak{m}) \cap \mathrm{Dom}(\Delta_{\mathsf{d},\mathfrak{m}}) \text{ with } \Delta_{\mathsf{d},\mathfrak{m}}\varphi \in \mathrm{L}^\infty(X, \mathfrak{m}). \tag{3.6}$$

Then we have

$$\lim_{h \to 0} \frac{\mathsf{P}_{t-(s+h)}f - \mathsf{P}_{t-s}f}{h} = -\Delta_{\mathsf{d},\mathfrak{m}}\mathsf{P}_{t-s}f$$

strongly in $\mathrm{L}^2(X, \mathfrak{m})$ for all $s \in [0, t)$ by (2.45) and

$$\lim_{h \to 0} \frac{\mathsf{P}_{s+h}\varphi - \mathsf{P}_s\varphi}{h} = \Delta_{\mathsf{d},\mathfrak{m}}\mathsf{P}_s\varphi$$

weakly* in $\mathrm{L}^\infty(X, \mathfrak{m})$ for all $s \in [0, t]$ by (2.37) and again by (2.45). Hence we get that

$$\frac{\partial}{\partial s}\mathsf{A}_t[f; \varphi](s) = \int_X \left(-\mathsf{P}_{t-s}f \, \Delta_{\mathsf{d},\mathfrak{m}}\mathsf{P}_{t-s}f \, \mathsf{P}_s\varphi + \frac{1}{2}(\mathsf{P}_{t-s}f)^2 \, \Delta_{\mathsf{d},\mathfrak{m}}\mathsf{P}_s\varphi \right) \mathrm{d}\mathfrak{m}$$

for all $s \in [0, t)$. Since $\mathsf{P}_{t-s}f \in \mathrm{L}^\infty(X, \mathfrak{m})$ by (3.6) according to (2.34), we have $(\mathsf{P}_{t-s}f)^2 \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ and thus, thanks to the integration-by-part formula (2.44) and the chain rule (2.48) for the Laplacian, we can compute

$$\int_X (\mathsf{P}_{t-s}f)^2 \, \Delta_{\mathsf{d},\mathfrak{m}}\mathsf{P}_s\varphi \, \mathrm{d}\mathfrak{m} = \int_X \Delta_{\mathsf{d},\mathfrak{m}}(\mathsf{P}_{t-s}f)^2 \, \mathsf{P}_s\varphi \, \mathrm{d}\mathfrak{m}$$

$$= -2 \int_X \mathsf{P}_{t-s}f \, \Delta_{\mathsf{d},\mathfrak{m}}\mathsf{P}_{t-s}f \, \mathsf{P}_s\varphi \, \mathrm{d}\mathfrak{m} + 2 \int_X \Gamma(\mathsf{P}_{t-s}f) \, \mathsf{P}_s\varphi \, \mathrm{d}\mathfrak{m}$$

for all $s \in [0, t)$ and (3.5) follows.

Now let $f \in \mathrm{L}^2(X, \mathfrak{m})$ and keep φ as in (3.6). Let $f_n = -n \vee f \wedge n \in \mathrm{L}^2(X, \mathfrak{m}) \cap \mathrm{L}^\infty(X, \mathfrak{m})$ for all $n \in \mathbb{N}$ and note that $f_n \to f$ in $\mathrm{L}^2(X, \mathfrak{m})$ as $n \to +\infty$. By (3.5) applied to f_n and φ, we know that

$$\mathsf{A}_t[f_n; \varphi](s_1) - \mathsf{A}_t[f_n; \varphi](s_0) = \int_{s_0}^{s_1} \mathsf{B}_t[f_n; \varphi](s) \, \mathrm{d}s \tag{3.7}$$

for all $0 \le s_0 < s_1 < t$. Since

$$||\mathsf{DP}_{t-s}f_n|_w - |\mathsf{DP}_{t-s}f|_w|^2 \le \Gamma(\mathsf{P}_{t-s}(f_n - f))$$

for all $n \in \mathbb{N}$ and $s \in [0, t)$ by (2.46), and since

$$\int_X \Gamma(\mathsf{P}_{t-s}(f_n - f)) \, \mathrm{d}\mathfrak{m} = 2\mathsf{Ch}(\mathsf{P}_{t-s}(f_n - f)) \le \frac{1}{t-s} \int_X |f_n - f|^2 \, \mathrm{d}\mathfrak{m}$$

for all $n \in \mathbb{N}$ and $s \in [0, t)$ by (2.40), we have that $\Gamma(\mathsf{P}_{t-s}f_n) \to \Gamma(\mathsf{P}_{t-s}f)$ in $\mathrm{L}^1(X, \mathfrak{m})$ as $n \to +\infty$ for all $s \in [0, t)$. Since also $\mathsf{P}_{t-s}f_n \to \mathsf{P}_{t-s}f$ in $\mathrm{L}^2(X, \mathfrak{m})$ as $n \to +\infty$ for all $s \in [0, t)$, we can pass to the limit as $n \to +\infty$ in (3.7) and prove (3.5) for all $f \in \mathrm{L}^2(X, \mathfrak{m})$ and φ as in (3.6).

Finally, let $f \in \mathrm{L}^2(X, \mathfrak{m})$ and $\varphi \in \mathrm{L}^2(X, \mathfrak{m}) \cap \mathrm{L}^\infty(X, \mathfrak{m})$. For all $\varepsilon > 0$, we set $\varphi_\varepsilon = \mathfrak{h}^\varepsilon \varphi$. By Lemma 3.2, we know that φ_ε is as in (3.6) for all $\varepsilon > 0$ and moreover $\varphi_\varepsilon \to \varphi$ weakly* in $\mathrm{L}^\infty(X, \mathfrak{m})$ as $\varepsilon \to 0^+$. By applying (3.5) to f and φ_ε in its integrated form and then passing to the limit as $\varepsilon \to 0^+$, we prove (3.5) for all $f \in \mathrm{L}^2(X, \mathfrak{m})$ and $\varphi \in \mathrm{L}^2(X, \mathfrak{m}) \cap \mathrm{L}^\infty(X, \mathfrak{m})$.

Step 4: the limit case $s = t$. Let $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$. By the Mean Value Theorem, we just need to prove that the continuity of the function B extends to $s = t$. This immediately follows since $\Gamma(\mathsf{P}_{t-s}f) \to \Gamma(f)$ in $\mathrm{L}^1(X, \mathfrak{m})$ as $s \to t^-$ thanks to Lemma 2.1. $\qquad \square$

3.3. $\mathsf{BE}_w(\mathsf{c}, \infty)$ condition.
We now come to the central definition of our paper. Here and in the following, we let

$$\mathsf{c} \colon [0, +\infty) \to (0, +\infty) \text{ be such that } \mathsf{c}, \mathsf{c}^{-1} \in \mathrm{L}^\infty([0, T]) \text{ for all } T > 0. \tag{3.8}$$

Definition 3.4 ($\mathsf{BE}_w(\mathsf{c}, \infty)$ condition). We say that $(X, \mathsf{d}, \mathfrak{m})$ satisfies the *weak Bakry–Émery curvature condition* with respect to the function $\mathsf{c} \colon [0, +\infty) \to (0, +\infty)$ in (3.8), $\mathsf{BE}_w(\mathsf{c}, \infty)$ for short, if for all $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ and $t \ge 0$ the function $\mathsf{P}_t f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ satisfies

$$\Gamma(\mathsf{P}_t f) \le \mathsf{c}^2(t) \, \mathsf{P}_t \Gamma(f) \quad \mathfrak{m}\text{-a.e. in } X. \tag{3.9}$$

Although not strictly necessary, we always assume that $\mathsf{c}(0) = 1$ for simplicity.

Clearly, if $\mathsf{c}(t) = e^{-Kt}$ for $t \ge 0$, then (3.9) states that $\Gamma(\mathsf{P}_t f) \le e^{-2Kt} \mathsf{P}_t \Gamma(f)$ $\mathfrak{m}$-a.e. in X for all $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$, which is precisely the standard Bakry–Émery curvature condition $\mathsf{BE}(K, \infty)$. We also observe that (3.9) naturally rephrases condition (G_2) in [97, Theorem 2.2(ii)] in our more general framework for $\tilde{d} = \mathsf{c}(t) \, d$ whenever $t \ge 0$.

Note that, if (3.9) holds for some everywhere finite measurable function $\mathsf{c} \colon [0, +\infty) \to [0, +\infty)$ (and so not necessarily locally positively bounded from above and below as in (3.8)), then we can replace it with another measurable function $\mathsf{c}_\star \colon [0, +\infty) \to [0, +\infty)$ which is optimal in the following sense: if $t > 0$ and $\mathsf{c}_\star(t) > 0$, then for all $\varepsilon > 0$ there exists $f_\varepsilon \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ such that

$$\mathfrak{m}\left(\left\{x \in X : \Gamma(\mathsf{P}_t f_\varepsilon)(x) \ge (\mathsf{c}_\star(t) - \varepsilon)^2 \, \mathsf{P}_t \Gamma(f_\varepsilon)(x)\right\}\right) > 0.$$

By (P.5), we immediately get that

$$\mathfrak{m}(\{x \in X : \Gamma(f)(x) > 0\}) > 0$$

whenever $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ is not $\mathfrak{m}$-equivalent to a constant function, and so $\mathsf{c}_\star(0) = 1$. In the following result we collect the elementary properties of $\mathsf{c}_\star$.

Lemma 3.5 (Properties of $\mathsf{c}_\star$). *The function* $\mathsf{c}_\star \colon [0, +\infty] \to [0, +\infty)$ *satisfies:*

(i) $\mathsf{c}_\star(s + t) \leq \mathsf{c}_\star(s)\,\mathsf{c}_\star(t)$ *for all* $s, t \geq 0$;
(ii) $\mathsf{c}_\star$ *is lower semicontinuous;*
(iii) $\mathsf{c}_\star(t) > 0$ *for all* $t \geq 0$.

Proof. Property (i) follows from the semigroup property of the heat flow and the optimality of $\mathsf{c}_\star$. Property (ii) is a consequence of Lemma 2.1 and again of the optimality of $\mathsf{c}_\star$. By (i), if $\mathsf{c}(t) = 0$ for some $t > 0$, then $\mathsf{c}(t') = 0$ for all $t' > t$. So let us set $t_0 = \inf\{t > 0 : \mathsf{c}_\star(t) = 0\}$. By (ii), we get that $\mathsf{c}_\star(t_0) = 0$. Since $\mathsf{c}_\star(0) = 1$, we must have that $t_0 > 0$. Now let $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}) \cap \mathrm{L}^1(X, \mathfrak{m})$ be non-negative, non-constant and such that $\int_X f \, \mathrm{d}\mathfrak{m} = 1$. Since $\mathsf{c}_\star(t_0) = 0$, we must have $\Gamma(\mathsf{P}_{t_0} f) = 0$ $\mathfrak{m}$-a.e. in X, so that $\mathsf{P}_{t_0} f(x) = a$ for all $x \in X$, for some $a \in \mathbb{R}$, by (P.5). By (2.38), we thus get that $a = 1$ and so we must have that $\mathfrak{m}(X) < +\infty$. Without loss of generality, we can assume $\mathfrak{m}(X) = 1$. Then, by (2.38) again and Jensen inequality, we get

$$1 = \left(\int_X f \, \mathrm{d}\mathfrak{m} \right)^2 = \left(\int_X \mathsf{P}_{t_0/2} f \, \mathrm{d}\mathfrak{m} \right)^2 \leq \int_X (\mathsf{P}_{t_0/2} f)^2 \, \mathrm{d}\mathfrak{m} = \int_X f \, \mathsf{P}_{t_0} f \, \mathrm{d}\mathfrak{m} = 1.$$

By the strict convexity of the square function, we thus get that $\mathsf{P}_{t_0/2} f = 1$ $\mathfrak{m}$-a.e. in X. Hence again $\int_X f \, \mathsf{P}_{t_0/2} f \, \mathrm{d}\mathfrak{m} = 1$, so that $\mathsf{P}_{t_0/2^n} f = 1$ $\mathfrak{m}$-a.e. in X for all $n \in \mathbb{N}$ by iterating the argument above. Thus $f = 1$ $\mathfrak{m}$-a.e. in X, contradicting the fact that the function f was taken non-constant. The proof is thus complete. $\qquad\square$

The following result is a simple consequence of well-known properties of subadditive functions (see [92, Chapter VII] for example), but we briefly sketch its proof here for the reader's convenience.

Lemma 3.6 (Local boundedness of $\mathsf{c}_\star$). *There exist* $t_\star \geq 0$ *and* $K \in \mathbb{R}$ *such that*

$$\mathsf{c}_\star(t) \leq e^{-Kt} \quad \text{for all } t \geq t_\star. \tag{3.10}$$

In addition, $\mathsf{c}_\star \in \mathrm{L}^\infty_{\mathrm{loc}}((0, +\infty))$ *and* $\mathsf{c}_\star^{-1} \in \mathrm{L}^\infty_{\mathrm{loc}}([0, +\infty))$.

Proof. We define $\varphi \colon [0, +\infty) \to \mathbb{R}$ by setting $\varphi(t) = \log \mathsf{c}_\star(t)$ for all $t \geq 0$. By Lemma 3.5, φ is well posed, lower semicontinuous and subadditive. By Fekete Lemma (see [92, Theorem 7.6.1] for example), we have that

$$\exists \lim_{t \to +\infty} \frac{\varphi(t)}{t} = \inf_{t > 0} \frac{\varphi(t)}{t} < +\infty, \tag{3.11}$$

from which we immediately deduce (3.10). By [92, Theorem 7.4.1], we have that $\varphi \in \mathrm{L}^\infty_{\mathrm{loc}}((0, +\infty))$. Since $\liminf_{t \to 0^+} \mathsf{c}_\star(t) \geq 1$ by Lemma 3.5(ii), we also get that $\mathsf{c}_\star(t) \geq M$ for all $t \in [0, \delta]$ for some $\delta, M > 0$, concluding the proof. $\qquad\square$

From Lemma 3.6, we easily deduce the following exponential upper bound for the optimal function $\mathsf{c}_\star$.

Corollary 3.7 (Exponential bound for $\mathsf{c}_\star$). *If* (3.9) *holds for some everywhere finite measurable function* $\mathsf{c} \colon [0, +\infty) \to [0, +\infty)$ *such that*

$$\limsup_{t \to 0^+} \mathsf{c}(t) < +\infty,$$

then the optimal function $c_\star \colon [0, +\infty) \to (0, +\infty)$ *fulfills* (3.8) *and is such that*

$$c_\star(t) \leq M e^{-Kt} \quad \text{for all } t \geq 0 \tag{3.12}$$

for some $M \geq 1$ *and* $K \in \mathbb{R}$.

By Corollary 3.7, in analogy with the classical Bakry–Émery condition, we may think of the (best) constant $K \in \mathbb{R}$ appearing in (3.12) as a bound from below of the generalized metric-measure Ricci curvature of the space $(X, \mathsf{d}, \mathfrak{m})$. In analogy with the usual RCD framework, we may say that $(X, \mathsf{d}, \mathfrak{m})$ is *negatively/zero/positively curved* if we can choose $K < 0 / K = 0 / K > 0$ in (3.12). By comparing (3.12) with (3.10) and (3.11), $\mathsf{BE}_w(c_\star, +\infty)$ condition behaves like $\mathsf{BE}(K, +\infty)$ for some limit $K \in \mathbb{R}$ as $t \to +\infty$.

3.4. **Poincaré inequalities.** Exploiting the differentiation formula in (3.5), we can prove the following consequence of Definition 3.4 in analogy with [16, Corollary 2.3]. See also [143, Theorem 1.1(3) and (4)] for the same inequalities in the Riemannian setting.

Proposition 3.8 (Poincaré inequalities)**.** *Assume* $(X, \mathsf{d}, \mathfrak{m})$ *satisfies* $\mathsf{BE}_w(c, \infty)$.

(i) *If* $f \in \mathrm{L}^2(X, \mathfrak{m})$ *and* $t > 0$, *then* $\mathsf{P}_t f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ *with*

$$2\,\mathrm{I}_{-2}(t)\,\Gamma(\mathsf{P}_t f) \leq \mathsf{P}_t(f^2) - (\mathsf{P}_t f)^2 \quad \mathfrak{m}\text{-}a.e.\ in\ X. \tag{3.13}$$

(ii) *If* $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ *and* $t > 0$, *then*

$$\mathsf{P}_t(f^2) - (\mathsf{P}_t f)^2 \leq 2\,\mathrm{I}_2(t)\,\mathsf{P}_t\Gamma(f) \quad \mathfrak{m}\text{-}a.e.\ in\ X. \tag{3.14}$$

Here and in the following, we let

$$\mathrm{I}_p(t) = \int_0^t c^p(s)\,\mathrm{d}s \tag{3.15}$$

for all $t \geq 0$ *and* $p \in \mathbb{R}$.

Proof. Fix $t > 0$ and $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$. By Lemma 3.3, we have

$$\begin{aligned}
\frac{1}{2}\int_X (\mathsf{P}_t(f^2) - (\mathsf{P}_t f)^2)\,\varphi\,\mathrm{d}\mathfrak{m} &= \frac{1}{2}\int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}\int_X (\mathsf{P}_{t-s}f)^2\,\mathsf{P}_s\varphi\,\mathrm{d}\mathfrak{m}\,\mathrm{d}s \\
&= \int_0^t \int_X \Gamma(\mathsf{P}_{t-s}f)\,\mathsf{P}_s\varphi\,\mathrm{d}\mathfrak{m}\,\mathrm{d}s \\
&= \int_0^t \int_X \mathsf{P}_s\Gamma(\mathsf{P}_{t-s}f)\,\varphi\,\mathrm{d}\mathfrak{m}\,\mathrm{d}s
\end{aligned}$$

for all $\varphi \in \mathrm{L}^2(X, \mathfrak{m}) \cap \mathrm{L}^\infty(X, \mathfrak{m})$. Now assume $\varphi \geq 0$. By the weak Bakry-Émery condition (3.9) and the semigroup property of the heat flow, we can estimate

$$c^{-2}(s)\int_X \Gamma(\mathsf{P}_t f)\,\varphi\,\mathrm{d}\mathfrak{m} \leq \int_X \mathsf{P}_s\Gamma(\mathsf{P}_{t-s}f)\,\varphi\,\mathrm{d}\mathfrak{m} \leq c^2(t-s)\int_X \mathsf{P}_t\Gamma(f)\,\varphi\,\mathrm{d}\mathfrak{m}$$

for all $s \in [0, t]$. Thus

$$\mathrm{I}_{-2}(t)\int_X \Gamma(\mathsf{P}_t f)\,\varphi\,\mathrm{d}\mathfrak{m} \leq \frac{1}{2}\int_X (\mathsf{P}_t(f^2) - (\mathsf{P}_t f)^2)\,\varphi\,\mathrm{d}\mathfrak{m} \leq \mathrm{I}_2(t)\int_X \mathsf{P}_t\Gamma(f)\,\varphi\,\mathrm{d}\mathfrak{m} \tag{3.16}$$

for all $\varphi \in \mathrm{L}^2(X, \mathfrak{m}) \cap \mathrm{L}^\infty(X, \mathfrak{m})$ such that $\varphi \geq 0$. In particular, we can choose $\varphi = \chi_E$ for any set $E \subset X$ with finite $\mathfrak{m}$-measure, so that inequalities (3.13) and (3.14) follow for all $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$.

Now assume $f \in \mathrm{L}^2(X, \mathfrak{m})$. By the density of $\mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ in $\mathrm{L}^2(X, \mathfrak{m})$, there exists $(f_n)_{n\in\mathbb{N}} \subset \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ such that $f_n \to f$ in $\mathrm{L}^2(X, \mathfrak{m})$ as $n \to +\infty$. As a consequence,

we have that $\mathsf{P}_t(f_n^2) \to \mathsf{P}_t(f^2)$ and $(\mathsf{P}_t f_n)^2 \to (\mathsf{P}_t f)^2$ in $\mathrm{L}^1(X, \mathfrak{m})$ as $n \to +\infty$. Moreover, by (2.40) and (2.46), we can estimate

$$\int_X \left| |\mathrm{D}\mathsf{P}_t f_n|_w - |\mathrm{D}\mathsf{P}_t f|_w \right|^2 \mathrm{d}\mathfrak{m} \leq \int_X \Gamma(\mathsf{P}_t(f_n - f)) \, \mathrm{d}\mathfrak{m}$$
$$= 2\mathsf{Ch}(\mathsf{P}_t(f_n - f))$$
$$\leq \frac{1}{t} \int_X |f_n - f|^2 \, \mathrm{d}\mathfrak{m}$$

for all $n \in \mathbb{N}$, so that $\Gamma(\mathsf{P}_t f_n) \to \Gamma(\mathsf{P}_t f)$ in $\mathrm{L}^1(X, \mathfrak{m})$ as $n \to +\infty$. By the first inequality in (3.16), we have that

$$\mathrm{I}_{-2}(t) \int_X \Gamma(\mathsf{P}_t f_n) \, \varphi \, \mathrm{d}\mathfrak{m} \leq \frac{1}{2} \int_X \left(\mathsf{P}_t(f_n^2) - (\mathsf{P}_t f_n)^2 \right) \varphi \, \mathrm{d}\mathfrak{m} \tag{3.17}$$

for all $\varphi \in \mathrm{L}^2(X, \mathfrak{m}) \cap \mathrm{L}^\infty(X, \mathfrak{m})$ such that $\varphi \geq 0$. Passing to the limit as $n \to +\infty$ in (3.17) and arguing as before, we get (3.13). □

3.5. **BE_w inequality for Lipschitz functions.** Thanks to Proposition 3.8, we can prove that the heat flow of a bounded Lipschitz functions in $\mathrm{L}^2(X, \mathfrak{m})$ has a Lipschitz representative with controlled Lipschitz constant.

Proposition 3.9 (BE_w for Lip-functions, I)**.** *Assume $(X, \mathsf{d}, \mathfrak{m})$ satisfies $\mathsf{BE}_w(\mathsf{c}, \infty)$. If $f \in \mathrm{Lip}_b(X, \mathsf{d}) \cap \mathrm{L}^2(X, \mathfrak{m})$, then $\mathsf{P}_t f \in \mathrm{Lip}_b(X, \mathsf{d}) \cap \mathrm{L}^2(X, \mathfrak{m})$ with*

$$\mathrm{Lip}(\mathsf{P}_t f) \leq \mathsf{c}(t) \, \mathrm{Lip}(f) \tag{3.18}$$

for all $t \geq 0$

Proof. Let $t \geq 0$ be fixed. By (2.34) and (3.13), we know that $\mathsf{P}_t f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}) \cap \mathrm{L}^\infty(X, \mathfrak{m})$ with

$$2 \, \mathrm{I}_{-2}(t) \, \Gamma(\mathsf{P}_t f) \leq \mathsf{P}_t(f^2) \leq \|f\|_{\mathrm{L}^\infty(X, \mathfrak{m})}^2 \quad \mathfrak{m}\text{-a.e. in } X.$$

Thus, recalling property (P.5), $\mathsf{P}_t f$ coincides $\mathfrak{m}$-a.e. in X with a (bounded) Lipschitz function. We now divide the proof in two steps.

Step 1. Assume that $\mathrm{supp}\, f$ is bounded. Then $f \in \mathrm{Lip}_b(X, \mathsf{d}) \cap \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$. Hence, by applying the weak Bakry–Émery condition (3.9) to f, we find that

$$\Gamma(\mathsf{P}_t f) \leq \mathsf{c}^2(t) \mathsf{P}_t \Gamma(f) \leq \mathsf{c}^2(t) \, \mathrm{Lip}(f)^2 \quad \mathfrak{m}\text{-a.e. in } X.$$

Thus, again by property (P.5), we get (3.18).

Step 2. Now fix $x_0 \in X$ and let $R > 0$. We let $\eta_{x_0, R} \colon X \to [0, 1]$ be such that

$$\eta_{x_0, R}(x) = \left(1 - \frac{\mathsf{d}(x, x_0)}{R} \right)^+ \quad \text{for all } x \in X. \tag{3.19}$$

Note that $\eta_{x_0, R} \in \mathrm{Lip}_b(X, \mathsf{d})$ with

$$\mathrm{supp}\, \eta_{x_0, R} \subset \overline{B_R(x_0)}, \quad |\mathrm{D}\eta_{x_0, R}| \leq \frac{1}{R} \chi_{\overline{B_R(x_0)}}, \quad \mathrm{Lip}(\eta_{x_0, R}) \leq \frac{1}{R}$$

for all $R > 0$. Let us set $f_n = f \eta_{x_0, n}$ for all $n \in \mathbb{N}$. Then $f_n \in \mathrm{Lip}_b(X, \mathsf{d}) \cap \mathrm{L}^2(X, \mathfrak{m})$ has bounded support and is such that

$$\|f_n\|_{\mathrm{L}^\infty(X, \mathsf{d})} \leq \|f\|_{\mathrm{L}^\infty(X, \mathsf{d})}, \quad \mathrm{Lip}(f_n) \leq \mathrm{Lip}(f) + \frac{1}{n} \|f\|_{\mathrm{L}^\infty(X, \mathsf{d})}$$

for all $n \in \mathbb{N}$. By Step 1, we get that

$$\mathrm{Lip}(\mathsf{P}_t f_n) \leq \mathsf{c}(t) \, \mathrm{Lip}(f_n) \tag{3.20}$$

for all $n \in \mathbb{N}$. Since $(f_n)_{n \in \mathbb{N}}$ is equi-bounded and equi-Lipschitz, by (2.34) and (3.20) also $(\mathsf{P}_t f_n)_{n \in \mathbb{N}}$ is equi-bounded and equi-Lipschitz. Hence $(\mathsf{P}_t f_n)_{n \in \mathbb{N}}$ converges locally uniformly to a (bounded) Lipschitz function g_t with $\mathrm{Lip}(g_t) \leq \mathsf{c}(t) \, \mathrm{Lip}(f)$. Since $f_n \to f$ in $\mathrm{L}^2(X, \mathfrak{m})$ as $n \to +\infty$, then $\mathsf{P}_t f_n \to \mathsf{P}_t f$ in $\mathrm{L}^2(X, \mathfrak{m})$ as $n \to +\infty$ and thus, up to subsequences, $\mathsf{P}_t f_n \to \mathsf{P}_t f$ $\mathfrak{m}$-a.e. in X. But then $g_t = \mathsf{P}_t f$ $\mathfrak{m}$-a.e. in X and thus (3.18) readily follows. $\qquad \square$

3.6. **Dual heat semigroup.** Thanks to (2.38), we can define the *dual heat semigroup* $\mathsf{H}_t \colon \mathscr{P}^{\mathrm{ac}}(X) \to \mathscr{P}^{\mathrm{ac}}(X)$ for all $t \geq 0$ on absolutely continuous probability measures by setting

$$\mathsf{H}_t \mu = (\mathsf{P}_t f) \, \mathfrak{m} \quad \text{for all } \mu = f \mathfrak{m} \in \mathscr{P}^{\mathrm{ac}}(X). \tag{3.21}$$

Note that $(\mathsf{H}_t)_{t \geq 0}$ is a *linearly convex* semigroup on $\mathscr{P}^{\mathrm{ac}}(X)$, in the sense that

$$\mathsf{H}_{s+t} \mu = \mathsf{H}_s(\mathsf{H}_t \mu) \tag{3.22}$$

and

$$\mathsf{H}_t((1 - \lambda)\mu + \lambda\nu) = (1 - \lambda)\mathsf{H}_t\mu + \lambda\mathsf{H}_t\nu \quad \text{for all } \lambda \in [0, 1] \tag{3.23}$$

whenever $s, t \geq 0$ and $\mu, \nu \in \mathscr{P}^{\mathrm{ac}}(X)$. Note that (3.21) is well posed without assuming the $\mathsf{BE}_w(\mathsf{c}, \infty)$ condition.

The following result (which still does not require the $\mathsf{BE}_w(\mathsf{c}, \infty)$ condition) proves that the dual heat semigroup preserves the finiteness of the second moments of the measures in the domain of the entropy.

Lemma 3.10 (Second moment estimate). *If $\mu = f\mathfrak{m} \in \mathrm{Dom}(\mathsf{Ent}_\mathfrak{m})$, then $\mu_t = \mathsf{H}_t\mu \in \mathscr{P}_2^{\mathrm{ac}}(X)$ with*

$$\int_X \mathsf{d}^2(x, x_0) \, \mathrm{d}\mu_t \leq e^{4t} \left(\mathsf{Ent}_\mathfrak{m}(\mu) + 2 \int_X \mathsf{d}^2(x, x_0) \, \mathrm{d}\mu \right) \tag{3.24}$$

for all $t \geq 0$, whenever $x_0 \in X$ is given. In particular, $\mathsf{H}_t(\mathrm{Dom}(\mathsf{Ent}_\mathfrak{m})) \subset \mathscr{P}_p^{\mathrm{ac}}(X)$ for all $p \in [1, 2]$ and $t \geq 0$.

Proof. Fix $x_0 \in X$ and set $V(x) = \mathsf{d}(x, x_0)$ for all $x \in X$. Then $V \in \mathrm{Lip}(X)$ with $\mathrm{Lip}(V) \leq 1$ and (3.24) follows by [14, Theorem 4.20] if $f \in \mathrm{L}^2(X, \mathfrak{m})$. The conclusion then follows by considering $f_n = q_n^{-1}(f \wedge n) \in \mathrm{L}^2(X, \mathfrak{m})$, $q_n = \int_X f \wedge n \, \mathrm{d}\mathfrak{m}$, $n \in \mathbb{N}$, and passing to the limit as $n \to +\infty$ by the Monotone Convergence Theorem. Since $\mathscr{P}_2(X) \subset \mathscr{P}_p(X)$ by Jensen inequality, the proof is complete. $\qquad \square$

The following simple result provides a useful sufficient condition to extend the dual heat semigroup to a W_p-Lipschitz map on the whole p-Wasserstein space for $p \in [1, 2]$.

Lemma 3.11 (Lip-extension of H_t on $\mathscr{P}_p(X)$ for $p \in [1, 2]$). *Let $p \in [1, 2]$ and let $C \colon [0, +\infty) \to (0, +\infty)$ be locally bounded. If the dual heat semigroup defined in (3.21) satisfies*

$$W_p(\mathsf{H}_t\mu, \mathsf{H}_t\nu) \leq C(t) \, W_p(\mu, \nu) \quad \text{for all } \mu, \nu \in \mathscr{D}$$

for some W_p-dense subset $\mathscr{D}$ of $\mathscr{P}_p(X)$, then it uniquely extends to a W_p-Lipschitz map (for which we retain the same notation) $\mathsf{H}_t \colon \mathscr{P}_p(X) \to \mathscr{P}_p(X)$ such that

$$W_p(\mathsf{H}_t\mu, \mathsf{H}_t\nu) \leq C(t) \, W_p(\mu, \nu) \quad \text{for all } \mu, \nu \in \mathscr{P}_p(X).$$

In addition, the maps $(\mathsf{H}_t)_{t \geq 0} \colon \mathscr{P}_p(X) \to \mathscr{P}_p(X)$ still satisfy (3.22) and (3.23) for all $\mu, \nu \in \mathscr{P}_p(X)$.

Proof. The extension of the dual heat semigroup readily follows from the W_p-density of $\mathscr{D}$ in $\mathscr{P}_p(X)$ and the completeness of the p-Wasserstein space. The validity of (3.22) and (3.23) for all $\mu, \nu \in \mathscr{P}_p(X)$ is a direct consequence of the joint convexity of the p-Wasserstein distance. $\square$

In the following result, we prove that the $\mathsf{BE}_w(\mathsf{c}, \infty)$ condition implies that the dual heat semigroup can be extended to a W_1-Lipschitz map on the whole 1-Wasserstein space. See [16, Proposition 3.2(i)], [97, Proposition 3.7] and [98, Theorem 2.2(ii)].

Proposition 3.12 ($\mathsf{BE}_w(\mathsf{c}, \infty) \Rightarrow \mathsf{H}_t$ is W_1-Lip)**.** *Assume $(X, \mathsf{d}, \mathfrak{m})$ satisfies $\mathsf{BE}_w(\mathsf{c}, \infty)$. For $t \geq 0$, the dual heat semigroup (3.21) uniquely extends to a W_1-Lipschitz map with*

$$W_1(\mathsf{H}_t\mu, \mathsf{H}_t\nu) \leq \mathsf{c}(t)\, W_1(\mu, \nu) \quad \text{for all } \mu, \nu \in \mathscr{P}_1(X). \tag{3.25}$$

Proof. Let $t \geq 0$ be fixed. Thanks to Lemma 3.11, we just need to prove that

$$W_1(\mathsf{H}_t\mu, \mathsf{H}_t\nu) \leq \mathsf{c}(t)\, W_1(\mu, \nu) \quad \text{for all } \mu, \nu \in \mathrm{Dom}(\mathsf{Ent}_\mathfrak{m}). \tag{3.26}$$

So let $\mu, \nu \in \mathrm{Dom}(\mathsf{Ent}_\mathfrak{m})$ with $\mu = f\mathfrak{m}$ and $\nu = g\mathfrak{m}$. Let $\varphi \in \mathrm{Lip}(X)$ with $\mathrm{Lip}(\varphi) \leq 1$ and bounded support. Then $\varphi \in \mathrm{Lip}_b(X) \cap \mathrm{L}^2(X, \mathfrak{m})$ and so $\varphi_t = \mathsf{P}_t\varphi \in \mathrm{Lip}_b(X) \cap \mathrm{L}^2(X, \mathfrak{m})$ with $\mathrm{Lip}(\varphi_t) \leq \mathsf{c}(t)$ for all $t \geq 0$ by Proposition 3.9. Thus, setting $\mu_t = \mathsf{H}_t\mu$ and $\nu_t = \mathsf{H}_t\nu$, we can estimate

$$\int_X \varphi \, \mathrm{d}(\mu_t - \nu_t) = \int_X \varphi \, f_t \, \mathrm{d}\mathfrak{m} - \int_X \varphi \, g_t \, \mathrm{d}\mathfrak{m} = \int_X \varphi_t \, f \, \mathrm{d}\mathfrak{m} - \int_X \varphi_t \, g \, \mathrm{d}\mathfrak{m}$$

$$= \int_X \varphi_t \, \mathrm{d}(\mu - \nu) \leq \mathsf{c}(t)\, W_1(\mu, \nu)$$

for all $t \geq 0$ by (2.10). Passing to the supremum on all $\varphi \in \mathrm{Lip}(X)$ with $\mathrm{Lip}(\varphi) \leq 1$ and bounded support, by (2.12) we get (3.26) and the proof is complete. $\square$

3.7. **Pointwise version of the heat semigroup.** Let $(\mathsf{H}_t)_{t \geq 0}$ be the dual heat semigroup defined from $\mathscr{P}^{\mathrm{ac}}(X)$ to itself as in (3.21) and assume that, for some $p \in [1, 2]$,

$$(\mathsf{H}_t)_{t \geq 0} \text{ admits a unique } W_p\text{-continuous extension from } \mathscr{P}_p(X) \text{ to itself.} \tag{3.27}$$

If (3.27) holds, then for all $t \geq 0$ we can define the (everywhere defined) *pointwise version* of the heat semigroup

$$\tilde{\mathsf{P}}_t f(x) = \int_X f \, \mathrm{d}\mathsf{H}_t\delta_x, \quad x \in X, \tag{3.28}$$

whenever $f \colon X \to \overline{\mathbb{R}}$ is either a bounded or a non-negative Borel function. Note that, by the very definition (3.28), $(\tilde{\mathsf{P}}_t)_{t \geq 0}$ defines a linear semigroup of L^∞-contractions on bounded Borel functions, in the sense that

$$\tilde{\mathsf{P}}_{s+t}f = \tilde{\mathsf{P}}_s(\tilde{\mathsf{P}}_t f)$$

and

$$\|\tilde{\mathsf{P}}_t f\|_{\mathrm{L}^\infty(X, \mathfrak{m})} \leq \|f\|_{\mathrm{L}^\infty(X, \mathfrak{m})} \tag{3.29}$$

whenever $f \in \mathrm{L}^\infty(X, \mathfrak{m})$ is Borel and $s, t \geq 0$. The following result lists the main properties of $(\tilde{\mathsf{P}}_t)_{t \geq 0}$, see [16, Proposition 3.2(ii) and (iii)].

Proposition 3.13 (Properties of $\tilde{\mathsf{P}}_t$)**.** *Assume (3.27) holds for some $p \in [1, 2]$.*

    (i) *If $f \in \mathrm{L}^\infty(X, \mathfrak{m})$ is lower (resp., upper) semicontinuous, then $\tilde{\mathsf{P}}_t f$ is lower (resp., upper) semicontinuous for all $t \geq 0$. As a consequence, if $f \in \mathrm{C}_b(X)$, then $\tilde{\mathsf{P}}_t f \in \mathrm{C}_b(X)$ for all $t \geq 0$.*

    (ii) *If $f \in \mathrm{L}^\infty(X, \mathfrak{m})$ is Borel, then $\tilde{\mathsf{P}}_t f = \mathsf{P}_t f$ $\mathfrak{m}$-a.e. in X for all $t \geq 0$.*

(iii) If $f \in L^\infty(X, \mathfrak{m})$ is Borel and $\mu \in \mathscr{P}_p(X)$, then

$$\int_X \tilde{\mathsf{P}}_t f \, \mathrm{d}\mu = \int_X f \, \mathrm{d}\mathsf{H}_t \mu \tag{3.30}$$

for all $t \geq 0$.

Proof. Let $t \geq 0$ be fixed. We prove the three statements separately.

Proof of (i). By the linearity of $\tilde{\mathsf{P}}_t$ and P_t, we can assume that $f \geq 0$ without loss of generality. Let f be lower semicontinuous. If $x_n \to x$ in X as $n \to +\infty$, then $\delta_{x_n} \to \delta_x$ in $\mathscr{P}_p(X)$ as $n \to +\infty$ and thus $\mathsf{H}_t \delta_{x_n} \to \mathsf{H}_t \delta_x$ in $\mathscr{P}_p(X)$ as $n \to +\infty$ by (3.27). Hence $\mathsf{H}_t \delta_{x_n} \rightharpoonup \mathsf{H}_t \delta_x$ in $\mathscr{P}(X)$ as $n \to +\infty$ and thus

$$\mathsf{H}_t \delta_x(\{f > t\}) \leq \liminf_{n \to +\infty} \mathsf{H}_t \delta_{x_n}(\{f > t\}).$$

Thus

$$\tilde{\mathsf{P}}_t f(x) = \int_0^{+\infty} \mathsf{H}_t \delta_x(\{f > t\}) \, \mathrm{d}t \leq \liminf_{n \to +\infty} \int_0^{+\infty} \mathsf{H}_t \delta_{x_n}(\{f > t\}) \, \mathrm{d}t = \liminf_{n \to +\infty} \tilde{\mathsf{P}}_t f(x_n)$$

by Fatou Lemma. The proof is similar in the case f is upper semicontinuous. This proves (i).

Proof of (ii). We divide the proof in two steps.

Step 1. Let $g \in L^1(X, \mathfrak{m})$ have compact support $K = \operatorname{supp} g$ and be such that $\mu = g\mathfrak{m} \in \mathscr{P}_p(X)$. We claim that

$$\mathsf{H}_t(g\mathfrak{m}) = \int_X g(x) \, \mathsf{H}_t \delta_x \, \mathrm{d}\mathfrak{m}(x) \quad \text{in } \mathscr{P}_p(X). \tag{3.31}$$

Indeed, given $n \in \mathbb{N}$, we can find n points $p_1^n, \dots, p_n^n \in K$ such that

$$K \subset \bigcup_{i=1}^n B\left(p_i^n, \tfrac{1}{n}\right).$$

Let us set

$$A_1^n = B\left(p_1^n, \tfrac{1}{n}\right), \quad A_i^n = B\left(x_i^n, \tfrac{1}{n}\right) \setminus \bigcup_{j<i} B\left(p_j^n, \tfrac{1}{n}\right) \text{ for all } i = 2, \dots, n,$$

so that $\operatorname{diam} A_i^n \leq \tfrac{2}{n}$ for all $i = 1, \dots, n$. Without loss of generality, we can assume that $A_i^n \neq \emptyset$ for all $i = 1, \dots, n$, so that we can choose some $x_i^n \in A_i^n$ for all $i = 1, \dots, n$. Let us set

$$a_{i,n} = \int_{A_i^n} g \, \mathrm{d}\mathfrak{m} \quad \text{for all } i = 1, \dots, n,$$

so that $\sum_{i=1}^n a_{i,n} = 1$. We thus define

$$\mu_n = \sum_{i=1}^n a_{i,n} \delta_{x_i^n} \in \mathscr{P}_p(X) \quad \text{for all } n \in \mathbb{N}.$$

We have that $\mu_n \xrightarrow{W_p} \mu$ as $n \to +\infty$. Indeed, if $\varphi \in C(X)$, then φ is uniformly continuous on K and thus

$$\limsup_{n \to +\infty} \left| \int_X \varphi \, \mathrm{d}\mu_n - \int_X \varphi \, \mathrm{d}\mu \right| \leq \limsup_{n \to +\infty} \sum_{i=1}^n \int_{A_i^n} |\varphi(x_i^n) - \varphi(x)| \, \mathrm{d}\mu(x) = 0.$$

By (3.27) we thus get that $\nu_n = \mathsf{H}_t \mu_n \xrightarrow{W_p} \nu = \mathsf{H}_t \mu$ as $n \to +\infty$. Let us set

$$\tilde{\nu} = \int_X \mathsf{H}_t \delta_x \, \mathrm{d}\mu(x) \in \mathscr{P}_p(X),$$

which is well posed thanks to (3.27) and the compactness of K. By using twice the joint convexity of the p-Wasserstein distance (see [141, Theorem 4.8]), we can estimate

$$W_p(\nu_n, \tilde{\nu}) = W_p \left(\sum_{i=1}^n \int_{A_i^n} \mathsf{H}_t \delta_{x_i^n} \, \mathrm{d}\mathfrak{m}(x), \sum_{i=1}^n \int_{A_i^n} \mathsf{H}_t \delta_x \, \mathrm{d}\mathfrak{m}(x) \right)$$

$$\leq \sum_{i=1}^n \mathfrak{m}(A_i^n) \, W_p \left(\fint_{A_i^n} \mathsf{H}_t \delta_{x_i^n} \, \mathrm{d}\mathfrak{m}(x), \fint_{A_i^n} \mathsf{H}_t \delta_x \, \mathrm{d}\mathfrak{m}(x) \right)$$

$$\leq \sum_{i=1}^n \fint_{A_i^n} W_p \left(\mathsf{H}_t \delta_{x_i^n}, \mathsf{H}_t \delta_x \right) \mathrm{d}\mathfrak{m}(x)$$

for all $n \in \mathbb{N}$, so that $\nu_n \xrightarrow{W_p} \tilde{\nu}$ as $n \to +\infty$ by the uniform W_p-continuity of the map $x \mapsto \mathsf{H}_t \delta_x$ on the compact set K. Thus $\tilde{\nu} = \nu$ in $\mathscr{P}_p(X)$ and the claim (3.31) follows.

Step 2. Thanks to Step 1 and Fubini Theorem, we can compute

$$\int_X \mathsf{P}_t f \, g \, \mathrm{d}\mathfrak{m} = \int_X f \, \mathsf{P}_t g \, \mathrm{d}\mathfrak{m}$$

$$= \int_X f \, \mathrm{d}\mathsf{H}_t(g\mathfrak{m})$$

$$= \int_X g(x) \int_X f(y) \, \mathrm{d}\mathsf{H}_t \delta_x(y) \, \mathrm{d}\mathfrak{m}(x)$$

$$= \int_X \tilde{\mathsf{P}}_t f \, g \, \mathrm{d}\mathfrak{m}$$

for all $g \in \mathrm{L}^1(X, \mathfrak{m})$ with compact support and such that $\mu = g\mathfrak{m} \in \mathscr{P}_p(X)$. Hence

$$\int_X \mathsf{P}_t f \, g \, \mathrm{d}\mathfrak{m} = \int_X \tilde{\mathsf{P}}_t f \, g \, \mathrm{d}\mathfrak{m}$$

for all non-negative $g \in \mathrm{L}^1(X, \mathfrak{m})$ and (ii) immediately follows.

Proof of (iii). We divide the proof in two steps.

Step 1. Assume $f \in \mathrm{C}_b(X)$. By (ii), we thus know that $\tilde{\mathsf{P}}_t f = \mathsf{P}_t f$ $\mathfrak{m}$-a.e. in X, so that we can compute

$$\int_X \tilde{\mathsf{P}}_t f \, \mathrm{d}\mu = \int_X \mathsf{P}_t f \, g \, \mathrm{d}\mathfrak{m} = \int_X f \, \mathsf{P}_t g \, \mathrm{d}\mathfrak{m} = \int_X f \, \mathrm{d}\mathsf{H}_t \mu \tag{3.32}$$

for all $\mu = g\mathfrak{m} \in \mathscr{P}_p^{\mathrm{ac}}(X)$. Now if $\mu \in \mathscr{P}_p(X)$, then we can find $\mu_n \in \mathscr{P}_p^{\mathrm{ac}}(X)$ such that $\mu_n \xrightarrow{W_p} \mu$ as $n \to +\infty$. By (3.27), we also have $\mathsf{H}_t \mu_n \xrightarrow{W_p} \mathsf{H}_t \mu$ as $n \to +\infty$, and thus by (3.32) and (i) we get

$$\int_X \tilde{\mathsf{P}}_t f \, \mathrm{d}\mu = \lim_{n \to +\infty} \int_X \tilde{\mathsf{P}}_t f \, \mathrm{d}\mu_n = \lim_{n \to +\infty} \int_X f \, \mathrm{d}\mathsf{H}_t \mu_n = \int_X f \, \mathrm{d}\mathsf{H}_t \mu,$$

proving (3.30) whenever $f \in \mathrm{C}_b(X)$.

Step 2. Let $K \subset X$ be a non-empty bounded closed set. For each $n \in \mathbb{N}$, let us set

$$f_n(x) = [1 - n \, \mathrm{dist}(x, K)]^+ \quad \text{for all } x \in X.$$

Then $f_n \in \mathrm{C}_b(X)$ and $\chi_K \leq f_n \leq \chi_H$ for all $n \in \mathbb{N}$, where $H = \{x \in X : \mathrm{dist}(x, K) \leq 1\}$, and $f_n(x) \downarrow \chi_K(x)$ for all $x \in X$ as $n \to +\infty$. Hence $\tilde{\mathsf{P}}_t f_n(x) \downarrow \tilde{\mathsf{P}}_t \chi_K(x)$ for all $x \in X$ as $n \to +\infty$ and thus

$$\int_X \tilde{\mathsf{P}}_t \chi_K \, \mathrm{d}\mathfrak{m} = \lim_{n \to +\infty} \int_X \tilde{\mathsf{P}}_t f_n \, \mathrm{d}\mu = \lim_{n \to +\infty} \int_X f_n \, \mathrm{d}\mathsf{H}_t \mu = \int_X \chi_K \, \mathrm{d}\mathsf{H}_t \mu$$

by Step 1 and the Monotone Convergence Theorem. Thus (iii) follows by the Monotone Class Theorem (see [70, Theorem 5.2.2] for example). □

3.8. BE_w **inequality for Lipschitz functions, refined.** We now refine Proposition 3.9 to the following Proposition 3.14, where we prove an everywhere pointwise gradient bound for the heat flow starting from Lipschitz functions.

Proposition 3.14 (BE_w for Lip-functions, II)**.** *Assume* $(X, \mathsf{d}, \mathfrak{m})$ *satisfies* $\mathsf{BE}_w(\mathsf{c}, \infty)$*. If* $f \in \mathrm{Lip}_b(X, \mathsf{d}) \cap \mathrm{L}^2(X, \mathfrak{m})$ *with* $|\mathrm{D}^* f| \in \mathrm{L}^2(X, \mathfrak{m})$*, then* $\tilde{\mathsf{P}}_t f \in \mathrm{Lip}_b(X, \mathsf{d}) \cap \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ *with*

$$|\mathrm{D}^* \tilde{\mathsf{P}}_t f|^2(x) \leq \mathsf{c}^2(t) \tilde{\mathsf{P}}_t(|\mathrm{D}^* f|^2)(x) \quad \textit{for all } x \in X \qquad (3.33)$$

for all $t \geq 0$*.*

In the proof of Proposition 3.14, we need the following technical result, see [16, Proposition 3.11]. For the reader's convenience, we recall its short proof here.

Lemma 3.15 (Reverse slope estimate)**.** *Let* $f \in \mathrm{C}_b(X) \cap \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$*. If* $\Gamma(f) \leq G^2$ $\mathfrak{m}$*-a.e. in* X *for some upper semicontinuous* $G \in \mathrm{L}^\infty(X, \mathfrak{m})$*, then* $f \in \mathrm{Lip}(X)$ *and* $|\mathrm{D}^* f|(x) \leq G(x)$ *for all* $x \in X$*.*

Proof. Since $G \in \mathrm{L}^\infty(X, \mathfrak{m})$, we immediately get $f \in \mathrm{Lip}(X)$ from property (P.5). Let $x \in X$ be fixed. For all $\varepsilon > 0$, set $G_\varepsilon = \sup_{y \in B_\varepsilon(x)} \zeta(y)$ and define

$$\psi_\varepsilon(y) = \min\left\{\max\{|f(y) - f(x)|, G_\varepsilon \mathsf{d}(y, x)\}, G_\varepsilon[\varepsilon - \mathsf{d}(y, x)]^+\right\} \quad \text{for all } y \in X.$$

Then $\psi_\varepsilon \in \mathrm{Lip}(X)$ with $\mathrm{supp}\,\psi \subset \overline{B_\varepsilon(x)}$, so that $\psi_\varepsilon \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$. By (2.26) and (2.27), we have $|\mathrm{D}\psi_\varepsilon|_w \leq \max\{\zeta, G_\varepsilon\}$ $\mathfrak{m}$-a.e. in X. Since $\psi_\varepsilon(y) = 0$ for $\mathsf{d}(y, x) \geq \varepsilon$, we must have that $|\mathrm{D}\psi_\varepsilon|_w \leq G_\varepsilon$ $\mathfrak{m}$-a.e. in X. Again by (P.5), we get $\mathrm{Lip}(\psi_\varepsilon) \leq G_\varepsilon$. Since $\psi_\varepsilon(x) = 0$, we conclude that

$$\psi_\varepsilon(y) \leq G_\varepsilon \mathsf{d}(y, x) \quad \text{for all } y \in X. \qquad (3.34)$$

Now, if $\mathsf{d}(y, x) < \frac{\varepsilon}{2}$, then $[\varepsilon - \mathsf{d}(y, x)]^+ > \frac{\varepsilon}{2}$ and $\psi_\varepsilon(y) < \frac{G_\varepsilon \varepsilon}{2}$ by (3.34), so that

$$|f(y) - f(x)| \leq \psi_\varepsilon(y) \leq G_\varepsilon \mathsf{d}(y, x).$$

Hence

$$|\mathrm{D}f|(x) = \limsup_{y \to x} \frac{|f(y) - f(x)|}{\mathsf{d}(y, x)} \leq \limsup_{y \to x} \frac{G_\varepsilon \mathsf{d}(y, x)}{\mathsf{d}(y, x)} = G_\varepsilon.$$

Since ζ is upper semicontinuous, we have $\lim_{\varepsilon \to 0^+} G_\varepsilon = \zeta(x)$ and thus $|\mathrm{D}f|(x) \leq \zeta(x)$ whenever $x \in X$. Since (X, d) is a length space by Proposition 2.2, again by the upper semicontinuity of ζ we also get $|\mathrm{D}^* f| \leq \zeta$. □

Proof of Proposition 3.14. Let $t \geq 0$ be fixed. By Proposition 3.9, we know that $\mathsf{P}_t f \in \mathrm{Lip}_b(X, \mathsf{d}) \cap \mathrm{L}^2(X, \mathfrak{m})$. By Proposition 3.13, we thus know that the continuous representative of $\mathsf{P}_t f$ coincides with $\tilde{\mathsf{P}}_t f$. Since $|\mathrm{D}^* f| \in \mathrm{L}^2(X, \mathfrak{m})$, we have $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ and thus, by (2.29), $|\mathrm{D}f|_w \leq |\mathrm{D}^* f|$ $\mathfrak{m}$-a.e. in X. Therefore, thanks to (2.35) and (3.18), we can estimate

$$\Gamma(\tilde{\mathsf{P}}_t f) \leq \mathsf{c}^2(t) \mathsf{P}_t \Gamma(f) \leq \mathsf{c}^2(t) \mathsf{P}_t(|\mathrm{D}^* f|^2) \quad \mathfrak{m}\text{-a.e. in } X$$

and hence, since $\tilde{\mathsf{P}}_t(|\mathrm{D}^* f|^2) = \mathsf{P}_t(|\mathrm{D}^* f|^2)$ $\mathfrak{m}$-a.e. in X by Proposition 3.13(ii), we get

$$\Gamma(\tilde{\mathsf{P}}_t f) \leq \mathsf{c}^2(t) \tilde{\mathsf{P}}_t(|\mathrm{D}^* f|^2) \quad \mathfrak{m}\text{-a.e. in } X.$$

Since the function $x \mapsto \tilde{\mathsf{P}}_t(|\mathrm{D}^* f|^2)$ is bounded and upper semicontinuous by Proposition 3.13(i), inequality (3.33) immediately follows by Lemma 3.15. □

3.9. **Kuwada duality.** We now come the the main result of this section, the equivalence between the weak Bakry–Émery inequality and the W_2-contractivity property of the (dual) heat semigroup. This duality property is well known for the $\mathsf{BE}(K, \infty)$ condition and is due to Kuwada, see the pioneering works [97–99]. This duality is also known for the stronger $\mathsf{BE}(K, N)$ condition (with $N < +\infty$), see [72]. In a very general framework, this equivalence has been obatined in [16, Theorem 3.5 and Corollary 3.18].

Theorem 3.16 (Kuwada duality)**.** *The following are equivalent.*
  (i) $(X, \mathsf{d}, \mathfrak{m})$ *satisfies* $\mathsf{BE}_w(\mathsf{c}, \infty)$.
  (ii) *There exists a W_2-dense subset $\mathscr{D}$ of $\mathscr{P}_2^{\mathrm{ac}}(X)$ such that $\mathsf{H}_t(\mathscr{D}) \subset \mathscr{P}_2^{\mathrm{ac}}(X)$ and*

$$W_2((\mathsf{P}_t f)\mathfrak{m}, (\mathsf{P}_t g)\mathfrak{m}) \leq \mathsf{c}(t)\, W_2(f\mathfrak{m}, g\mathfrak{m}) \quad \text{whenever } f\mathfrak{m}, g\mathfrak{m} \in \mathscr{D} \tag{3.35}$$

  for all $t \geq 0$.

If either (i) or (ii) holds, then for all $t \geq 0$ the dual heat semigroup (3.21) uniquely extends to a map $\mathsf{H}_t \colon \mathscr{P}_2(X) \to \mathscr{P}_2(X)$ such that

$$W_2(\mathsf{H}_t\mu, \mathsf{H}_t\nu) \leq \mathsf{c}(t)\, W_2(\mu, \nu) \quad \text{for all } \mu, \nu \in \mathscr{P}_2(X). \tag{3.36}$$

In the proof of Theorem 3.16, we will need the following two technical results.

The first one will be use in the proof of the implication (i) $\Rightarrow$ (ii) and was proved for the first time in [97]. In the present framework, this result was proved in [16, Lemma 3.4]. For the reader's convenience, we provide a proof of it below. Here and in the following, we let

$$\mathrm{Lip}_\star(X) = \{f \in \mathrm{Lip}(X) : \mathrm{supp}\, f \text{ is bounded and } f \geq 0\}. \tag{3.37}$$

Thanks to (2.29), we immediately see that $\mathrm{Lip}_\star(X) \subset \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$. Moreover, from its very definition (2.4), we see that the Hopf–Lax semigroup satisfies $Q_s(\mathrm{Lip}_\star(X)) \subset \mathrm{Lip}_\star(X)$ for all $s \geq 0$, since if $f \in \mathrm{Lip}_\star(X)$ then $\mathrm{supp}(Q_s f) \subset \mathrm{supp}\, f$ for all $s \geq 0$.

Lemma 3.17 (Kuwada estimate)**.** *Assume $(X, \mathsf{d}, \mathfrak{m})$ satisfies $\mathsf{BE}_w(\mathsf{c}, \infty)$. If $f \in \mathrm{Lip}_\star(X)$, then*

$$\tilde{\mathsf{P}}_t Q_1 f(y) - \tilde{\mathsf{P}}_t f(x) \leq \frac{1}{2}\, \mathsf{c}^2(t)\, \mathsf{d}^2(y, x) \tag{3.38}$$

for all $x, y \in X$ and $t \geq 0$.

In the proof of Lemma 3.17, we will use the following generalization of Fatou Lemma. For its proof, we refer the reader to [16, Lemma 3.3].

Lemma 3.18 (Generalized Fatou Lemma)**.** *If $\mu_n \in \mathscr{P}(X)$ weakly converges to $\mu \in \mathscr{P}(X)$ and $(f_n)_{n\in\mathbb{N}}$ are Borel equi-bounded functions such that*

$$\limsup_{n\to+\infty} f_n(x_n) \leq f(x) \quad \text{whenever } x_n \to x \text{ as } n \to +\infty$$

for some Borel function f, then

$$\limsup_{n\to+\infty} \int_X f_n \, \mathrm{d}\mu_n \leq \int_X f \, \mathrm{d}\mu.$$

Proof of Lemma 3.17. Let $t \geq 0$ and $x, y \in X$ be fixed. Since $Q_s f \in \mathrm{Lip}_\star(X)$, we clearly have $Q_s f \in \mathrm{Lip}_b(X) \cap \mathrm{L}^2(X, \mathfrak{m})$ with $|\mathrm{D}^* Q_s f| \in \mathrm{L}^2(X, \mathfrak{m})$ for all $s \geq 0$. By Proposition 3.13(i), Proposition 3.9 and Proposition 3.14, we thus have that

$$\mathrm{Lip}(\tilde{\mathsf{P}}_t Q_s f) \leq \mathsf{c}(t)\, \mathrm{Lip}(Q_s f) \tag{3.39}$$

for all $s \geq 0$ and

$$|\mathrm{D}^* \tilde{\mathsf{P}}_t Q_s f|^2(x) \leq \mathsf{c}^2(t)\, \tilde{\mathsf{P}}_t(|\mathrm{D}^* Q_s f|^2)(x) \tag{3.40}$$

for all $s \geq 0$ and $x \in X$. Now let $\gamma \in \mathrm{AC}([0,1]; X)$ be such that $\gamma_0 = x$ and $\gamma_1 = y$. We claim that $s \mapsto \tilde{\mathsf{P}}_t Q_s f(\gamma_s) \in \mathrm{AC}([0,1]; \mathbb{R})$. Indeed, by (2.6) and (3.39), we can estimate

$$
\begin{aligned}
|\tilde{\mathsf{P}}_t Q_{s_1} f(\gamma_{s_1}) - \tilde{\mathsf{P}}_t Q_{s_0} f(\gamma_{s_0})| &\leq |\tilde{\mathsf{P}}_t Q_{s_1} f(\gamma_{s_1}) - \tilde{\mathsf{P}}_t Q_{s_1} f(\gamma_{s_0})| + |\tilde{\mathsf{P}}_t Q_{s_1} f(\gamma_{s_0}) - \tilde{\mathsf{P}}_t Q_{s_0} f(\gamma_{s_0})| \\
&\leq \mathrm{Lip}(\tilde{\mathsf{P}}_t Q_{s_1} f)\, \mathsf{d}(\gamma_{s_1}, \gamma_{s_0}) + \int_X |Q_{s_1} f - Q_{s_0} f|\, \mathrm{d}\mathsf{H}_t \gamma_{s_0} \\
&\leq 2\mathsf{c}(t)\, \mathrm{Lip}(f) \int_{s_0}^{s_1} |\dot{\gamma}_s|\, \mathrm{d}s + 2\, \mathrm{Lip}(f)^2\, (s_1 - s_0)
\end{aligned}
$$

for all $0 \leq s_0 < s_1 \leq 1$. We can now write

$$
\frac{\tilde{\mathsf{P}}_t Q_{s+h} f(\gamma_{s+h}) - \tilde{\mathsf{P}}_t Q_s f(\gamma_s)}{h} = \int_X \frac{Q_{s+h} f - Q_s f}{h}\, \mathrm{d}\mathsf{H}_t \gamma_{s+h} + \frac{\tilde{\mathsf{P}}_t Q_s f(\gamma_{s+h}) - \tilde{\mathsf{P}}_t Q_s f(\gamma_s)}{h}
$$

for all $0 \leq s < s + h \leq 1$. On the one hand, we have

$$
\begin{aligned}
\limsup_{h \to 0^+} \int_X \frac{Q_{s+h} f - Q_s f}{h}\, \mathrm{d}\mathsf{H}_t \gamma_{s+h} &\leq \int_X \frac{\mathrm{d}^+}{\mathrm{d}s} Q_s f\, \mathrm{d}\mathsf{H}_t \gamma_s \\
&= -\frac{1}{2} \int_X |\mathrm{D}Q_s f|^2\, \mathrm{d}\mathsf{H}_t \gamma_s \\
&= -\frac{1}{2} \tilde{\mathsf{P}}_t (|\mathrm{D}Q_s f|^2)(\gamma_s)
\end{aligned}
$$

by Lemma 3.18 and (2.5) for all $s \in [0,1]$. On the other hand, by the upper gradient property of the asymptotic Lipschitz constant for Lipschitz functions, we can estimate

$$
\begin{aligned}
\limsup_{h \to 0^+} \frac{|\tilde{\mathsf{P}}_t Q_s f(\gamma_{s+h}) - \tilde{\mathsf{P}}_t Q_s f(\gamma_s)|}{h} &\leq \limsup_{h \to 0^+} \frac{1}{h} \int_s^{s+h} |\mathrm{D}^* \tilde{\mathsf{P}}_t Q_s f|(\gamma_r)\, |\dot{\gamma}_r|\, \mathrm{d}r \\
&= |\mathrm{D}^* \tilde{\mathsf{P}}_t Q_s f|(\gamma_s)\, |\dot{\gamma}_s|
\end{aligned}
$$

for $\mathscr{L}^1$-a.e. $s \in [0,1]$. By (3.40) and Young inequality, we have

$$
\begin{aligned}
|\mathrm{D}^* \tilde{\mathsf{P}}_t Q_s f|(\gamma_s)\, |\dot{\gamma}_s| &\leq \mathsf{c}(t)\, |\dot{\gamma}_s|\, \sqrt{\tilde{\mathsf{P}}_t (|\mathrm{D}^* Q_s f|^2)(\gamma_s)} \\
&\leq \frac{1}{2} \mathsf{c}^2(t)\, |\dot{\gamma}_s|^2 + \frac{1}{2} \tilde{\mathsf{P}}_t (|\mathrm{D}^* Q_s f|^2)(\gamma_s)
\end{aligned}
$$

for all $s \in [0,1]$. By (2.7), we thus have that

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s} \tilde{\mathsf{P}}_t Q_s f(\gamma_s) &\leq -\frac{1}{2} \tilde{\mathsf{P}}_t (|\mathrm{D}Q_s f|^2)(\gamma_s) + \frac{1}{2} \mathsf{c}^2(t)\, |\dot{\gamma}_s|^2 + \frac{1}{2} \tilde{\mathsf{P}}_t (|\mathrm{D}^* Q_s f|^2)(\gamma_s) \\
&= \frac{1}{2} \mathsf{c}^2(t)\, |\dot{\gamma}_s|^2
\end{aligned}
$$

for $\mathscr{L}^1$-a.e. $s \in [0,1]$, so that

$$
\tilde{\mathsf{P}}_t Q_1 f(y) - \tilde{\mathsf{P}}_t f(x) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s} \tilde{\mathsf{P}}_t Q_s f(\gamma_s)\, \mathrm{d}s \leq \frac{1}{2} \mathsf{c}^2(t) \int_0^1 |\dot{\gamma}_s|^2\, \mathrm{d}s
$$

and (3.38) follows by minimizing with respect to all curves $\gamma \in \mathrm{AC}([0,1]; X)$ such that $\gamma_0 = x$ and $\gamma_1 = y$. This concludes the proof. $\qquad\square$

The second preliminary result is a well known result proved for the first time by Lisini in [108]. In this general framework, this result was proved in [16, Lemma 4.12] (se also [83, Theorem 2.1]). For the reader's convenience, we provide a proof of it below.

Lemma 3.19 (Lisini Theorem for Lip-functions). *If $s \mapsto \mu_s \in \mathrm{AC}^2([0,1], \mathscr{P}_2(X))$, then $s \mapsto \int_X \varphi \, \mathrm{d}\mu_s \in \mathrm{AC}^2([0,1]; \mathbb{R})$ with*

$$\left| \int_X \varphi \, \mathrm{d}\mu_1 - \int_X \varphi \, \mathrm{d}\mu_0 \right| \leq \int_0^1 \left(\int_X |\mathrm{D}\varphi|^2 \, \mathrm{d}\mu_s \right)^{1/2} |\dot{\mu}_s| \, \mathrm{d}s \tag{3.41}$$

for all $\varphi \in \mathrm{Lip}_b(X)$. In particular, we have

$$\left| \frac{\mathrm{d}}{\mathrm{d}s} \int_X \varphi \, \mathrm{d}\mu_s \right|^2 \leq |\dot{\mu}_s|^2 \int_X |\mathrm{D}\varphi|^2 \, \mathrm{d}\mu_s \quad \textit{for } \mathscr{L}^1\textit{-a.e. } s \in [0,1] \tag{3.42}$$

for all $\varphi \in \mathrm{Lip}_b(X)$.

Proof. Let $\varphi \in \mathrm{Lip}_b(X)$ be fixed. By Lisini Theorem, see [108, Theorem 5] or [83, Theorem 2.1], there exists $\eta \in \mathscr{P}(\mathscr{C})$, $\mathscr{C} = \mathrm{C}([0,1], (X, \mathsf{d}))$, concentrated on $\mathrm{AC}([0,1], X)$, such that

$$(\mathsf{e}_s)_\sharp \eta = \mu_s \quad \text{for all } s \in [0,1], \tag{3.43}$$

where $\mathsf{e}_s \colon \mathscr{C} \to X$ is the evaluation map at time $s \in [0,1]$, and

$$\int_\mathscr{C} |\dot{\gamma}_s|^2 \, \mathrm{d}\eta(\gamma) = |\dot{\mu}_s|^2 \quad \text{for } \mathscr{L}^1\text{-a.e. } s \in [0,1]. \tag{3.44}$$

By the upper gradient property of the slope (recall (2.3)), by (3.43), (3.44) and Hölder inequality, we can estimate

$$\left| \int_X \varphi \, \mathrm{d}\mu_1 - \int_X \varphi \, \mathrm{d}\mu_0 \right| = \left| \int_\mathscr{C} (\varphi(\gamma_1) - \varphi(\gamma_0)) \, \mathrm{d}\eta(\gamma) \right|$$

$$\leq \int_0^1 \int_\mathscr{C} |\mathrm{D}\varphi|(\gamma_s) \, |\dot{\gamma}_s| \, \mathrm{d}\eta(\gamma) \, \mathrm{d}s$$

$$\leq \int_0^1 \left(\int_\mathscr{C} |\mathrm{D}\varphi|^2(\gamma_s) \, \mathrm{d}\eta(\gamma) \right)^{1/2} \left(\int_\mathscr{C} |\dot{\gamma}_s|^2 \, \mathrm{d}\eta(\gamma) \right)^{1/2} \mathrm{d}s$$

$$= \int_0^1 \left(\int_X |\mathrm{D}\varphi|^2 \, \mathrm{d}\mu_s \right)^{1/2} |\dot{\mu}_s| \, \mathrm{d}s,$$

proving (3.41). Inequality (3.42) follows easily. $\qquad\square$

We are now ready to prove the main result of this section.

Proof of Theorem 3.16. We prove the two implications separately.

Proof of $(i) \Rightarrow (ii)$. Fix $t \geq 0$. We divide the proof in three steps.

Step 1: definition of H_t and $\tilde{\mathsf{P}}_t$. We define $\mathsf{H}_t \colon \mathscr{P}^{\mathrm{ac}}(X) \to \mathscr{P}^{\mathrm{ac}}(X)$ by setting

$$\mathsf{H}_t \mu = (\mathsf{P}_t f) \, \mathfrak{m} \quad \text{for all } \mu = f\mathfrak{m} \in \mathscr{P}^{\mathrm{ac}}(X), \tag{3.45}$$

as in (3.21). By Proposition 3.12, this map can be extended to a map $\mathsf{H}_t \colon \mathscr{P}_1(X) \to \mathscr{P}_1(X)$ which satisfies (3.25) with $C(t) = \mathsf{c}(t)$. Hence (3.27) holds with $p = 1$ and we can thus define

$$\tilde{\mathsf{P}}_t f(x) = \int_X f \, \mathrm{d}\mathsf{H}_t \delta_x, \quad x \in X,$$

whenever $f \colon X \to \overline{\mathbb{R}}$ is either a bounded or a non-negative Borel function, as in (3.28).

Step 2: W_2-estimate for H_t on Dirac deltas. Let $x, y \in X$. By Lemma 3.17, we can estimate

$$\int_X Q_1 \varphi \, \mathrm{d}\mathsf{H}_t \delta_y - \int_X \varphi \, \mathrm{d}\mathsf{H}_t \delta_x = \tilde{\mathsf{P}}_t Q_1 \varphi(y) - \tilde{\mathsf{P}}_t \varphi(x) \leq \frac{1}{2} \mathsf{c}^2(t) \, \mathsf{d}^2(y, x) \tag{3.46}$$

for all $\varphi \in \mathrm{Lip}_\star(X)$. Hence, by (2.13) and taking the supremum on all $\varphi \in \mathrm{Lip}_\star(X)$ in (3.46), we get

$$W_2(\mathsf{H}_t\delta_y, \mathsf{H}_t\delta_x) \leq \mathsf{c}(t)\,\mathsf{d}(y,x) \tag{3.47}$$

whenever $x, y \in X$.

  Step 3: W_2 estimate for H_t on $\mathscr{P}_2(X)$. If $\mu \in \mathscr{P}_2(X)$, then we can write

$$\mu = \int_X \delta_x\,\mathrm{d}\mu(x) \tag{3.48}$$

and thus, by Proposition 3.13(iii), we we can also write

$$\mathsf{H}_t\mu = \int_X \mathsf{H}_t\delta_x\,\mathrm{d}\mu(x). \tag{3.49}$$

Now let $\mu, \nu \in \mathscr{P}_2(X)$. If $\pi \in \mathsf{Plan}(\mu,\nu)$, then we may use a Measurable Selection Theorem (see [141, Corollary 5.22] or [52, Theorem 6.9.2] for example) to select in a π-measurable way an optimal plan

$$\omega_{x,y}^t \in \mathsf{OptPlan}(\mathsf{H}_t\delta_x, \mathsf{H}_t\delta_y) \quad \text{for all } x, y \in X. \tag{3.50}$$

By (3.48) and (3.49), we thus get that

$$\Omega_t = \int_{X \times X} \omega_{x,y}^t\,\mathrm{d}\pi(x,y) \in \mathsf{Plan}(\mathsf{H}_t\mu, \mathsf{H}_t\nu).$$

Hence, by (2.8), the optimality of (3.50) and by (3.47), we can estimate

$$\begin{aligned}
W_2^2(\mathsf{H}_t\mu, \mathsf{H}_t\nu) &\leq \int_{X \times X} \mathsf{d}^2(u,v)\,\mathrm{d}\Omega_t(u,v) \\
&= \int_{X \times X}\int_{X \times X} \mathsf{d}^2(u,v)\,\mathrm{d}\omega_{x,y}^t(u,v)\,\mathrm{d}\pi(x,y) \\
&= \int_{X \times X} W_2^2(\mathsf{H}_t\delta_x, \mathsf{H}_t\delta_y)\,\mathrm{d}\pi(x,y) \\
&\leq \mathsf{c}^2(t) \int_{X \times X} \mathsf{d}^2(x,y)\,\mathrm{d}\pi(x,y)
\end{aligned}$$

whenever $\pi \in \mathsf{Plan}(\mu,\nu)$. Again by (2.8), we thus get (3.36). By (3.45) and Lemma 3.10, this proves (ii) with $\mathscr{D} = \mathrm{Dom}(\mathsf{Ent}_\mathfrak{m})$.

  Proof of $(ii) \Rightarrow (i)$. Fix $t \geq 0$. We divide the proof in three steps.

  Step 1: definition of H_t and $\tilde{\mathsf{P}}_t$. We define $\mathsf{H}_t \colon \mathscr{D} \to \mathscr{P}_2^{\mathrm{ac}}(X)$ by setting

$$\mathsf{H}_t(f\mathfrak{m}) = (\mathsf{P}_t f)\,\mathfrak{m} \quad \text{for all } f\mathfrak{m} \in \mathscr{D} \tag{3.51}$$

as in (3.21). Thanks to Lemma 3.11, by (3.35) we can extend the map (3.51) to a map $\mathsf{H}_t \colon \mathscr{P}_2(X) \to \mathscr{P}_2(X)$ (for which we retain the same notation) such that

$$W_2(\mathsf{H}_t\mu, \mathsf{H}_t\nu) \leq \mathsf{c}(t)\,W_2(\mu,\nu) \quad \text{for all } \mu, \nu \in \mathscr{P}_2(X).$$

Hence (3.27) holds for $p = 2$ and we can thus define

$$\tilde{\mathsf{P}}_t f(x) = \int_X f\,\mathrm{d}\mathsf{H}_t\delta_x, \quad x \in X,$$

whenever $f \colon X \to \overline{\mathbb{R}}$ is either a bounded or a non-negative Borel function, as in (3.28).

  Step 2: BE_w on Lip-functions via $\tilde{\mathsf{P}}_t$. Let $f \in \mathrm{Lip}_b(X) \cap \mathrm{L}^2(X,\mathfrak{m})$ with $|\mathrm{D}^* f| \in \mathrm{L}^2(X,\mathfrak{m})$. We claim that

$$\Gamma(\mathsf{P}_t f) \leq \mathsf{c}^2(t)\,\mathsf{P}_t|\mathrm{D}^* f|^2 \quad \mathfrak{m}\text{-a.e. in } X. \tag{3.52}$$

Indeed, thanks to Proposition 3.13(ii), we have $\tilde{\mathsf{P}}_t f = \mathsf{P}_t f$ $\mathfrak{m}$-a.e. in X, so that $\tilde{\mathsf{P}}_t f \in L^2(X, \mathfrak{m}) \cap L^\infty(X, \mathfrak{m})$ in particular. Fix $x, y \in X$ and let $\gamma \in AC^2([0, 1]; X)$ such that $\gamma_0 = x$ and $\gamma_1 = y$. We thus have that

$$s \mapsto \mu_s = \mathsf{H}_t \delta_{\gamma_s} \in AC^2([0, 1]; \mathscr{P}_2(X)),$$

since we can estimate

$$W_2(\mu_{s_1}, \mu_{s_0}) \le \mathsf{c}(t)\, W_2(\delta_{\gamma_{s_0}}, \delta_{\gamma_{s_1}}) \le \mathsf{c}(t)\, \mathsf{d}(\gamma_{s_1}, \gamma_{s_0}) \le \mathsf{c}(t) \int_{s_0}^{s_1} |\dot{\gamma}_r|\, \mathrm{d}r$$

for all $0 \le s_0 < s_1 \le 1$, which immediately gives

$$|\dot{\mu}_s| \le \mathsf{c}(t)\, |\dot{\gamma}_s| \quad \text{for } \mathscr{L}^1\text{-a.e. } s \in [0, 1]. \tag{3.53}$$

By Lemma 3.19 and (3.53), we thus get

$$\begin{aligned}
|\tilde{\mathsf{P}}_t f(y) - \tilde{\mathsf{P}}_t f(x)| &= \left| \int_X f\, \mathrm{d}\mu_1 - \int_X f\, \mathrm{d}\mu_0 \right| \\
&\le \int_0^1 \left(\int_X |\mathrm{D}f|^2\, \mathrm{d}\mu_s \right)^{1/2} |\dot{\mu}_s|\, \mathrm{d}s \\
&\le \mathsf{c}(t) \int_0^1 \left(\int_X |\mathrm{D}f|^2\, \mathrm{d}\mu_s \right)^{1/2} |\dot{\gamma}_s|\, \mathrm{d}s \\
&= \mathsf{c}(t) \int_0^1 \left(\tilde{\mathsf{P}}_t |\mathrm{D}f|^2(\gamma_s) \right)^{1/2} |\dot{\gamma}_s|\, \mathrm{d}s.
\end{aligned}$$

Thanks to (3.29), we thus get

$$|\tilde{\mathsf{P}}_t f(y) - \tilde{\mathsf{P}}_t f(x)| \le \mathsf{c}(t)\, \mathrm{Lip}(f) \int_0^1 |\dot{\gamma}_s|\, \mathrm{d}s,$$

so that $\tilde{\mathsf{P}}_t f \in \mathrm{Lip}(X)$ with $\mathrm{Lip}(\tilde{\mathsf{P}}_t f) \le \mathsf{c}(t)\, \mathrm{Lip}(f)$ by the length property of (X, d) (recall Proposition 2.2). In addition, again by the length property of (X, d), we have

$$\left| \tilde{\mathsf{P}}_t f(y) - \tilde{\mathsf{P}}_t f(x) \right| \le \mathsf{c}(t)\, \mathsf{d}(y, x)\, \sup\left\{ \left(\tilde{\mathsf{P}}_t |\mathrm{D}^* f|^2(z) \right)^{1/2} : \mathsf{d}(z, y) \le 2\mathsf{d}(y, x) \right\} \tag{3.54}$$

for all $x, y \in X$. Since $x \mapsto |\mathrm{D}^* f|(x)$ is upper semicontinuous and bounded, by Proposition 3.13(i) also $x \mapsto \left(\tilde{\mathsf{P}}_t |\mathrm{D}^* f|^2(x) \right)^{1/2}$ is upper semicontinuous. Therefore, taking the $\limsup$ as $y \to x$ in (3.54), we get

$$|\mathrm{D}^* \tilde{\mathsf{P}}_t f|(x) \le \mathsf{c}(t) \left(\tilde{\mathsf{P}}_t |\mathrm{D}^* f|^2(x) \right)^{1/2} \quad \text{for all } x \in X.$$

Since $|\mathrm{D}^* f| \in L^2(X, \mathfrak{m}) \cap L^\infty(X, \mathfrak{m})$ is Borel, we have that $\tilde{\mathsf{P}}_t |\mathrm{D}^* f|^2 = \mathsf{P}_t |\mathrm{D}^* f|^2$ $\mathfrak{m}$-a.e. in X by Proposition 3.13(i), and thus $\tilde{\mathsf{P}}_t |\mathrm{D}^* f|^2 \in L^2(X, \mathfrak{m})$. Hence $\tilde{\mathsf{P}}_t f \in \mathrm{Lip}_b(X) \cap L^2(X, \mathfrak{m})$ with $|\mathrm{D}^* \tilde{\mathsf{P}}_t f| \in L^2(X, \mathfrak{m})$, so that $\tilde{\mathsf{P}}_t f \in W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ by (2.22), with $\Gamma(\tilde{\mathsf{P}}_t f) \le |\mathrm{D}^* \tilde{\mathsf{P}}_t f|^2$ $\mathfrak{m}$-a.e. in X by (2.29). Since $\tilde{\mathsf{P}}_t f = \mathsf{P}_t f$ $\mathfrak{m}$-a.e. in X, we must have $\Gamma(\tilde{\mathsf{P}}_t f) = \Gamma(\mathsf{P}_t f)$ (recall (2.21) and again the definition of minimal weak gradient). Claim (3.52) is thus proved.

Step 3: approximation. Let $f \in W^{1,2}(X, \mathsf{d}, \mathfrak{m})$. By (2.30), we can find $f_n \in \mathrm{Lip}_b(X) \cap L^2(X, \mathfrak{m})$ such that $f_n \to f$ and $|\mathrm{D}^* f| \to |\mathrm{D}f|_w$ in $L^2(X, \mathfrak{m})$ as $n \to +\infty$. By Step 2, we have

$$\Gamma(\mathsf{P}_t f_n) \le \mathsf{c}^2(t)\, \mathsf{P}_t |\mathrm{D}^* f_n|^2 \quad \mathfrak{m}\text{-a.e. in } X \tag{3.55}$$

for all $n \in \mathbb{N}$. Since $\mathsf{P}_t |\mathrm{D}^* f_n|^2 \to \mathsf{P}_t \Gamma(\mathsf{P}_t f)$ in $\mathrm{L}^1(X, \mathfrak{m})$ as $n \to +\infty$ by (2.33) and $\Gamma(\mathsf{P}_t f_n) \to \Gamma(\mathsf{P}_t f)$ by (2.40) and again (2.33), up to possibly pass to a subsequence, we can pass to the limit as $n \to +\infty$ in (3.55) and get (3.9). This proves (i). $\qquad \square$

Remark 3.20 (Errata to the proof of [16, Theorem 3.5]). In [16, Section 3.2], instead of (3.33), the authors consider the pointwise inequality (see [16, Equation (3.16)])

$$|\mathrm{D}\mathsf{P}_t f|^2(x) \le \mathsf{c}^2(t)\,\tilde{\mathsf{P}}_t(|\mathrm{D}f|^2)(x) \quad \text{for all } x \in X, \tag{3.56}$$

whenever $f \in \mathrm{Lip}_b(X) \cap \mathrm{L}^2(X, \mathfrak{m})$. In the proof of [16, Theorem 3.5], the authors then state that inequality (3.56), together with inequality (3.18) (which corresponds to [16, Equation (3.15)]) are implied by the W_2-contractivity property of the dual heat semigroup. Since they do not use this implication in their paper, the authors do not provide a proof of this statement and only refer to [15, Theorem 6.2] and to [97]. However, the proof of [15, Theorem 6.2] uses the fact that $\tilde{\mathsf{P}}_t|\mathrm{D}f|^2 \in \mathrm{C}(X)$ for all $t > 0$, i.e. the L^∞-*to*-C-*regularization property* of $(\tilde{\mathsf{P}}_t)_{t>0}$ previously proved in [15, Theorem 6.1] thanks to the EVI_K property of the gradient flow of the entropy. In the general framework considered in [16, Section 3.2], as well as in the present one, the L^∞-to-C-regularization property of $(\tilde{\mathsf{P}}_t)_{t>0}$ is not available, and thus the continuity of the function $x \mapsto \tilde{\mathsf{P}}_t(|\mathrm{D}f|^2)(x)$ for $t > 0$ is not guaranteed. For this reason, the implication

$$W_2\text{-contractivity of } (\mathsf{H}_t)_{t \ge 0} \implies (3.18) \text{ and } (3.56)$$

stated in [16, Theorem 3.5] is not completely justified. One can get rid of this problem by replacing [16, Equation (3.16)] with (3.33) and arguing as we have done in the proof of the implication (ii) $\Rightarrow$ (i) of Theorem 3.16 above thanks to the upper semicontinuity of the asymptotic Lipschitz constant (recall its definition in (2.2)), without affecting the validity of all the other results of [16, Section 3.2]. We let the interested reader check the details.

3.10. **Strong Feller property and densities of the dual heat semigroup.** The following result deals with the regularization property of the pointwise heat semigroup $(\tilde{\mathsf{P}}_t)_{t>0}$ on $\mathrm{L}^2 \cap \mathrm{L}^\infty$-functions, see the proof of the implication (i) $\Rightarrow$ (v) in [16, Theorem 3.17]. We briefly provide its proof below for the reader's convenience.

Corollary 3.21 (Strong Feller property). *Assume $(X, \mathsf{d}, \mathfrak{m})$ satisfies $\mathsf{BE}_w(\mathsf{c}, \infty)$. If $f \in \mathrm{L}^2(X, \mathfrak{m}) \cap \mathrm{L}^\infty(X, \mathfrak{m})$ is Borel, then $\tilde{\mathsf{P}}_t f \in \mathrm{Lip}_b(X)$ with*

$$\sqrt{2\,\mathrm{I}_{-2}(t)}\,\mathrm{Lip}(\tilde{\mathsf{P}}_t f) \le \|f\|_{\mathrm{L}^\infty(X, \mathfrak{m})} \tag{3.57}$$

for all $t > 0$.

Proof. Let $t > 0$ be fixed. Assume $f \in \mathrm{C}_b(X) \cap \mathrm{L}^2(X, \mathfrak{m})$. Then $\mathsf{P}_t f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ by Proposition 3.8(i) with $2\,\mathrm{I}_{-2}(t)\,\Gamma(\mathsf{P}_t f) \le \|f\|_{\mathrm{L}^\infty(X, \mathfrak{m})}^2$ $\mathfrak{m}$-a.e. in X. Hence $\mathsf{P}_t f$ has a Lipschitz representative by (P.5). Thanks to Theorem 3.16 and Proposition 3.13(i), the Lipschitz representative of $\mathsf{P}_t f$ must coincide with $\tilde{\mathsf{P}}_t f$. We thus get that $\tilde{\mathsf{P}}_t f \in \mathrm{Lip}_b(X)$ satisfies (3.57). Hence, arguing as in Step 2 of the proof of Proposition 3.13(iii), we get that $\tilde{\mathsf{P}}_t \chi_K \in \mathrm{Lip}_b(X)$ satisfies (3.57) whenever $K \subset X$ is a non-empty bounded closed set. The conclusion thus follows by the Monotone Class Theorem (see [70, Theorem 5.2.2] for example), since (3.57) allows to convert monotone equibounded convergence of a sequence $(f_n)_{n \in \mathbb{N}}$ into pointwise convergence on X of the sequence $(\tilde{\mathsf{P}}_t f_n)_{n \in \mathbb{N}}$. $\qquad \square$

An important consequence of Corollary 3.21 is the absolute continuity property of the dual heat semigroup $(\mathsf{H}_t)_{t>0}$ on measures in $\mathscr{P}_2(X)$, see the proof of [16, Theorem 3.17]. We provide a sketch of its proof below for the reader's convenience.

Corollary 3.22 ($\mathsf{H}_t(\mathscr{P}_2(X)) \subset \mathscr{P}_2^{\mathrm{ac}}(X)$ for $t > 0$). *Assume* $(X, \mathsf{d}, \mathfrak{m})$ *satisfies* $\mathsf{BE}_w(\mathsf{c}, \infty)$. *If* $\mu \in \mathscr{P}_2(X)$, *then* $\mathsf{H}_t\mu \ll \mathfrak{m}$ *for all* $t > 0$.

Proof. Let $t > 0$ be fixed. Let $\mu \in \mathscr{P}_2(X)$ and let $A \subset X$ be a Borel set with $\mathfrak{m}(A) = 0$. Then $\chi_A \in \mathrm{L}^2(X, \mathfrak{m}) \cap \mathrm{L}^\infty(X, \mathfrak{m})$ and so $\tilde{\mathsf{P}}_t\chi_A \in \mathrm{Lip}_b(X)$ by Corollary 3.21. By Proposition 3.13(ii), we must have that $\tilde{\mathsf{P}}_t f = \mathsf{P}_t f = 0$ $\mathfrak{m}$-a.e. in X, and thus $\tilde{\mathsf{P}}_t f(x) = 0$ for all $x \in X$. Hence

$$\mathsf{H}_t\mu(A) = \int_X \chi_A \,\mathrm{d}\mathsf{H}_t\mu = \int_X \tilde{\mathsf{P}}_t\chi_A \,\mathrm{d}\mu = 0$$

by Theorem 3.16 and Proposition 3.13(iii). The proof is complete. $\qquad\square$

Remark 3.23 (Extension of $(\mathsf{H}_t)_{t\geq 0}$ on $\mathscr{P}(X)$). Although we do not need such a generality for our purposes, it is possible to show that the dual heat semigroup can be extended to a weakly continuous map $(\mathsf{H}_t)_{t\geq 0} \colon \mathscr{P}(X) \to \mathscr{P}(X)$ such that (3.36) holds for all $\mu, \nu \in \mathscr{P}(X)$ with $W_2(\mu, \nu) < +\infty$. Moreover, the validity of Proposition 3.13(iii) and of Corollary 3.22 extends to any $\mu \in \mathscr{P}(X)$. We refer the interested reader to [16, Section 3.2] for the details.

By Corollary 3.22, for all $x \in X$ there exists a non-negative density $\mathsf{p}_t[x] \in \mathrm{L}^1(X, \mathfrak{m})$ such that

$$\mathsf{H}_t\delta_x = \mathsf{p}_t[x]\,\mathfrak{m} \quad \text{for all } t > 0. \tag{3.58}$$

Therefore, accordingly with (3.28), if $f \colon X \to \overline{\mathbb{R}}$ is either a bounded or a non-negative Borel function, we can then write

$$\tilde{\mathsf{P}}_t f(x) = \int_X f(y) \,\mathrm{d}\mathsf{H}_t\delta_x(y) = \int_X f(y) \,\mathsf{p}_t[x](y) \,\mathrm{d}\mathfrak{m}(y), \tag{3.59}$$

for all $t > 0$, so that the definition of $(\tilde{\mathsf{P}}_t f)_{t>0}$ does not depend on the particular choice of the representative of f. By linearity, $(\tilde{\mathsf{P}}_t f)_{t>0}$ is thus well defined whenever $f \colon X \to \overline{\mathbb{R}}$ is a one-side bounded measurable function.

Lemma 3.24 (Properties of $(\mathsf{p}_t[\,\cdot\,])_{t>0}$). *Assume* $(X, \mathsf{d}, \mathfrak{m})$ *satisfies* $\mathsf{BE}_w(\mathsf{c}, \infty)$ *and let* $t > 0$. *The following hold.*

    (i) $\tilde{\mathsf{P}}_s(\mathsf{p}_t[x]) = \mathsf{p}_{s+t}[x]$ $\mathfrak{m}$-a.e. in X for all $x \in X$ and $s \geq 0$.
    (ii) $\mathsf{p}_t[x](y) = \mathsf{p}_t[y](x)$ for $\mathfrak{m}$-a.e. $x, y \in X$.

Proof. Let $t > 0$ be fixed. We prove the two statements separately.

Proof of (i). Let $\varphi \in \mathrm{L}^1(X, \mathfrak{m}) \cap \mathrm{L}^\infty(X, \mathfrak{m})$ be a Borel non-negative function and set $\varphi = \bar{\varphi} \,\|\varphi\|_{\mathrm{L}^1(X,\mathfrak{m})}$. We can compute

$$
\begin{aligned}
\int_X \varphi \,\mathsf{p}_{s+t}[y] \,\mathrm{d}\mathfrak{m} &= \int_X \varphi \,\mathrm{d}\mathsf{H}_{s+t}\delta_y && \text{by (3.59), (3.22) and Lemma 3.11} \\
&= \int_X \tilde{\mathsf{P}}_s\varphi \,\mathrm{d}\mathsf{H}_t\delta_y && \text{by (3.28)} \\
&= \int_X \mathsf{P}_s\varphi \,\mathsf{p}_t[y] \,\mathrm{d}\mathfrak{m} && \text{by Proposition 3.13(ii) and (3.59)} \\
&= \|\varphi\|_{\mathrm{L}^1(X,\mathfrak{m})} \int_X \mathsf{p}_t[y] \,\mathrm{d}\mathsf{H}_s(\bar{\varphi}\,\mathfrak{m}) && \text{by (3.21)}
\end{aligned}
$$

$$= \int_X \tilde{\mathsf{P}}_s(\mathsf{p}_t[y])\,\varphi\,\mathrm{d}\mathfrak{m} \qquad\qquad \text{by Proposition 3.13(iii) and (3.59)}$$

for all $s \geq 0$, so that (i) immediately follows.

Proof of (ii). Let $\varphi, \psi \in \mathrm{L}^1(X, \mathfrak{m}) \cap \mathrm{L}^\infty(X, \mathfrak{m})$ be two Borel non-negative functions. By Tonelli Theorem, (3.59) and Proposition 3.13(ii), we can compute

$$\int_X \int_X \varphi(x)\,\psi(y)\,\mathsf{p}_t[x](y)\,\mathrm{d}\mathfrak{m}(x)\,\mathrm{d}\mathfrak{m}(y) = \int_X \varphi(x) \int_X \psi(y)\,\mathsf{p}_t[x](y)\,\mathrm{d}\mathfrak{m}(y)\,\mathrm{d}\mathfrak{m}(x)$$

$$= \int_X \varphi(x)\,\tilde{\mathsf{P}}_t\psi(x)\,\mathrm{d}\mathfrak{m}(x)$$

$$= \int_X \tilde{\mathsf{P}}_t\varphi(y)\,\psi(y)\,\mathrm{d}\mathfrak{m}(y)$$

$$= \int_X \psi(y) \int_X \varphi(x)\,\mathsf{p}_t[y](x)\,\mathrm{d}\mathfrak{m}(x)\,\mathrm{d}\mathfrak{m}(y)$$

$$= \int_X \int_X \varphi(x)\,\psi(y)\,\mathsf{p}_t[y](x)\,\mathrm{d}\mathfrak{m}(x)\,\mathrm{d}\mathfrak{m}(y),$$

so that (ii) immediately follows. $\square$

3.11. BE_w **inequality for Lipschitz functions, again.** We conclude this section with the following result, which provides a refined version of Proposition 3.9 and Proposition 3.14, see the proof of [16, Theorem 3.17]. We give its proof below for the reader's convenience.

Proposition 3.25 (BE_w for Lip-functions, III)**.** *Assume* $(X, \mathsf{d}, \mathfrak{m})$ *satisfies* $\mathsf{BE}_w(\mathsf{c}, \infty)$*. If* $f \in \mathrm{Lip}_b(X) \cap \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$*, then* $\mathsf{P}_t f \in \mathrm{Lip}_b(X) \cap \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ *with*

$$|\mathrm{D}^*\mathsf{P}_t f|^2(x) \leq \mathsf{c}^2(t)\,\tilde{\mathsf{P}}_t\Gamma(f)(x) \quad \text{for all } x \in X$$

whenever $t > 0$*.*

Proof. Let $t > 0$ be fixed. By Proposition 3.9 we already know that $\mathsf{P}_t f \in \mathrm{Lip}_b(X) \cap \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$. From (3.9), (2.29), Proposition 3.13(ii) and (3.59) we get that

$$\Gamma(\mathsf{P}_t f) \leq \mathsf{c}^2(t)\,\mathsf{P}_t\Gamma(f) = \mathsf{c}^2(t)\,\tilde{\mathsf{P}}_t\Gamma(f) \quad \mathfrak{m}\text{-a.e. in } X.$$

Since $\Gamma(f) \in \mathrm{L}^1(X, \mathfrak{m}) \cap \mathrm{L}^\infty(X, \mathfrak{m})$, by Corollary 3.21 and again (3.59) we must have that $\tilde{\mathsf{P}}_t\Gamma(f) \in \mathrm{C}_b(X)$. The conclusion thus follows from Lemma 3.15. $\square$

Corollary 3.26 (Weak reverse slope estimate for $(\mathsf{P}_t)_{t>0}$)**.** *Assume* $(X, \mathsf{d}, \mathfrak{m})$ *satisfies* $\mathsf{BE}_w(\mathsf{c}, \infty)$*. If* $f \in \mathrm{L}^2(X, \mathfrak{m}) \cap \mathrm{L}^\infty(X, \mathfrak{m})$*, then*

$$|\mathrm{D}^*\mathsf{P}_t f|^2 \leq \underline{\mathsf{c}}(0^+)^2\,\Gamma(\mathsf{P}_t f) \quad \mathfrak{m}\text{-a.e. in } X \tag{3.60}$$

for all $t > 0$*, where* $\underline{\mathsf{c}}(0^+) = \liminf\limits_{s \to 0^+} \mathsf{c}(s) \geq 1$*.*

Proof. Let $t > 0$ be fixed and let $(\varepsilon_n)_{n \in \mathbb{N}} \subset (0, t)$ be such that $\varepsilon_n \downarrow 0$ as $n \to +\infty$ and $\underline{\mathsf{c}}(0^+) = \lim\limits_{n \to +\infty} \mathsf{c}(\varepsilon_n)$. By Corollary 3.21, we have $f_n = \mathsf{P}_{t-\varepsilon_n} f \in \mathrm{Lip}_b(X) \cap \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ for all $n \in \mathbb{N}$ and thus, by Proposition 3.25, Proposition 3.13(ii) and (3.59), we can estimate

$$|\mathrm{D}^*\mathsf{P}_t f|^2 = |\mathrm{D}^*\mathsf{P}_{\varepsilon_n} f_n|^2 \leq \mathsf{c}^2(\varepsilon_n)\,\tilde{\mathsf{P}}_{\varepsilon_n}\Gamma(f_n) = \mathsf{c}^2(\varepsilon_n)\,\mathsf{P}_{\varepsilon_n}\Gamma(f_n) \quad \mathfrak{m}\text{-a.e. in } X$$

for all $n \in \mathbb{N}$. Since $\Gamma(f_n) \to \Gamma(f)$ in $\mathrm{L}^1(X, \mathfrak{m})$ as $n \to +\infty$ by (2.41) in Lemma 2.1, the conclusion follows by passing to the limit as $n \to +\infty$. $\square$

As a completely natural (although painful) drawback of the weakness of the $\mathsf{BE}_w(\mathsf{c}, \infty)$ property, if the function c in (3.8) is such that

$$\underline{\mathsf{c}}(0^+) = \liminf_{t \to 0^+} \mathsf{c}(t) > 1, \tag{3.61}$$

then Corollary 3.26 provides no useful information, since $|\mathsf{D}^*\mathsf{P}_t f| \geq |\mathsf{DP}_t f| \geq |\mathsf{DP}_t f|_w$ $\mathfrak{m}$-a.e. in X whenever $t > 0$ by (2.29) and Corollary 3.21. In Section 5 (precisely, in the proof of Lemma 5.3), similarly to [16], we will need the following regularization property of the heat semigroup.

Definition 3.27 (Heat-smoothing admissible space)**.** We say that an admissible metric-measure space is *heat-smoothing* if

$$f \in \mathrm{L}^\infty(X, \mathfrak{m}) \cap \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}) \implies |\mathsf{DP}_t f| = |\mathsf{DP}_t f|_w \ \mathfrak{m}\text{-a.e. in } X \text{ for all } t > 0. \tag{3.62}$$

Note that, if $\underline{\mathsf{c}}(0^+) = 1$, then inequality (3.60) in Corollary 3.26 immediately implies (3.62) (actually, in the stronger form assuming $f \in \mathrm{L}^2(X, \mathfrak{m}) \cap \mathrm{L}^\infty(X, \mathfrak{m})$ only).

4. Fisher information and $\mathrm{L} \log \mathrm{L}$-regolarization

In this section, we recall some useful properties of the Fisher information and the entropy functional in admissible metric-measure spaces. We only detail the proofs of the results which rely on the $\mathsf{BE}_w(\mathsf{c}, \infty)$ condition.

4.1. **Fisher information, entropy and Kuwada Lemma.** Let us set

$$\mathrm{L}^1_+(X, \mathfrak{m}) = \left\{ f \in \mathrm{L}^1(X, \mathfrak{m}) : f \geq 0 \ \mathfrak{m}\text{-a.e. in } X \right\}$$

the convex cone of non-negative L^1-functions. As in [14, Definition 4.9], the *Fisher information* $\mathsf{F} \colon \mathrm{L}^1_+(X, \mathfrak{m}) \to [0, +\infty]$ is defined for all $f \in \mathrm{L}^1_+(X, \mathfrak{m})$ as

$$\mathsf{F}(f) = \begin{cases} 4\mathsf{Ch}(\sqrt{f}) & \text{if } \sqrt{f} \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}) \\ +\infty & \text{otherwise.} \end{cases}$$

In particular, we have

$$\mathrm{Dom}(\mathsf{F}) = \left\{ f \in \mathrm{L}^1_+(X, \mathfrak{m}) : \sqrt{f} \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}) \right\}.$$

Since $f_n \to f$ in $\mathrm{L}^1_+(X, \mathfrak{m})$ as $n \to +\infty$ implies that $\sqrt{f_n} \to \sqrt{f}$ in $\mathrm{L}^2(X, \mathfrak{m})$ as $n \to +\infty$, the Fisher information F is lower semicontinuous in $\mathrm{L}^1_+(X, \mathfrak{m})$. Thanks to the locality property (2.23) and the chain rule (2.25), if $f \in \mathrm{Dom}(\mathsf{F})$ then $f_n = f \wedge n \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ with

$$|\mathsf{D}f_n|_w = 2\sqrt{f_n}\,|\mathsf{D}\sqrt{f}|_w\,\chi_{\{f \leq n\}} \in \mathrm{L}^1(X, \mathfrak{m})$$

for all $n \in \mathbb{N}$ and $|\mathsf{D}f_n|_w \uparrow 2\sqrt{f}\,|\mathsf{D}\sqrt{f}|_w$ in $\mathrm{L}^1(X, \mathfrak{m})$ as $n \to +\infty$. Hence we can write

$$\mathsf{F}(f) = \lim_{n \to +\infty} 4 \int_X |\mathsf{D}\sqrt{f}|_w^2\,\chi_{\{0 < f \leq n\}}\,\mathrm{d}\mathfrak{m}$$

$$= \lim_{n \to +\infty} \int_X \frac{4 f_n\,|\mathsf{D}\sqrt{f}|_w^2}{f}\,\chi_{\{0 < f \leq n\}}\,\mathrm{d}\mathfrak{m}$$

$$= \lim_{n \to +\infty} \int_{\{f > 0\}} \frac{|\mathsf{D}f_n|^2}{f}\,\mathrm{d}\mathfrak{m}.$$

Thus, accordingly with [14, Lemma 4.10], if $f \in \mathrm{Dom}(\mathsf{F})$ then we define

$$|\tilde{\mathrm{D}}f|_w = 2\sqrt{f}\,|\mathrm{D}\sqrt{f}|_w = \lim_{n\to+\infty} |\mathrm{D}f_n|_w \qquad (4.1)$$

and

$$\mathsf{F}(f) = \int_{\{f>0\}} \frac{\tilde{\Gamma}(f)}{f}\,\mathrm{d}\mathfrak{m}, \ \text{ where } \tilde{\Gamma}(f) = |\tilde{\mathrm{D}}f|_w^2. \qquad (4.2)$$

In particular, thanks to the convexity of the weak gradient and the convexity of the function $(x,y) \mapsto y^2/x$ on $(0,+\infty)\times\mathbb{R}$, we also get that the Fisher information is convex on $\mathrm{L}_+^1(X,\mathfrak{m})$, see [14, Lemma 4.10].

The following result is a part of the statement of [16, Lemma 4.2] and provides some simple but extremely useful estimates involving the Fisher information, the entropy functional and the second moments of the measure. The proof goes as the one of [16, Lemma 4.2] (see also [14, Theorems 4.16(b) and 4.20]) and we thus omit it.

Lemma 4.1 (Entropy and Fisher information along $(\mathsf{H}_t)_{t\geq0}$). *Let $\mu = f\mathfrak{m} \in \mathrm{Dom}(\mathsf{Ent}_\mathfrak{m})$ and set $f_t = \mathsf{P}_t f$ and $\mu_t = f_t\mathfrak{m}$ for all $t \geq 0$.*

(i) For all $t \geq 0$, we have $\mathsf{Ent}_\mathfrak{m}(\mu_t) \leq \mathsf{Ent}_\mathfrak{m}(\mu)$.
(ii) If $x_0 \in X$ and $T > 0$, then

$$\int_0^T \mathsf{F}(f_t)\,\mathrm{d}t + 2\int_0^T \int_X \mathsf{d}^2(x,x_0)\,\mathrm{d}\mu_t(x)\,\mathrm{d}t \leq 2e^{4T}\left(\mathsf{Ent}_\mathfrak{m}(\mu) + 2\int_X \mathsf{d}^2(x,x_0)\,\mathrm{d}\mu(x)\right). \quad (4.3)$$

Thanks to Lemma 4.1, we have the following fundamental result. We refer the interested reader to [14, Lemma 6.1] (see also [84, Proposition 3.7]) for its proof.

Lemma 4.2 (Estimate on the W_2-velocity). *If $\mu = f\mathfrak{m} \in \mathrm{Dom}(\mathsf{Ent}_\mathfrak{m})$ with $f \in \mathrm{L}^2(X,\mathfrak{m})$, then $t \mapsto \mu_t = f_t\mathfrak{m} \in \mathrm{AC}_{\mathrm{loc}}^2([0,+\infty); \mathscr{P}_2(X))$, where $f_t = \mathsf{P}_t f$ for all $t \geq 0$, with*

$$|\dot{\mu}_t|^2 \leq \mathsf{F}(f_t) = \int_{\{f_t>0\}} \frac{|\mathrm{D}f_t|_w^2}{f_t}\,\mathrm{d}\mathfrak{m}$$

for $\mathscr{L}^1$-a.e. $t > 0$.

As a simple consequence of Lemma 4.2 and the $\mathsf{BE}_w(\mathsf{c},\infty)$ condition, we can prove the following W_2-continuity property of the dual heat flow. For the same result under the standard $\mathsf{BE}(K,\infty)$ condition, see the last part of [16, Lemma 4.2].

Lemma 4.3 (W_2-continuity of $t \mapsto \mathsf{H}_t$). *Assume $(X,\mathsf{d},\mathfrak{m})$ satisfies $\mathsf{BE}_w(\mathsf{c},\infty)$. If $\mu \in \mathscr{P}_2(X)$, then $t \mapsto \mathsf{H}_t\mu \in \mathrm{C}([0,+\infty); \mathscr{P}_2(X))$. Equivalently, if $\mu \in \mathscr{P}_2(X)$ then $t \mapsto \mathsf{H}_t\mu$ is weakly continuous on $[0,+\infty)$ and $t \mapsto \int_X \mathsf{d}^2(x,x_0)\,\mathrm{d}\mathsf{H}_t\mu(x)$ is continuous on $[0,+\infty)$ whenever $x_0 \in X$ is given.*

Proof. If $\mu = f\mathfrak{m} \in \mathrm{Dom}(\mathsf{Ent}_\mathfrak{m})$ with $f \in \mathrm{L}^2(X,\mathfrak{m})$, then the W_2-continuity of the map $t \mapsto \mathsf{H}_t\mu$ follows immediately from Lemma 4.2. If $\mu \in \mathscr{P}_2(X)$, then we can find $(\mu_n)_{n\in\mathbb{N}} \subset \mathscr{P}_2^{\mathrm{ac}}(X)$ such that $\mu_n \xrightarrow{W_2} \mu$ as $n \to +\infty$. Possibly performing a truncation argument, we can also assume that $\mathsf{Ent}(\mu_n) < +\infty$ and $\mu_n = f_n\mathfrak{m}$ with $f_n \in \mathrm{L}^2(X,\mathfrak{m})$ for all $n \in \mathbb{N}$. If $t \geq 0$, then we can estimate

$$\limsup_{s\to t} W_2(\mathsf{H}_s\mu, \mathsf{H}_t\mu) \leq W_2(\mathsf{H}_s\mu_n, \mathsf{H}_s\mu) + W_2(\mathsf{H}_t\mu_n, \mathsf{H}_t\mu) + \limsup_{s\to t} W_2(\mathsf{H}_s\mu_n, \mathsf{H}_t\mu_n)$$

$$\leq 2M\,W_2(\mu_n, \mu),$$

where we have set $M = \sup_{s\in[\frac{t}{2},t+1]} \mathsf{c}(s)$. The conclusion thus follows by passing to the limit as $n \to +\infty$. $\qquad\square$

4.2. **Log-Harnack and** $\mathrm{L}\log\mathrm{L}$ **estimates.** The rest of this section is dedicated to the proof of the following fundamental regularization property of the dual heat semigroup, see [16, Theorem 4.8] for the same result in the standard $\mathsf{BE}(K,\infty)$ setting.

Theorem 4.4 ($\mathrm{L}\log\mathrm{L}$ regularization). *Assume* $(X,\mathsf{d},\mathfrak{m})$ *satisfies* $\mathsf{BE}_w(\mathsf{c},\infty)$. *If* $\mu \in \mathscr{P}_2(X)$, *then*

$$\mathsf{Ent}_\mathfrak{m}(\mathsf{H}_t\mu) \leq \frac{1}{2\,\mathrm{I}_{-2}(t)}\left(r^2 + \int_X \mathsf{d}^2(x,x_0)\,\mathrm{d}\mu(x)\right) - \log\mathfrak{m}(B_r(x_0)) \tag{4.4}$$

for all $x_0 \in X$ *and* $r,t > 0$. *In particular,* $\mathsf{H}_t(\mathscr{P}_2(X)) \subset \mathrm{Dom}(\mathsf{Ent}_\mathfrak{m})$ *for all* $t > 0$.

To prove Theorem 4.4, we follow the same strategy adopted in [16, Section 4.2]. Before the proof of Theorem 4.4, we need two preliminary results.

The first one is the following generalization of the differentiation formula proved in Lemma 3.3. The proof goes as that of [16, Lemma 4.5] with minor modifications, so we omit it.

Lemma 4.5 (General differentiation formula). *Assume* $(X,\mathsf{d},\mathfrak{m})$ *satisfies* $\mathsf{BE}_w(\mathsf{c},\infty)$ *and let* $\omega \in \mathrm{C}^2([0,+\infty))$. *If* $f \in \mathrm{Lip}_b(X) \cap \mathrm{W}^{1,2}(X,\mathsf{d},\mathfrak{m})$ *and* $\mu \in \mathscr{P}_2(X)$, *then for all* $t > 0$ *we have*

$$s \mapsto G(s) = \int_X \omega(\mathsf{P}_{t-s}f)\,\mathrm{d}\mathsf{H}_s\mu \in \mathrm{C}([0,t]) \cap \mathrm{C}^1((0,t))$$

with

$$G'(s) = \int_X \omega''(\mathsf{P}_{t-s}f)\,\Gamma(\mathsf{P}_{t-s}f)\,\mathrm{d}\mathsf{H}_s\mu$$

for all $s \in [0,t]$.

The second preliminary result is an adaptation to the abstract setting of an inequality proved for the first time in the Riemannian framework by Wang, see [143, Theorem 1.1(6)]. Lemma 4.6 below is the reformulation, under the more general $\mathsf{BE}_w(\mathsf{c},\infty)$ condition, of [16, Lemma 4.6]. Although its proof is very similar to that of [16, Lemma 4.6], we detail it here for the reader's convenience.

Lemma 4.6 (Wang log-Harnack inequality). *Assume* $(X,\mathsf{d},\mathfrak{m})$ *satisfies* $\mathsf{BE}_w(\mathsf{c},\infty)$ *and let* $\varepsilon > 0$. *If* $f \in \mathrm{L}^1(X,\mathfrak{m})$ *is non-negative, then*

$$\tilde{\mathsf{P}}_t(\log(f+\varepsilon))(y) \leq \log(\tilde{\mathsf{P}}_t f(x) + \varepsilon) + \frac{\mathsf{d}^2(x,y)}{4\,\mathrm{I}_{-2}(t)}, \tag{4.5}$$

for all $x,y \in X$ *and* $t > 0$, *where* I_{-2} *is as in* (3.15).

Proof. Let $\varepsilon > 0$ and $t > 0$ be fixed. We divide the proof in three steps.

Step 1. Assume $f \in \mathrm{Lip}_b(X) \cap \mathrm{W}^{1,2}(X,\mathsf{d},\mathfrak{m})$. Fix $x,y \in X$ and let $\gamma \in \mathrm{AC}([0,1];\mathbb{R})$ be such that $\gamma_0 = x$ and $\gamma_1 = y$. We set

$$\vartheta(r) = \frac{\mathrm{I}_{-2}(r)}{\mathrm{I}_{-2}(t)} \in [0,1] \quad \text{for all } r \in [0,t].$$

Note that $\vartheta \in \mathrm{Lip}([0,t])$. We also set $\omega_\varepsilon(r) = \log(r+\varepsilon) - \log\varepsilon$ for all $r \geq 0$. Note that $\omega \in \mathrm{C}^2([0,+\infty))$ with $\omega(0) = 0$. We claim that

$$s \mapsto \int_X \omega_\varepsilon(\mathsf{P}_{t-s}f)\,\mathrm{d}\mathsf{H}_s\delta_{\gamma_{\vartheta(s)}} \in \mathrm{AC}([0,t];\mathbb{R}). \tag{4.6}$$

To this aim, we set

$$G_\varepsilon(s,r) = \int_X \omega_\varepsilon(\mathsf{P}_{t-s}f)\,\mathrm{d}\mathsf{H}_s\delta_{\gamma_{\vartheta(r)}} \quad \text{for all } s,r \in [0,t].$$

On the one hand, by Lemma 4.5 applied with $\mu = \delta_{\gamma_{\vartheta(r)}}$ for each $r \in [0,t]$, we get that

$$|\partial_s G_\varepsilon(s,r)| = \left| \int_X \omega_\varepsilon''(\mathsf{P}_{t-s}f)\,\Gamma(\mathsf{P}_{t-s}f)\,\mathrm{d}\mathsf{H}_s\delta_{\gamma_{\vartheta(r)}} \right| \le \frac{\mathsf{c}(t-s)}{\varepsilon^2}\,\mathrm{Lip}(f)$$

for all $s \in (0,t)$ and $r \in [0,t]$ by (2.29) and Proposition 3.9, so that $s \mapsto G(s,r)$ is Lipschitz on $[0,t]$ uniformly in $r \in [0,t]$. On the other hand, by (2.12), Proposition 3.9, Jensen inequality and Theorem 3.16, we can estimate

$$\begin{aligned}
|G_\varepsilon(s,r_1) - G_\varepsilon(s,r_0)| &= \left| \int_X \omega_\varepsilon(\mathsf{P}_{t-s}f)\,\mathrm{d}\left(\mathsf{H}_s\delta_{\gamma_{\vartheta(r_1)}} - \mathsf{H}_s\delta_{\gamma_{\vartheta(r_0)}}\right) \right| \\
&\le \mathrm{Lip}(\omega_\varepsilon(\mathsf{P}_{t-s}f))\,W_1\left(\mathsf{H}_s\delta_{\gamma_{\vartheta(r_1)}}, \mathsf{H}_s\delta_{\gamma_{\vartheta(r_0)}}\right) \\
&\le \frac{\mathsf{c}(t-s)}{\varepsilon}\,\mathrm{Lip}(f)\,W_2\left(\mathsf{H}_s\delta_{\gamma_{\vartheta(r_1)}}, \mathsf{H}_s\delta_{\gamma_{\vartheta(r_0)}}\right) \\
&\le \frac{\mathsf{c}(t-s)\mathsf{c}(s)}{\varepsilon}\,\mathrm{Lip}(f)\,W_2\left(\delta_{\gamma_{\vartheta(r_1)}}, \delta_{\gamma_{\vartheta(r_0)}}\right) \\
&\le \frac{\mathsf{c}(t-s)\mathsf{c}(s)}{\varepsilon}\,\mathrm{Lip}(f)\,\mathsf{d}(\gamma_{\vartheta(r_1)}, \gamma_{\vartheta(r_0)})
\end{aligned}$$

for all $0 \le r_0 \le r_1 \le t$ and $s \in [0,t]$, so that $r \mapsto G(s,r) \in \mathrm{AC}([0,t];\mathbb{R})$ uniformly in $s \in [0,t]$. This prove the claim in (4.6). Now write

$$G_\varepsilon(s,r) = F_s^\varepsilon(\gamma_{\vartheta(r)}), \quad F_s^\varepsilon(x) = \tilde{\mathsf{P}}_s\left(\omega_\varepsilon(\mathsf{P}_{t-s}f)\right)(x) \text{ for all } x \in X.$$

Then, whenever $s \in [0,t]$, we can estimate

$$\partial_r G_\varepsilon(s,r) \le |\mathrm{D}F_s^\varepsilon|(\gamma_{\vartheta(r)})\,|\dot{\gamma}_{\vartheta(r)}|\,|\vartheta'(r)|$$

for $\mathscr{L}^1$-a.e. $r \in [0,t]$. By Proposition 3.25, we have

$$|\mathrm{D}F_s^\varepsilon|^2(\gamma_{\vartheta(r)}) = \left|\mathrm{D}\tilde{\mathsf{P}}_s\left(\omega_\varepsilon(\mathsf{P}_{t-s}f)\right)\right|^2(\gamma_{\vartheta(r)}) \le \mathsf{c}^2(s)\,\tilde{\mathsf{P}}_s\Gamma\left(\omega_\varepsilon(\mathsf{P}_{t-s}f)\right)(\gamma_{\vartheta(r)}) \tag{4.7}$$

for all $r \in [0,t]$. Recalling (3.58), by the chain rule (2.25) we can write

$$\begin{aligned}
\tilde{\mathsf{P}}_s\Gamma\left(\omega_\varepsilon(\mathsf{P}_{t-s}f)\right)(\gamma_{\vartheta(r)}) &= \int_X \Gamma\left(\omega_\varepsilon(\mathsf{P}_{t-s}f)\right)\,\mathrm{d}\mathsf{H}_s\gamma_{\vartheta(r)} \\
&= \int_X \Gamma\left(\omega_\varepsilon(\mathsf{P}_{t-s}f)\right)\,\mathsf{p}_s[\gamma_{\vartheta(r)}]\,\mathrm{d}\mathfrak{m} \\
&= \int_X (\omega_\varepsilon'(\mathsf{P}_{t-s}f))^2\,\Gamma(\mathsf{P}_{t-s}f)\,\mathsf{p}_s[\gamma_{\vartheta(r)}]\,\mathrm{d}\mathfrak{m} \\
&= -\int_X \omega_\varepsilon''(\mathsf{P}_{t-s}f)\,\Gamma(\mathsf{P}_{t-s}f)\,\mathrm{d}\mathsf{H}_s\gamma_{\vartheta(r)}
\end{aligned} \tag{4.8}$$

for all $s \in (0,t)$ and $r \in [0,t]$, since $(\omega_\varepsilon')^2 = -\omega_\varepsilon''$. Therefore, by Young inequality, we get that

$$\partial_r G_\varepsilon(s,r) \le \mathsf{c}^{-2}(s)\,|\mathrm{D}F_s^\varepsilon|^2(\gamma_{\vartheta(r)}) + \frac{\mathsf{c}^2(s)}{4}\,|\dot{\gamma}_{\vartheta(r)}|^2\,|\vartheta'(r)|^2$$

for all $s \in (0,t)$ and $\mathscr{L}^1$-a.e. $r \in [0,t]$. By combining (4.7) with (4.8), we conclude that

$$\frac{\mathrm{d}}{\mathrm{d}s} \int_X \omega_\varepsilon(\mathsf{P}_{t-s}f)\,\mathrm{d}\mathsf{H}_s\delta_{\gamma_{\vartheta(s)}} = \partial_s G_\varepsilon(s,s) + \partial_r G_\varepsilon(s,s)$$

$$\leq \int_X \omega_\varepsilon''(\mathsf{P}_{t-s}f)\, \Gamma(\mathsf{P}_{t-s}f)\, \mathrm{d}\mathsf{H}_s\delta_{\gamma_{\vartheta(s)}}$$

$$+ \mathsf{c}^{-2}(s)\, |\mathrm{D}F_s^\varepsilon|^2(\gamma_{\vartheta(s)}) + \frac{\mathsf{c}^2(s)}{4}\, |\dot{\gamma}_{\vartheta(s)}|^2\, |\vartheta'(s)|^2$$

$$\leq \frac{\mathsf{c}^2(s)}{4}\, |\dot{\gamma}_{\vartheta(s)}|^2\, |\vartheta'(s)|^2$$

for $\mathscr{L}^1$-a.e. $s \in [0,t]$, so that

$$\int_X \omega_\varepsilon(f)\, \mathrm{d}\mathsf{H}_t\delta_y - \int_X \omega_\varepsilon(\mathsf{P}_t f)\, \mathrm{d}\delta_x = \int_0^t \frac{\mathrm{d}}{\mathrm{d}s} \int_X \omega_\varepsilon(\mathsf{P}_{t-s}f)\, \mathrm{d}\mathsf{H}_s\delta_{\gamma_{\vartheta(s)}}\, \mathrm{d}s$$

$$\leq \frac{1}{4}\int_0^t \mathsf{c}^2(s)\, |\dot{\gamma}_{\vartheta(r)}|^2\, |\vartheta'(s)|^2\, \mathrm{d}s$$

$$= \frac{1}{4\,\mathrm{I}_{-2}(t)} \int_0^1 |\dot{\gamma}_\tau|^2\, \mathrm{d}\tau.$$

Recalling Proposition 2.2, we immediately deduce that

$$\tilde{\mathsf{P}}_t(\omega_\varepsilon(f))(y) \leq \omega_\varepsilon(\tilde{\mathsf{P}}_t f)(x) + \frac{\mathsf{d}^2(x,y)}{4\,\mathrm{I}_{-2}(t)}, \tag{4.9}$$

for all $x, y \in X$, whenever $f \in \mathrm{Lip}_b(X) \cap \mathrm{W}^{1,2}(X,\mathsf{d},\mathfrak{m})$.

Step 2. Assume $f \in \mathrm{L}^1(X,\mathfrak{m}) \cap \mathrm{L}^\infty(X,\mathfrak{m})$. Since $\mathrm{W}^{1,2}(X,\mathsf{d},\mathfrak{m})$ is dense in $\mathrm{L}^2(X,\mathfrak{m})$, by (2.28) we can find $(f_n)_{n\in\mathbb{N}} \subset \mathrm{Lip}_b(X) \cap \mathrm{W}^{1,2}(X,\mathsf{d},\mathfrak{m})$ such that $f_n \to f$ $\mathfrak{m}$-a.e. in X as $n \to +\infty$. Since $f \in \mathrm{L}^\infty(X,\mathfrak{m})$ is non-negative, by (2.26) and (2.27) we can also assume that $0 \leq f_n \leq \|f\|_{\mathrm{L}^\infty(X,\mathfrak{m})}$ for all $n \in \mathbb{N}$. By (3.59) and the Dominated Convergence Theorem, we thus get that $\tilde{\mathsf{P}}_t f_n(x) \to \tilde{\mathsf{P}}_t f(x)$ for all $x \in X$ as $n \to +\infty$. Hence, by Fatou Lemma and by (4.9) in Step 1, we get

$$\tilde{\mathsf{P}}_t(\omega_\varepsilon(f))(y) \leq \liminf_{n\to+\infty} \tilde{\mathsf{P}}_t(\omega_\varepsilon(f_n))(y)$$

$$\leq \lim_{n\to+\infty} \omega_\varepsilon(\tilde{\mathsf{P}}_t f_n)(x) + \frac{\mathsf{d}^2(x,y)}{4\,\mathrm{I}_{-2}(t)}$$

$$= \omega_\varepsilon(\tilde{\mathsf{P}}_t f)(x) + \frac{\mathsf{d}^2(x,y)}{4\,\mathrm{I}_{-2}(t)},$$

for all $x, y \in X$, proving (4.9) whenever $f \in \mathrm{L}^1(X,\mathfrak{m}) \cap \mathrm{L}^\infty(X,\mathfrak{m})$.

Step 3. Assume $f \in \mathrm{L}^1(X,\mathfrak{m})$. Then $f_n = f \wedge n \in \mathrm{L}^1(X,\mathfrak{m}) \cap \mathrm{L}^\infty(X,\mathfrak{m})$ for all $n \in \mathbb{N}$ and thus the conclusion follows by Step 2 and the Monotone Convergence Theorem. $\square$

An simple but interesting consequence of Lemma 4.6 is the following result, see [16, Corollary 4.7] for the same result in the standard $\mathsf{BE}(K,\infty)$ setting.

Corollary 4.7 (Wang inequality for $\mathsf{p}_t[\,\cdot\,]$). *Assume $(X,\mathsf{d},\mathfrak{m})$ satisfies $\mathsf{BE}_w(\mathsf{c},\infty)$. For every $\varepsilon, t > 0$ and every $y \in X$, we have*

$$\int_X \mathsf{p}_t[y]\log(\mathsf{p}_t[y] + \varepsilon)\, \mathrm{d}\mathfrak{m} \leq \log(\mathsf{p}_{2t}[y](x) + \varepsilon) + \frac{\mathsf{d}^2(x,y)}{4\,\mathrm{I}_{-2}(t)}$$

for $\mathfrak{m}$-a.e. $x \in X$. In particular, if $\mathfrak{m} \in \mathscr{P}(X)$, then for all $t > 0$ and $y \in X$ we have

$$\mathsf{p}_{2t}[y](x) \geq \exp\left(-\frac{\mathsf{d}^2(x,y)}{4\,\mathrm{I}_{-2}(t)}\right) \tag{4.10}$$

for $\mathfrak{m}$-a.e. $x \in X$.

48 G. STEFANI

Proof. Thanks to Lemma 3.24(i), the result immediately follows by applying Lemma 4.6 to $f = \mathsf{p}_t[y] \in \mathrm{L}^1(X, \mathfrak{m})$ and passing to the limit as $\varepsilon \to 0^+$. $\qquad\square$

We are now ready to prove the main result of this section.

Proof of Theorem 4.4. Let $x_0 \in X$ and $r, t > 0$ be fixed. We divide the proof in two steps.

Step 1. Assume $\mu = f\mathfrak{m}$ for some non-negative $f \in \mathrm{L}^1(X, \mathfrak{m})$. Let us set $f_t = \mathsf{P}_t f$ and $\tilde{f}_t = \tilde{\mathsf{P}}_t f$. Note that $\tilde{f}_t = f_t$ $\mathfrak{m}$-a.e. in X since, as in the proof of Corollary 4.7, we can compute

$$\int_X \tilde{f}_t \, \varphi \, \mathrm{d}\mathfrak{m} = \|\varphi\|_{\mathrm{L}^1(X,\mathfrak{m})} \int_X f \, \mathrm{d}\mathsf{H}_t(\bar{\varphi} \, \mathfrak{m}) = \int_X f \, \mathsf{P}_t\varphi \, \mathrm{d}\mathfrak{m} = \int_X f_t \, \varphi \, \mathrm{d}\mathfrak{m}$$

for all non-negative Borel $\varphi \in \mathrm{L}^1(X, \mathfrak{m}) \cap \mathrm{L}^\infty(X, \mathfrak{m})$, where $\varphi = \bar{\varphi} \, \|\varphi\|_{\mathrm{L}^1(X,\mathfrak{m})}$. By (3.58) and Jensen inequality, we can estimate

$$\tilde{f}_t(x) \log(\tilde{f}_t(x) + \varepsilon) = \left(\int_X f \, \mathrm{d}\mathsf{H}_t\delta_x \right) \log \left(\varepsilon + \int_X f \, \mathrm{d}\mathsf{H}_t\delta_x \right)$$
$$= \left(\int_X \mathsf{p}_t[x] \, \mathrm{d}\mu \right) \log \left(\varepsilon + \int_X \mathsf{p}_t[x] \, \mathrm{d}\mu \right) \qquad (4.11)$$
$$\leq \int_X \mathsf{p}_t[x](y) \log(\mathsf{p}_t[x](y) + \varepsilon) \, \mathrm{d}\mu(y)$$

for all $x \in X$ and $\varepsilon > 0$. Integrating (4.11), by Lemma 3.24(ii), Tonelli Theorem and Corollary 4.7 we thus get

$$\mathsf{Ent}_{\mathfrak{m}}(\mathsf{H}_t\mu) = \int_X f_t \log(f_t + \varepsilon) \, \mathrm{d}\mathfrak{m} \leq \int_X \left(\int_X \mathsf{p}_t[x](y) \log(\mathsf{p}_t[x](y) + \varepsilon) \, \mathrm{d}\mu(y) \right) \mathrm{d}\mathfrak{m}(x)$$
$$= \int_X \left(\int_X \mathsf{p}_t[y](x) \log(\mathsf{p}_t[y](x) + \varepsilon) \, \mathrm{d}\mathfrak{m}(x) \right) \mathrm{d}\mu(y)$$
$$\leq \int_X \left(\log(\mathsf{p}_{2t}[y](z) + \varepsilon) + \frac{\mathsf{d}^2(z,y)}{4\,\mathrm{I}_{-2}(t)} \right) \mathrm{d}\mu(y)$$
$$(4.12)$$

for $\mathfrak{m}$-a.e. $z \in X$. Now let $q = \mathfrak{m}(B_r(x_0))$ and define $\nu = q^{-1}\mathfrak{m}\llcorner B_r(x_0)$. Integrating (4.12), by Tonelli Theorem and Jensen inequality we get

$$\mathsf{Ent}_{\mathfrak{m}}(\mathsf{H}_t\mu) \leq \int_X \int_X \left(\log(\mathsf{p}_{2t}[y](z) + \varepsilon) + \frac{\mathsf{d}^2(z,y)}{4\,\mathrm{I}_{-2}(t)} \right) \mathrm{d}\mu(y) \, \mathrm{d}\nu(z)$$
$$= \int_X \int_X \log(\mathsf{p}_{2t}[y](z) + \varepsilon) \, \mathrm{d}\mu(y) \, \mathrm{d}\nu(z) + \frac{1}{2\,\mathrm{I}_{-2}(t)} \left(r^2 + \int_X \mathsf{d}^2(y, x_0) \, \mathrm{d}\mu(y) \right)$$
$$\leq \log \left(\varepsilon + \int_X \int_X \mathsf{p}_{2t}[y](z) \, \mathrm{d}\nu(z) \, \mathrm{d}\mu(y) \right) + \frac{1}{2\,\mathrm{I}_{-2}(t)} \left(r^2 + \int_X \mathsf{d}^2(y, x_0) \, \mathrm{d}\mu(y) \right)$$
$$= \log \left(\varepsilon + \int_X \frac{1}{q} \int_{B_r(x_0)} \mathsf{p}_{2t}[y](z) \, \mathrm{d}\mathfrak{m}(z) \, \mathrm{d}\mu(y) \right) + \frac{1}{2\,\mathrm{I}_{-2}(t)} \left(r^2 + \int_X \mathsf{d}^2(y, x_0) \, \mathrm{d}\mu(y) \right)$$
$$\leq \log(q^{-1} + \varepsilon) + \frac{1}{2\,\mathrm{I}_{-2}(t)} \left(r^2 + \int_X \mathsf{d}^2(y, x_0) \, \mathrm{d}\mu(y) \right)$$

Passing to the limit as $\varepsilon \to 0^+$, we prove (4.4) whenever $\mu = f\mathfrak{m}$ for some non-negative $f \in \mathrm{L}^1(X, \mathfrak{m})$.

Step 2. Now let $\mu \in \mathscr{P}_2(X)$. We can find $(\mu_n)_{n\in\mathbb{N}} \subset \mathscr{P}_2^{\mathrm{ac}}(X)$ such that $\mu_n \xrightarrow{W_2} \mu$ as $n \to +\infty$. Thanks to the lower semicontinuity property of the entropy and the properties

of Wasserstein distance, by Step 1 we get that

$$\mathsf{Ent}_\mathfrak{m}(\mathsf{H}_t\mu) \leq \liminf_{n\to+\infty} \mathsf{Ent}_\mathfrak{m}(\mathsf{H}_t\mu_n)$$

$$\leq \frac{1}{2\,\mathsf{I}_{-2}(t)} \left(r^2 + \lim_{n\to+\infty} \int_X \mathsf{d}^2(x, x_0)\,\mathrm{d}\mu_n(x) \right) - \log\mathfrak{m}(B_r(x_0))$$

$$= \frac{1}{2\,\mathsf{I}_{-2}(t)} \left(r^2 + \int_X \mathsf{d}^2(x, x_0)\,\mathrm{d}\mu(x) \right) - \log\mathfrak{m}(B_r(x_0))$$

and the proof is complete. $\qquad\square$

5. Entropic inequalities in groups

We now proceed with the core argument of the proof of the entropic inequalities, adapting the action and the entropy estimates established in [16, Section 4.3] (see also [72, Section 4.2] for a closely related approach in the finite dimensional case). Since we frequently consider curves $s \mapsto \mu_s = f_s\mathfrak{m} \in \mathrm{AC}^2([0,1]; (\mathscr{P}_2(X), W_2))$ with $s \mapsto f_s \in \mathrm{C}^1([0,1]; \mathrm{L}^p(X, \mathfrak{m}))$ for some $p \in [1, +\infty)$, we shall denote by $s \mapsto \dot{f}_s \in \mathrm{C}([0,1]; \mathrm{L}^p(X, \mathfrak{m}))$ the functional derivative in $\mathrm{L}^p(X, \mathfrak{m})$, keeping the notation $s \mapsto |\dot{\mu}_s|$ for the metric derivative in $(\mathscr{P}_2(X), W_2)$.

5.1. Strongly regular curves.
Instead of considering *regular curves* in $\mathscr{P}_2(X)$ as in [16, Definition 4.10], we deal with *strongly regular curves* defined as follows.

Definition 5.1 (Strongly regular curve in $\mathscr{P}_2(X)$)**.** We say that a curve $s \mapsto \mu_s \in \mathrm{AC}^2([0,1]; \mathscr{P}_2(X))$ is *strongly regular* if $\mu_s = f_s\mathfrak{m}$ for all $s \in [0,1]$ with

$$s \mapsto f_s \in \mathrm{C}^1([0,1]; \mathrm{L}^2(X, \mathfrak{m})). \tag{5.1}$$

Note that the L^2-integrability property in (5.1) immediately gives that

$$\sup_{s\in[0,1]} \mathsf{Ent}_\mathfrak{m}(\mu_s) < +\infty \tag{5.2}$$

whenever $s \mapsto \mu_s$ is a strongly regular curve. Under the standard $\mathsf{BE}(K, N)$ condition, the uniform upper control of the entropy along *regular curves* and, most importantly, the absolute continuity property of *regular curves* at each time with respect to the reference measure $\mathfrak{m}$, are gained from the absolute continuity of heated measures and the $\mathrm{L}\log\mathrm{L}$-regularization property of the dual heat flow $(\mathsf{H}_t)_{t\geq 0}$. Having the more general $\mathsf{BE}_w(\mathsf{c}, \infty)$ condition at disposal, the absolute continuity of heated measures and the $\mathrm{L}\log\mathrm{L}$-regularization property are still available, recall Corollary 3.22 and Theorem 4.4, but the application of the dual heat flow to an arbitrary curve $s \mapsto \mu_s \in \mathrm{AC}^2([0,1]; \mathscr{P}_2(X))$ drastically affects its Wasserstein velocity. In more precise terms, by (3.36) we immediately get that the heated curve $s \mapsto \mu_s^t = \mathsf{H}_t\mu_s \in \mathrm{AC}^2([0,1]; \mathscr{P}_2(X))$ satisfies

$$|\dot{\mu}_s^t| \leq \mathsf{c}(t)\,|\dot{\mu}_s| \quad \text{for } \mathscr{L}^1\text{-a.e. } s \in [0,1] \tag{5.3}$$

for all $t \geq 0$, but, since the general strategy developed in [16, 72] requires the approximation of $s \mapsto \mu_s$ with more regular AC^2-curves in $\mathscr{P}_2(X)$ having Wasserstein velocity closer and closer to the velocity of the original curve (see equation (4.20) in the statement of [16, Proposition 4.11]), the velocity estimate in (5.3) becomes useful only if (at least)

$$\underline{\mathsf{c}}(0^+) = \liminf_{t\to+\infty} \mathsf{c}(t) = 1,$$

a condition which is not available in sub-Riemannian manifolds. For this reason, taking inspiration from the regularization procedure performed in [20, Theorem 4.8], instead of relying on the contraction property of the dual heat flow, in our group-modeled framework (see Section 5.5 below) we will regularize arbitrary AC^2-curves in $\mathscr{P}_2(X)$ via the group (left-)convolution with suitable L^1-kernels (see also Remark 5.14 below for a strictly related discussion).

Since our strategy will be deeply based on the L^2-regularity property (5.1) of strongly regular curves, we will frequently take advantage of the following product rule for the functional derivative of Lipschitz curves in $\mathrm{L}^2(X, \mathfrak{m})$. We leave its elementary proof to the interested reader.

Lemma 5.2 (Product rule for Lip-curves in L^2). *If $s \mapsto a_s, b_s \in \mathrm{Lip}([0,1]; \mathrm{L}^2(X, \mathfrak{m}))$, then $s \mapsto \int_X a_s\, b_s\, \mathrm{d}\mathfrak{m} \in \mathrm{Lip}([0,1]; \mathbb{R})$ with*

$$\frac{\mathrm{d}}{\mathrm{d}s} \int_X a_s\, b_s\, \mathrm{d}\mathfrak{m} = \int_X \dot{a}_s\, b_s\, \mathrm{d}\mathfrak{m} + \int_X a_s\, \dot{b}_s\, \mathrm{d}\mathfrak{m}$$

for $\mathscr{L}^1$-a.e. $s \in [0,1]$.

5.2. **Action along strongly regular curves.** We begin by fixing two important functions. Here and in the rest of the section, we let

$$\vartheta \in \mathrm{C}^1([0,1]; [0,1]) \text{ with } \vartheta(i) = i, \ i = 0, 1,$$

and

$$\eta \in \mathrm{C}^2([0,1]; [0, +\infty)) \text{ with } \dot{\eta} \geq 0 \text{ and } \eta(s) > 0 \text{ for all } s > 0.$$

In the following result, analogous to [16, Lemma 4.13] and [72, Lemma 4.12], we compute the derivative of the action along a strongly regular curve. Recall that $\mathrm{Lip}_\star(X)$ was defined in (3.37) as the set of non-negative Lipschitz functions with bounded support.

Lemma 5.3 (Derivative of action along strongly regular curves). *Assume $(X, \mathsf{d}, \mathfrak{m})$ satisfies $\mathsf{BE}_w(\mathsf{c}, \infty)$. Let $s \mapsto \mu_s = f_s \mathfrak{m} \in \mathrm{AC}^2([0,1]; \mathscr{P}_2(X))$ be a strongly regular curve and define*

$$s \mapsto \tilde{\mu}_s = \mathsf{H}_{\eta(s)} \mu_{\vartheta(s)} = \tilde{f}_s \mathfrak{m} \in \mathrm{C}([0,1]; \mathscr{P}_2(X)). \tag{5.4}$$

If $\varphi \in \mathrm{Lip}_\star(X)$ and $\varphi_s = Q_s \varphi$ for all $s \in [0,1]$, then

$$s \mapsto \int_X \varphi_s\, \mathrm{d}\tilde{\mu}_s \in \mathrm{Lip}([0,1]; \mathbb{R}) \tag{5.5}$$

with

$$\frac{\mathrm{d}}{\mathrm{d}s} \int_X \varphi_s\, \mathrm{d}\tilde{\mu}_s = -\frac{1}{2} \int_X |\mathrm{D}\varphi_s|^2\, \mathrm{d}\tilde{\mu}_s - \dot{\eta}(s) \int_X \Gamma(\tilde{f}_s, \varphi_s)\, \mathrm{d}\mathfrak{m} + \dot{\vartheta}(s) \int_X \dot{f}_{\vartheta(s)}\, \mathsf{P}_{\eta(s)} \varphi_s\, \mathrm{d}\mathfrak{m} \tag{5.6}$$

for $\mathscr{L}^1$-a.e. $s \in (0,1)$.

Note that the function η equals $\eta(s) = st$ for all $s \in [0,1]$, $t \geq 0$, in [16, Lemma 4.13], while in [72, Lemma 4.12] η is an increasing C^1 time-change satisfying $\eta(0) = 0$ depending on the dimension $N \in (0, +\infty)$. For similar variations of curves via semigroup operators, we refer the reader to [67, 124].

In the proof of Lemma 5.3, having in mind the product rule provided by Lemma 5.2, we will use the following result about the L^2-functional derivative of the Hopf–Lax semigroup. For the reader's convenience, we give a sketch of its proof (see also the discussion in the first three paragraphs of the proof of [14, Lemma 6.1]).

Lemma 5.4 (Hopf–Lax semigroup is Lip in L^2). *If $f \in \mathrm{Lip}_\star(X)$, then $s \mapsto Q_s f \in \mathrm{Lip}([0, +\infty); L^2(X, \mathfrak{m}))$ with*

$$\frac{\mathrm{d}^+}{\mathrm{d}s} Q_s f = -\frac{1}{2} |\mathrm{D}Q_s f| \quad in\ L^2(X, \mathfrak{m}) \tag{5.7}$$

for $\mathscr{L}^1$-a.e. $s > 0$.

Proof. If $f \in \mathrm{Lip}_\star(X)$, then also $Q_s f \in \mathrm{Lip}_\star(X)$ with $\mathrm{supp}(Q_s f) \subset \mathrm{supp}\, f$ for all $s \geq 0$. Hence by (2.6) we can estimate

$$\|Q_{s+h} f - Q_s f\|_{L^2(X, \mathfrak{m})} \leq \sqrt{\mathfrak{m}(\mathrm{supp}\, f)} \cdot 2\, \mathrm{Lip}(f)^2\, h$$

for all $0 < s \leq s + h$, so that the conclusion follows by (2.5) and the Dominated Convergence Theorem. $\qquad\square$

Proof of Lemma 5.3. By Theorem 3.16(ii), we can estimate

$$
\begin{aligned}
W_2(\tilde\mu_{s_0}, \tilde\mu_{s_1}) &= W_2(\mathsf{H}_{\eta(s_0)} \mu_{\vartheta(s_0)}, \mathsf{H}_{\eta(s_1)} \mu_{\vartheta(s_1)}) \\
&\leq W_2(\mathsf{H}_{\eta(s_0)} \mu_{\vartheta(s_0)}, \mathsf{H}_{\eta(s_0)} \mu_{\vartheta(s_1)}) + W_2(\mathsf{H}_{\eta(s_0)} \mu_{\vartheta(s_1)}, \mathsf{H}_{\eta(s_1)} \mu_{\vartheta(s_1)}) \\
&\leq \mathsf{c}(\eta(s_0))\, W_2(\mu_{\vartheta(s_0)}, \mu_{\vartheta(s_1)}) + W_2(\mathsf{H}_{\eta(s_0)} \mu_{\vartheta(s_1)}, \mathsf{H}_{\eta(s_1)} \mu_{\vartheta(s_1)})
\end{aligned}
$$

for all $s_0, s_1 \in [0, 1]$, so that (5.4) readily follows from Lemma 4.3. Since we have $s \mapsto f_s \in C^1([0, 1]; L^2(X, \mathfrak{m}))$ by (5.1) in Definition 5.1, we also have $s \mapsto \tilde{f}_s = \mathsf{P}_{\eta(s)} f_{\vartheta(s)} \in C^1((0, 1]; L^2(X, \mathfrak{m}))$ with

$$\frac{\mathrm{d}}{\mathrm{d}s} \tilde{f}_s = \dot\eta(s)\, \Delta_{\mathsf{d}, \mathfrak{m}} \mathsf{P}_{\eta(s)} f_{\vartheta(s)} + \dot\vartheta(s)\, \mathsf{P}_{\eta(s)} \dot{f}_{\vartheta(s)} \quad in\ L^2(X, \mathfrak{m}) \tag{5.8}$$

for all $s \in (0, 1]$. Hence, by Lemma 5.2 and Lemma 5.4, we get (5.5) and we can compute

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s} \int_X \varphi_s \,\mathrm{d}\tilde\mu_s &= \frac{\mathrm{d}}{\mathrm{d}s} \int_X \varphi_s\, \tilde{f}_s \,\mathrm{d}\mathfrak{m} = \int_X \tilde{f}_s \frac{\mathrm{d}}{\mathrm{d}s} \varphi_s \,\mathrm{d}\mathfrak{m} + \int_X \varphi_s \frac{\mathrm{d}}{\mathrm{d}s} \tilde{f}_s \,\mathrm{d}\mathfrak{m} \\
&= -\frac{1}{2} \int_X |\mathrm{D}\varphi_s| \,\mathrm{d}\tilde\mu_s + \int_X \left(\dot\eta(s)\, \Delta_{\mathsf{d}, \mathfrak{m}} \tilde{f}_s + \dot\vartheta(s)\, \mathsf{P}_{\eta(s)} \dot{f}_{\vartheta(s)} \right) \varphi_s \,\mathrm{d}\mathfrak{m} \\
&= -\frac{1}{2} \int_X |\mathrm{D}\varphi_s| \,\mathrm{d}\tilde\mu_s - \dot\eta(s) \int_X \Gamma(\tilde{f}_s, \varphi_s) \,\mathrm{d}\mathfrak{m} + \dot\vartheta(s) \int_X \dot{f}_{\vartheta(s)}\, \mathsf{P}_{\eta(s)} \varphi_s \,\mathrm{d}\mathfrak{m}
\end{aligned}
$$

for $\mathscr{L}^1$-a.e. $s \in (0, 1)$ by (5.7) and (2.44), since $\varphi_s \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$ for all $s \in [0, 1]$, proving (5.6). This concludes the proof. $\qquad\square$

5.3. Entropy along strongly regular curves.

We now introduce some notation. For every $\varepsilon > 0$, we let

$$\ell_\varepsilon(r) = \log(\varepsilon + r) \quad \text{for all } r \geq 0,$$

and for all $\mu \in \mathscr{P}_2(X)$ we define

$$E_\varepsilon(\mu) = \begin{cases} \displaystyle\int_X \ell_\varepsilon(f) \,\mathrm{d}\mu & \text{if } \mu = f\mathfrak{m}, \\ +\infty & \text{otherwise.} \end{cases} \tag{5.9}$$

Note that

$$E_\varepsilon(\mu) \geq \mathsf{Ent}_{\mathfrak{m}}(\mu) \quad \text{for all } \mu \in \mathscr{P}_2(X) \tag{5.10}$$

and, moreover, that if $\mu = f\mathfrak{m} \in \mathscr{P}_2(X)$ with $f \in L^2(X, \mathfrak{m})$ then

$$E_\varepsilon(\mu) = \log\varepsilon + \int_X (\ell_\varepsilon(f) - \log\varepsilon) \,\mathrm{d}\mu \leq \log\varepsilon + \frac{1}{\varepsilon} \int_X f^2 \,\mathrm{d}\mathfrak{m} < +\infty. \tag{5.11}$$

For this reason, we set

$$\hat{\ell}_\varepsilon(r) = \ell_\varepsilon(r) - \log\varepsilon \quad \text{for all } r \geq 0.$$

We also let

$$p_\varepsilon(r) = \hat{\ell}_\varepsilon(r) + r\,\hat{\ell}_\varepsilon'(r) \quad \text{for all } r \geq 0.$$

Note that

$$p_\varepsilon'(r) = 2\ell_\varepsilon'(r) + r\,\ell_\varepsilon''(r) = \frac{r + 2\varepsilon}{(r + \varepsilon)^2} \quad \text{for all } r \geq 0. \tag{5.12}$$

Finally, note that if $\varepsilon \in (0, 1]$ then $\log(\varepsilon + r) \leq \log(1 + r) \leq r$ for all $r \geq 0$. Thus, if $\mu = f\mathfrak{m} \in \mathscr{P}_2(X)$ for some $f \in \mathrm{L}^2(X, \mathfrak{m})$, then $[\log(\varepsilon + f)]^- \leq [\log f]^- \in \mathrm{L}^1(X, \mu)$ by (2.15) and $[\log(\varepsilon + f)]^+ \leq f \in \mathrm{L}^1(X, \mu)$, so that

$$\mathsf{Ent}_\mathfrak{m}(\mu) \leq E_\varepsilon(\mu) \leq |E_\varepsilon(\mu)| \leq \int_X [\log f]^-\, \mathrm{d}\mu + \int_X f^2\, \mathrm{d}\mathfrak{m}$$

and

$$\mathsf{Ent}_\mathfrak{m}(\mu) = \lim_{\varepsilon \to 0^+} E_\varepsilon(\mu), \tag{5.13}$$

by the Dominated Convergence Theorem.

In the following result, analogous to [16, Lemma 4.15] and [72, Lemma 4.13], we compute the derivative of the truncated entropy E_ε defined in (5.9) along a strongly regular curve.

Lemma 5.5 (Derivative of E_ε along strongly regular curves). *Assume $(X, \mathsf{d}, \mathfrak{m})$ satisfies $\mathsf{BE}_w(\mathsf{c}, \infty)$ and let $\varepsilon > 0$. Under the same assumptions of Lemma 5.3 and the notation above, we have*

$$s \mapsto E_\varepsilon(\tilde{\mu}_s) \in \mathrm{C}^1((0, 1]; \mathbb{R}) \tag{5.14}$$

with

$$\frac{\mathrm{d}}{\mathrm{d}s} E_\varepsilon(\tilde{\mu}_s) \leq -\dot{\eta}(s) \int_X \Gamma(g_s^\varepsilon)\, \mathrm{d}\tilde{\mu}_s + \dot{\vartheta}(s) \int_X \dot{f}_{\vartheta(s)} \mathsf{P}_{\eta(s)} g_s^\varepsilon\, \mathrm{d}\mathfrak{m} \tag{5.15}$$

for all $s \in (0, 1]$, where $g_s^\varepsilon = p_\varepsilon(\tilde{f}_s)$ for all $s \in [0, 1]$.

Proof. Let $\varepsilon > 0$ be fixed. Since $\hat{\ell}_\varepsilon \in \mathrm{C}^1([0, +\infty); [0, +\infty)) \cap \mathrm{Lip}([0, +\infty); [0, +\infty))$ with $\hat{\ell}_\varepsilon(0) = 0$, by (5.8) and the Mean Value Theorem we get that $s \mapsto \hat{\ell}_\varepsilon(\tilde{f}_s) \in \mathrm{C}^1((0, 1]; \mathrm{L}^2(X, \mathfrak{m}))$ with

$$\frac{\mathrm{d}}{\mathrm{d}s} \hat{\ell}_\varepsilon(\tilde{f}_s) = \hat{\ell}_\varepsilon'(\tilde{f}_s) \frac{\mathrm{d}}{\mathrm{d}s} \tilde{f}_s \quad \text{in } \mathrm{L}^2(X, \mathfrak{m})$$

for all $s \in (0, 1]$. Similarly, since $p_\varepsilon \in \mathrm{C}^1([0, +\infty); \mathbb{R}) \cap \mathrm{Lip}([0, +\infty); \mathbb{R})$ with $p_\varepsilon(0) = 0$, we also have that $s \mapsto g_s^\varepsilon = p_\varepsilon(\tilde{f}_s) \in \mathrm{C}^1((0, 1]; \mathrm{L}^2(X, \mathfrak{m}))$. In addition, recalling that $\tilde{f}_s = \mathsf{P}_{\eta(s)} f_{\vartheta(s)}$ for all $s \in [0, 1]$, by Lemma 2.1 we also have that $s \mapsto \tilde{f}_s \in \mathrm{C}((0, 1]; \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}))$ and thus also $s \mapsto g_s^\varepsilon \in \mathrm{C}((0, 1]; \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m}))$ by the chain rule (2.25). Hence, again by (5.8) and by Lemma 5.2, we get that

$$s \mapsto E_\varepsilon(\tilde{\mu}_s) = \log\varepsilon + \int_X \hat{\ell}_\varepsilon(\tilde{f}_s)\, \tilde{f}_s\, \mathrm{d}\mathfrak{m} \in \mathrm{C}^1((0, 1]; \mathbb{R}),$$

proving (5.14), and we can compute

$$\frac{\mathrm{d}}{\mathrm{d}s} E_\varepsilon(\tilde{\mu}_s) = \frac{\mathrm{d}}{\mathrm{d}s} \int_X \hat{\ell}_\varepsilon(\tilde{f}_s)\, \tilde{f}_s\, \mathrm{d}\mathfrak{m}$$

$$= \int_X \tilde{f}_s \frac{\mathrm{d}}{\mathrm{d}s} \hat{\ell}_\varepsilon(\tilde{f}_s) + \hat{\ell}_\varepsilon(\tilde{f}_s) \frac{\mathrm{d}}{\mathrm{d}s} \tilde{f}_s\, \mathrm{d}\mathfrak{m}$$

$$= \int_X p_\varepsilon(\tilde{f}_s) \frac{\mathrm{d}}{\mathrm{d}s} \tilde{f}_s \, \mathrm{d}\mathfrak{m}$$

$$= \int_X g_s^\varepsilon \left(\dot{\eta}(s) \, \Delta_{\mathsf{d},\mathfrak{m}} \tilde{f}_s + \dot{\vartheta}(s) \, \mathsf{P}_{\eta(s)} \dot{f}_{\vartheta(s)} \right) \mathrm{d}\mathfrak{m}$$

for all $s \in (0,1]$. By the integration-by-part formula (2.44) and the chain rule (2.47), we can write

$$\int_X g_s^\varepsilon \, \Delta_{\mathsf{d},\mathfrak{m}} \tilde{f}_s \, \mathrm{d}\mathfrak{m} = -\int_X \Gamma(g_s^\varepsilon, \tilde{f}_s) \, \mathrm{d}\mathfrak{m} = -\int_X \Gamma(p_\varepsilon(\tilde{f}_s), \tilde{f}_s) \, \mathrm{d}\mathfrak{m} = -\int_X p_\varepsilon'(\tilde{f}_s) \, \Gamma(\tilde{f}_s) \, \mathrm{d}\mathfrak{m}$$

Observing that $-p_\varepsilon'(r) \leq -r \, (p_\varepsilon'(r))^2$ for all $r \geq 0$, since $r \, p_\varepsilon'(r) \leq 1$ for all $r \geq 0$ by (5.12), again by the chain rule (2.25) we can estimate

$$\int_X g_s^\varepsilon \, \Delta_{\mathsf{d},\mathfrak{m}} \tilde{f}_s \, \mathrm{d}\mathfrak{m} = -\int_X p_\varepsilon'(\tilde{f}_s) \, \Gamma(\tilde{f}_s) \, \mathrm{d}\mathfrak{m}$$

$$\leq -\int_X \tilde{f}_s \, (p_\varepsilon'(\tilde{f}_s))^2 \, \Gamma(\tilde{f}_s) \, \mathrm{d}\mathfrak{m}$$

$$= -\int_X \Gamma(p_\varepsilon(\tilde{f}_s)) \, \mathrm{d}\tilde{\mu}_s$$

$$= -\int_X \Gamma(g_s^\varepsilon) \, \mathrm{d}\tilde{\mu}_s$$

for all $s \in (0,1]$, so that

$$\frac{\mathrm{d}}{\mathrm{d}s} \, E_\varepsilon(\tilde{\mu}_s) = \dot{\eta}(s) \int_X g_s^\varepsilon \, \Delta_{\mathsf{d},\mathfrak{m}} \tilde{f}_s \, \mathrm{d}\mathfrak{m} + \dot{\vartheta}(s) \int_X g_s^\varepsilon \, \mathsf{P}_{\eta(s)} \dot{f}_{\vartheta(s)} \, \mathrm{d}\mathfrak{m}$$

$$\leq -\dot{\eta}(s) \int_X \Gamma(g_s^\varepsilon) \, \mathrm{d}\tilde{\mu}_s + \dot{\vartheta}(s) \int_X \dot{f}_{\vartheta(s)} \, \mathsf{P}_{\eta(s)} g_s^\varepsilon \, \mathrm{d}\mathfrak{m}$$

for all $s \in (0,1]$, concluding the proof of (5.15). □

5.4. Action and entropy along regular curves.

We now come to the following crucial result connecting the action estimate obtained in Lemma 5.3 with the entropic inequality proved in Lemma 5.5. For the same result in the standard $\mathsf{BE}(K,N)$ framework, we refer the reader to [16, Theorem 4.16] and [72, Proposition 4.16].

Theorem 5.6 (Action and entropy along strongly regular curves). *Assume* $(X, \mathsf{d}, \mathfrak{m})$ *satisfies* $\mathsf{BE}_w(\mathsf{c}, \infty)$ *and is heat-smoothing as in Definition 3.27. Under the same assumptions of Lemma 5.3 and Lemma 5.5, if* $\varepsilon > 0$ *then*

$$\frac{1}{2} \, W_2^2(\tilde{\mu}_1, \tilde{\mu}_0) - \int_0^1 \ddot{\eta}(s) \, E_\varepsilon(\tilde{\mu}_s) \, \mathrm{d}s + \dot{\eta}(1) \, E_\varepsilon(\tilde{\mu}_1) \leq \dot{\eta}(0) \, E_\varepsilon(\tilde{\mu}_0) + \frac{1}{2 \, \mathrm{I}_{-2,\eta}(1)} \int_0^1 |\dot{\mu}_s|^2 \, \mathrm{d}s,$$

where

$$\mathrm{I}_{p,\eta}(s) = \int_0^s \mathsf{c}^p(\eta(r)) \, \mathrm{d}r \quad \textit{for all } s \in [0,1] \textit{ and } p \in \mathbb{R},$$

and

$$\vartheta(s) = \frac{\mathrm{I}_{-2,\eta}(s)}{\mathrm{I}_{-2,\eta}(1)} \quad \textit{for all } s \in [0,1]. \tag{5.16}$$

Proof. Let $\varepsilon > 0$ be fixed. On the one hand, recalling (2.29), by Lemma 5.3 we can estimate

$$\frac{\mathrm{d}}{\mathrm{d}s} \int_X \varphi_s \, \mathrm{d}\tilde{\mu}_s \leq -\frac{1}{2} \int_X \Gamma(\varphi_s) \, \mathrm{d}\tilde{\mu}_s - \dot{\eta}(s) \int_X \Gamma(\tilde{f}_s, \varphi_s) \, \mathrm{d}\mathfrak{m} + \dot{\vartheta}(s) \int_X \dot{f}_{\vartheta(s)} \, \mathsf{P}_{\eta(s)} \varphi_s \, \mathrm{d}\mathfrak{m}$$

for $\mathscr{L}^1$-a.e. $s \in (0,1)$. On the other hand, by Lemma 5.5, we can also estimate

$$\dot{\eta}(s)\frac{\mathrm{d}}{\mathrm{d}s}E_\varepsilon(\tilde{\mu}_s) \le -\frac{\dot{\eta}^2(s)}{2}\int_X \Gamma(g_s^\varepsilon)\,\mathrm{d}\tilde{\mu}_s + \dot{\vartheta}(s)\,\dot{\eta}(s)\int_X \dot{f}_{\vartheta(s)}\,\mathsf{P}_{\eta(s)}g_s^\varepsilon\,\mathrm{d}\mathfrak{m}$$

for all $s \in (0,1]$. Hence, we can estimate

$$\frac{\mathrm{d}}{\mathrm{d}s}\int_X \varphi_s\,\mathrm{d}\tilde{\mu}_s + \dot{\eta}(s)\frac{\mathrm{d}}{\mathrm{d}s}E_\varepsilon(\tilde{\mu}_s) \le \dot{\vartheta}(s)\int_X \dot{f}_{\vartheta(s)}\,\mathsf{P}_{\eta(s)}(\varphi_s + \dot{\eta}(s)\,g_s^\varepsilon)\,\mathrm{d}\mathfrak{m}$$
$$-\frac{1}{2}\int_X \left(\Gamma(\varphi_s) + \dot{\eta}^2(s)\,\Gamma(g_s^\varepsilon)\right)\mathrm{d}\tilde{\mu}_s - \dot{\eta}(s)\int_X \Gamma(\tilde{f}_s,\varphi_s)\,\mathrm{d}\mathfrak{m}$$

for $\mathscr{L}^1$-a.e. $s \in (0,1)$. Recalling that $r\,p'_\varepsilon(r) \le 1$ for all $r \ge 0$ by (5.12), by the chain rule (2.47) we can also estimate

$$\int_X \Gamma(g_s^\varepsilon,\varphi_s)\,\mathrm{d}\tilde{\mu}_s = \int_X \tilde{f}_s\,p'_\varepsilon(\tilde{f}_s)\Gamma(\tilde{f}_s,\varphi_s)\,\mathrm{d}\mathfrak{m} \le \int_X \Gamma(\tilde{f}_s,\varphi_s)\,\mathrm{d}\mathfrak{m}$$

for all $s \in (0,1]$, so that

$$\frac{1}{2}\int_X \left(\Gamma(\varphi_s) + \dot{\eta}^2(s)\,\Gamma(g_s^\varepsilon)\right)\mathrm{d}\tilde{\mu}_s + \dot{\eta}(s)\int_X \Gamma(\tilde{f}_s,\varphi_s)\,\mathrm{d}\mathfrak{m}$$
$$\ge \frac{1}{2}\int_X \left(\Gamma(\varphi_s) + \dot{\eta}^2(s)\,\Gamma(g_s^\varepsilon)\right)\mathrm{d}\tilde{\mu}_s + \dot{\eta}(s)\int_X \Gamma(g_s^\varepsilon,\varphi_s)\,\mathrm{d}\tilde{\mu}_s$$
$$= \frac{1}{2}\int_X \Gamma(\varphi_s + \dot{\eta}(s)\,g_s^\varepsilon)\,\mathrm{d}\tilde{\mu}_s$$

for all $s \in (0,1]$. Thus, by (3.9), we can estimate

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s}\int_X \varphi_s\,\mathrm{d}\tilde{\mu}_s + \dot{\eta}(s)\frac{\mathrm{d}}{\mathrm{d}s}E_\varepsilon(\tilde{\mu}_s) &\le \dot{\vartheta}(s)\int_X \dot{f}_{\vartheta(s)}\,\mathsf{P}_{\eta(s)}(\varphi_s + \dot{\eta}(s)\,g_s^\varepsilon)\,\mathrm{d}\mathfrak{m} \\
&\quad -\frac{1}{2}\int_X \Gamma(\varphi_s + \dot{\eta}(s)\,g_s^\varepsilon)\,\mathrm{d}\tilde{\mu}_s \\
&= \dot{\vartheta}(s)\int_X \dot{f}_{\vartheta(s)}\,\mathsf{P}_{\eta(s)}(\varphi_s + \dot{\eta}(s)\,g_s^\varepsilon)\,\mathrm{d}\mathfrak{m} \\
&\quad -\frac{1}{2}\int_X \mathsf{P}_{\eta(s)}\left(\Gamma(\varphi_s + \dot{\eta}(s)\,g_s^\varepsilon)\right)\mathrm{d}\mu_{\vartheta(s)} \\
&\le \dot{\vartheta}(s)\int_X \dot{f}_{\vartheta(s)}\,\mathsf{P}_{\eta(s)}(\varphi_s + \dot{\eta}(s)\,g_s^\varepsilon)\,\mathrm{d}\mathfrak{m} \\
&\quad -\frac{\mathsf{c}^{-2}(\eta(s))}{2}\int_X \Gamma\left(\mathsf{P}_{\eta(s)}(\varphi_s + \dot{\eta}(s)\,g_s^\varepsilon)\right)\mathrm{d}\mu_{\vartheta(s)}
\end{aligned}$$

(5.17)

for $\mathscr{L}^1$-a.e. $s \in (0,1)$. Since $\eta(s) > 0$ for all $s \in (0,1]$, we have

$$\varphi_s + \dot{\eta}(s)\,g_s^\varepsilon \in \mathrm{L}^\infty(X,\mathfrak{m}) \cap \mathrm{W}^{1,2}(X,\mathsf{d},\mathfrak{m})$$

for all $s \in (0,1]$ so that, thanks to heat-smoothing assumption (3.62),

$$\Gamma(\mathsf{P}_{\eta(s)}(\varphi_s + \dot{\eta}(s)g_s^\varepsilon)) = |\mathsf{DP}_{\eta(s)}(\varphi_s + \dot{\eta}(s)g_s^\varepsilon)|^2 \quad \mathfrak{m}\text{-a.e. in } X$$

for all $s \in (0, 1]$. Thus, by (5.1) in Definition 5.1, by Lemma 3.19 and by Young inequality, we can estimate

$$\dot{\vartheta}(s) \int_X \dot{f}_{\vartheta(s)} \, \mathsf{P}_{\eta(s)}(\varphi_s + \dot{\eta}(s) \, g_s^\varepsilon) \, \mathrm{d}\mathfrak{m} = \frac{\mathrm{d}}{\mathrm{d}r} \int_X \mathsf{P}_{\eta(s)}(\varphi_s + \dot{\eta}(s) \, g_s^\varepsilon) \, \mathrm{d}\mu_{\vartheta(r)} \Big|_{r=s}$$

$$\leq |\dot{\vartheta}(s)| \, |\dot{\mu}_{\vartheta(s)}| \left(\int_X |\mathsf{DP}_{\eta(s)}(\varphi_s + \dot{\eta}(s) g_s^\varepsilon)|^2 \, \mathrm{d}\mu_{\vartheta(s)} \right)^{\frac{1}{2}}$$

$$= |\dot{\vartheta}(s)| \, |\dot{\mu}_{\vartheta(s)}| \left(\int_X \Gamma(\mathsf{P}_{\eta(s)}(\varphi_s + \dot{\eta}(s) g_s^\varepsilon)) \, \mathrm{d}\mu_{\vartheta(s)} \right)^{\frac{1}{2}}$$

$$\leq \frac{\mathsf{c}^2(\eta(s))}{2} \, \dot{\vartheta}^2(s) \, |\dot{\mu}_{\vartheta(s)}|^2 + \frac{\mathsf{c}^{-2}(\eta(s))}{2} \int_X \Gamma(\mathsf{P}_{\eta(s)}(\varphi_s + \dot{\eta}(s) g_s^\varepsilon)) \, \mathrm{d}\mu_{\vartheta(s)}$$

(5.18)

for all $s \in (0, 1]$. In conclusion, by combining (5.17) with (5.18), we get

$$\frac{\mathrm{d}}{\mathrm{d}s} \int_X \varphi_s \, \mathrm{d}\tilde{\mu}_s + \dot{\eta}(s) \, \frac{\mathrm{d}}{\mathrm{d}s} \, E_\varepsilon(\tilde{\mu}_s) \leq \frac{\mathsf{c}^2(\eta(s))}{2} \, \dot{\vartheta}^2(s) \, |\dot{\mu}_{\vartheta(s)}|^2 \tag{5.19}$$

for $\mathscr{L}^1$-a.e. $s \in (0, 1)$. We now integrate (5.19) in $s \in [0, 1]$. For the left-hand side of (5.19), we have

$$\int_0^1 \frac{\mathrm{d}}{\mathrm{d}s} \int_X \varphi_s \, \mathrm{d}\tilde{\mu}_s \, \mathrm{d}s + \int_0^1 \dot{\eta}(s) \, \frac{\mathrm{d}}{\mathrm{d}s} \, E_\varepsilon(\tilde{\mu}_s) \, \mathrm{d}s$$

$$= \int_X \varphi_1 \, \mathrm{d}\tilde{\mu}_1 - \int_X \varphi_0 \, \mathrm{d}\tilde{\mu}_0 - \int_0^1 \ddot{\eta}(s) \, E_\varepsilon(\tilde{\mu}_s) \, \mathrm{d}s + \Big(\dot{\eta}(1) E_\varepsilon(\tilde{\mu}_1) - \dot{\eta}(0) E_\varepsilon(\tilde{\mu}_0) \Big). \tag{5.20}$$

For the right-hand side of (5.19), instead, we simply choose $\vartheta \colon [0, 1] \to [0, 1]$ as in (5.16), so that

$$\frac{1}{2} \int_0^1 \dot{\vartheta}^2(s) \, \mathsf{c}^2(\eta(s)) \, |\dot{\mu}_{\vartheta(s)}|^2 \, \mathrm{d}s = \frac{1}{2 \, \mathrm{I}_{-2,\eta}(1)} \int_0^1 |\dot{\mu}_{\vartheta(s)}|^2 \, \dot{\vartheta}(s) \, \mathrm{d}s = \frac{1}{2 \, \mathrm{I}_{-2,\eta}(1)} \int_0^1 |\dot{\mu}_s|^2 \, \mathrm{d}s. \tag{5.21}$$

Combining (5.20) with (5.21), we get

$$\int_X \varphi_1 \, \mathrm{d}\tilde{\mu}_1 - \int_X \varphi_0 \, \mathrm{d}\tilde{\mu}_0 - \int_0^1 \ddot{\eta}(s) \, E_\varepsilon(\tilde{\mu}_s) \, \mathrm{d}s + \dot{\eta}(1) \, E_\varepsilon(\tilde{\mu}_1)$$

$$\leq \dot{\eta}(0) \, E_\varepsilon(\tilde{\mu}_0) + \frac{1}{2 \, \mathrm{I}_{-2,\eta}(1)} \int_0^1 |\dot{\mu}_s|^2 \, \mathrm{d}s$$

whenever $\varepsilon > 0$, and the conclusion follows by taking the supremum on all $\varphi \in \mathrm{Lip}_\star(X)$ thanks to (2.13). □

Remark 5.7 (Errata to the proof of [16, Theorem 4.16]). We warn the reader that there is a typo in the last inequality of the long chain of inequalities in the proof of [16, Theorem 4.16]: in place of $\frac{1}{2}(\dot{\vartheta}_s)e^{-2Kst}|\dot{\varrho}_s|^2$, it should be written $\frac{1}{2}(\dot{\vartheta}_s)e^{-2Kst}|\dot{\varrho}_{\vartheta(s)}|^2$. Unfortunately, this typo induces the authors to make the wrong choice of $\vartheta(s)$ at the beginning of [16, p. 393], making the proofs of [16, Theorem 4.16 and Theorem 4.17] not completely corrected. The reader can easily fix all the computations needed in [16] by adapting the ones performed above in the proof of Theorem 5.6.

5.5. **Admissible groups.** We now focus our attention on some particular admissible metric-measure spaces that we call *admissible groups*.

Definition 5.8 (Admissible group). We say that an admissible metric-measure space $(X, \mathsf{d}, \mathfrak{m})$ is an *admissible group* if:

(i) the metric space (X, d) is locally compact;
(ii) the set X is a *topological group*, i.e. the group operations of *multiplication* $(x, y) \mapsto xy$ and *inversion* $x \mapsto x^{-1}$ are continuous;
(iii) d is *left-invariant*, i.e. $\mathsf{d}(zx, zy) = \mathsf{d}(x, y)$ for all $x, y, z \in X$;
(iv) $\mathfrak{m}$ is a *left-invariant Haar measure*, i.e. $\mathfrak{m}$ is a Radon measure such that $\mathfrak{m}(xE) = \mathfrak{m}(E)$ for all $x \in X$ and all Borel set $E \subset X$;
(v) X is *unimodular*, i.e. $\mathfrak{m}$ is also right-invariant.

For an agile introduction on topological groups and Haar measures, we refer the reader to [78, Section 11.1] and to [79, Chapter 2]. For a more general approach to the subject, see [73, Section 2.7].

Note that, since d is left-invariant, we can write $B_r(x) = xB_r(\mathsf{o})$ for all $x \in X$ and $r > 0$, where $\mathsf{o} \in X$ is the identity element. Since $\mathfrak{m}$ is right invariant, we thus get that $\mathfrak{m}(B_r(x)) = \mathfrak{m}(B_r(\mathsf{o}))$ for all $x \in X$ and $r > 0$, so that if (2.15) is satisfied for one $x_0 \in X$, then it is satisfied (with the same constants $A, B > 0$) for all $x_0 \in X$. Hence, from now on, we assume $x_0 = \mathsf{o}$ in (2.15) for simplicity.

Here and in the rest of the paper, we let $L_x(y) = x^{-1}y$, $x, y \in X$, be the left-translation map. The following result is a simple consequence of Definition 5.8 and the definitions given in Section 2. We leave its proof to the interested reader.

Proposition 5.9 (Invariance properties of metric-measure objects). *Let $(X, \mathsf{d}, \mathfrak{m})$ be an admissible group. Let $p \in [1, +\infty]$ and $x \in X$ be fixed. The following hold.*

(i) *If $p < +\infty$, then $W_p((L_x)_\sharp \mu, (L_x)_\sharp \nu) = W_p(\mu, \nu)$ for all $\mu, \nu \in \mathscr{P}_p(X)$.*
(ii) *If $f \in \mathrm{Lip}(X)$, then also $f \circ L_x \in \mathrm{Lip}(X)$, with $|\mathrm{D}(f \circ L_x)| = |\mathrm{D}f| \circ L_x$ and $|\mathrm{D}^*(f \circ L_x)| = |\mathrm{D}^*f| \circ L_x$.*
(iii) *If $f \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$, then also $f \circ L_x \in \mathrm{W}^{1,2}(X, \mathsf{d}, \mathfrak{m})$, with $\mathsf{Ch}(f \circ L_x) = \mathsf{Ch}(f)$ and $|\mathrm{D}(f \circ L_x)|_w = |\mathrm{D}f|_w \circ L_x$.*
(iv) *If $f \in \mathrm{Dom}(\Delta_{\mathsf{d},\mathfrak{m}})$, then also $f \circ L_x \in \mathrm{Dom}(\Delta_{\mathsf{d},\mathfrak{m}})$ with $\Delta_{\mathsf{d},\mathfrak{m}}(f \circ L_x) = (\Delta_{\mathsf{d},\mathfrak{m}}f) \circ L_x$.*
(v) *If $f \in \mathrm{L}^p(X, \mathfrak{m})$, then $\mathsf{P}_t(f \circ L_x) = (\mathsf{P}_t f) \circ L_x$ for all $t > 0$.*

5.6. **Convolution.** We let

$$(f \star g)(x) = \int_X f(xy^{-1})\, g(y)\, \mathrm{d}\mathfrak{m}(y), \quad x \in X,$$

be the *group convolution* of $f, g \in \mathrm{L}^1(X, \mathfrak{m})$, and we use the notation $f * g$ to denote the usual convolution in $\mathbb{R}^n$ of $f, g \in \mathrm{L}^1(\mathbb{R}^n, \mathscr{L}^n)$. Since X is unimodular, by a simple change of variables we can also write

$$(f \star g)(x) = \int_X f(y)\, g(y^{-1}x)\, \mathrm{d}\mathfrak{m}(y), \quad x \in X.$$

Thus, accordingly, for $f \in \mathrm{L}^1(X, \mathfrak{m})$ and $\mu \in \mathscr{P}(X)$ we write

$$(f \star \mu)(x) = \int_X f(xy^{-1})\, \mathrm{d}\mu(y), \quad x \in X.$$

For an account on the elementary properties of convolution in locally compact groups, we refer the reader to [79, Section 2.5] (in particular, recall Young inequality in [79, Proposition 2.40]).

The following result completes the information provided by Lemma 3.24.

Lemma 5.10 ($(\tilde{\mathsf{P}}_t)_{t>0}$ as a right-convolution). *Assume $(X, \mathsf{d}, \mathfrak{m})$ is an admissible group satisfying $\mathsf{BE}_w(\mathsf{c}, \infty)$. If $t > 0$, then*

$$\mathsf{p}_t[x](y) = \mathsf{p}_t[\mathsf{o}](y^{-1}x) \quad \text{for } \mathfrak{m}\text{-a.e. } x, y \in X.$$

Consequently, if $t > 0$, then we can write

$$\tilde{\mathsf{P}}_t f(x) = \int_X f(y)\, \mathsf{p}_t[\mathsf{o}](y^{-1}x)\, \mathrm{d}\mathfrak{m}(y) = (f \star \mathsf{p}_t[\mathsf{o}])(x)$$

for $\mathfrak{m}$-a.e. $x \in X$ and for all one-side bounded measurable functions $f \colon X \to \overline{\mathbb{R}}$.

Proof. Let $t > 0$ be fixed. We start by claiming that

$$\mathsf{H}_t((L_x)_\sharp \mu) = (L_x)_\sharp(\mathsf{H}_t \mu) \tag{5.22}$$

for all $\mu \in \mathscr{P}_2(X)$ and $x \in X$. Indeed, if $\mu \ll \mathfrak{m}$, then by (3.21) claim (5.22) is nothing but Proposition 5.9(v). Since $\mathscr{P}_2^{\mathrm{ac}}(X)$ is W_2-dense in $\mathscr{P}_2(X)$, claim (5.22) follows from Proposition 5.9(i) and Theorem 3.16(ii) by a simple approximation argument. Thanks to claim (5.22), we can compute

$$\begin{aligned}
\int_X f(y)\, \mathsf{p}_t[x](y)\, \mathrm{d}\mathfrak{m}(y) &= \int_X f(y)\, \mathrm{d}\mathsf{H}_t \delta_x(y) \\
&= \int_X f(y)\, \mathrm{d}\mathsf{H}_t((L_{x^{-1}})_\sharp \delta_\mathsf{o})(y) \\
&= \int_X f(y)\, \mathrm{d}(L_{x^{-1}})_\sharp(\mathsf{H}_t \delta_\mathsf{o})(y) \\
&= \int_X f(xy)\, \mathsf{p}_t[\mathsf{o}](y)\, \mathrm{d}\mathfrak{m}(y) \\
&= \int_X f(y)\, \mathsf{p}_t[\mathsf{o}](x^{-1}y)\, \mathrm{d}\mathfrak{m}(y)
\end{aligned}$$

for all $f \in \mathrm{L}^\infty(X, \mathfrak{m})$ and $x \in X$. Thus $\mathsf{p}_t[x](y) = \mathsf{p}_t[\mathsf{o}](x^{-1}y)$ for all $x \in X$ and $\mathfrak{m}$-a.e. $y \in X$, and the conclusion follows by Lemma 3.24(ii). $\qquad\square$

According to Lemma 5.10, we thus simply write $\mathsf{p}_t[\mathsf{o}] = \mathsf{p}_t$ for all $t > 0$ and we call $(\mathsf{p}_t)_{t>0}$ the (*metric-measure*) *heat kernel* of the (pointwise version of the) heat flow.

Remark 5.11 (Application of (4.10))**.** Assume $(X, \mathsf{d}, \mathfrak{m})$ is an admissible group satisfying $\mathsf{BE}_w(\mathsf{c}, \infty)$ with $\mathfrak{m} \in \mathscr{P}(X)$. From inequality (4.10) in Corollary 4.7 we immediately have

$$\mathsf{p}_{2t}(x) \geq \exp\left(-\frac{\mathsf{d}^2(x, \mathsf{o})}{4\, \mathrm{I}_{-2}(t)} \right) \tag{5.23}$$

for all $t > 0$ and $\mathfrak{m}$-a.e. $x \in X$. Inequality (5.23) applies in particular to the (sub-Riemmanian) $\mathbb{SU}(2)$ group, see Section 5.9.2 for the precise definition. Up to our knowledge, inequality (5.23) provides a new lower bound on the heat kernel in $\mathbb{SU}(2)$.

5.7. **Approximation by regular curves in admissible groups.** Let $(X, \mathsf{d}, \mathfrak{m})$ be an admissible group. We say that $\varrho \in \mathrm{L}^1(X, \mathfrak{m})$ is a *convolution kernel* if it is non-negative, renormalized, symmetric and has bounded support, i.e.

$$\varrho \geq 0, \quad \int_X \varrho\, \mathrm{d}\mathfrak{m} = 1, \quad \varrho(x^{-1}) = \varrho(x) \text{ for all } x \in X, \quad \operatorname{supp} \varrho \text{ is bounded.} \tag{5.24}$$

Since d is left-invariant, d-balls centered at $\mathsf{o} \in X$ are symmetric, in the sense that

$$x \in B_r(\mathsf{o}) \iff x^{-1} \in B_r(\mathsf{o})$$

whenever $x \in X$ and $r > 0$. Thus, for all $r > 0$, the function $\varrho_r \in \mathrm{L}^1(X, \mathfrak{m}) \cap \mathrm{L}^\infty(X, \mathfrak{m})$ defined by

$$\varrho_r(x) = \frac{\chi_{B_r(\mathsf{o})}(x)}{\mathfrak{m}(B_r(\mathsf{o}))}, \quad x \in X, \tag{5.25}$$

is a convolution kernel (and also an *approximate identity* as $r \to 0^+$, see [79, Proposition 2.44]).

The following result provides a simple but useful relation between test plans and convolution.

Lemma 5.12 (Convolution and plans). *Let $(X, \mathsf{d}, \mathfrak{m})$ be an admissible group and let $\varrho \in \mathrm{L}^1(X, \mathfrak{m})$ be as in (5.24). Let $\mu_1, \mu_2 \in \mathscr{P}(X)$ and define $\tilde{\mu}_1, \tilde{\mu}_2 \in \mathscr{P}(X)$ as $\tilde{\mu}_1 = (\varrho \star \mu_1)\mathfrak{m}$ and $\tilde{\mu}_2 = (\varrho \star \mu_2)\mathfrak{m}$. If $\pi \in \mathsf{Plan}(\mu_1, \mu_2)$, then the measure $\tilde{\pi}$ given by*

$$\int_{X \times X} \varphi(x_1, x_2) \, \mathrm{d}\tilde{\pi}(x_1, x_2) = \int_{X \times X} \int_X \varrho(y) \, \varphi(yx_1, yx_2) \, \mathrm{d}\mathfrak{m}(y) \, \mathrm{d}\pi(x_1, x_2) \qquad (5.26)$$

for all $\varphi \in \mathrm{C}_b(X \times X)$ is such that $\tilde{\pi} \in \mathsf{Plan}(\tilde{\mu}_1, \tilde{\mu}_2)$.

Proof. Note that $\tilde{\pi} \in \mathscr{P}(X \times X)$, since

$$\int_{X \times X} \mathrm{d}\tilde{\pi}(x_1, x_2) = \int_{X \times X} \int_X \varrho(y) \, \mathrm{d}\mathfrak{m}(y) \, \mathrm{d}\pi(x_1, x_2) = \int_{X \times X} \mathrm{d}\pi(x_1, x_2)$$

by (5.24) and (5.26). Let us now prove that $(p_i)_\sharp \tilde{\pi} = \mu_i$, $i = 1, 2$. Let $\varphi \in \mathrm{C}_b(X)$ and set $\psi_i = \varphi \circ p_i \in \mathrm{C}_b(X \times X)$, $i = 1, 2$. By (5.26) and recalling that X is unimodular, we can write

$$\begin{aligned}
\int_X \varphi \, \mathrm{d}(p_i)_\sharp \tilde{\pi} &= \int_{X \times X} \psi_i \, \mathrm{d}\tilde{\pi} \\
&= \int_{X \times X} \int_X \varrho(y) \, \psi_i(yx_1, yx_2) \, \mathrm{d}\mathfrak{m}(y) \, \mathrm{d}\pi(x_1, x_2) \\
&= \int_X \int_X \varrho(y) \, \varphi(yx) \, \mathrm{d}\mathfrak{m}(y) \, \mathrm{d}\mu_i(x) \\
&= \int_X \int_X \varrho(zx^{-1}) \, \varphi(z) \, \mathrm{d}\mathfrak{m}(z) \, \mathrm{d}\mu_i(x) \\
&= \int_X \varphi(z) \int_X \varrho(zx^{-1}) \, \mathrm{d}\mu_i(x) \, \mathrm{d}\mathfrak{m}(z) \\
&= \int_X (\varrho \star \mu_i)(z) \, \varphi(z) \, \mathrm{d}\mathfrak{m}(z) \\
&= \int_X \varphi \, \mathrm{d}\tilde{\mu}_i
\end{aligned}$$

thanks to Fubini Theorem, concluding the proof. $\qquad \square$

A fundamental consequence of Lemma 5.12 is the following estimate on the Wasserstein velocity of left-convoluted curves of measures.

Lemma 5.13 (Convolution and W_q-velocity). *Let $(X, \mathsf{d}, \mathfrak{m})$ be an admissible group and let $\varrho \in \mathrm{L}^1(X, \mathfrak{m})$ be as in (5.24). Let $p, q \in [1, +\infty)$ and let $I \subset \mathbb{R}$ be an interval. If $s \mapsto \mu_s \in \mathrm{AC}^q(I; \mathscr{P}_p(X))$, then also $s \mapsto \tilde{\mu}_s \in \mathrm{AC}^q(I; \mathscr{P}_p(X))$, where $\tilde{\mu}_s = (\varrho \star \mu_s)\mathfrak{m}$ for all $s \in I$, with $|\dot{\tilde{\mu}}_s| \le |\dot{\mu}_s|$ for $\mathscr{L}^1$-a.e. $s \in I$.*

Proof. Since X is unimodular and d is left-invariant, by Tonelli Theorem we can estimate

$$\begin{aligned}
\int_X \mathsf{d}^p(x, \mathsf{o}) \, \mathrm{d}\tilde{\mu}_s(x) &= \int_X \int_X \varrho(xy^{-1}) \, \mathsf{d}^p(x, \mathsf{o}) \, \mathrm{d}\mu_s(y) \, \mathrm{d}\mathfrak{m}(x) \\
&= \int_X \int_X \varrho(xy^{-1}) \, \mathsf{d}^p(x, \mathsf{o}) \, \mathrm{d}\mathfrak{m}(x) \, \mathrm{d}\mu_s(y) \\
&= \int_X \int_X \varrho(z) \, \mathsf{d}^p(zy, \mathsf{o}) \, \mathrm{d}\mathfrak{m}(z) \, \mathrm{d}\mu_s(y)
\end{aligned}$$

$$\leq 2^{p-1} \int_X \int_X \varrho(z) \left(\mathsf{d}^p(zy, z) + \mathsf{d}^p(z, \mathsf{o}) \right) \mathrm{d}\mathfrak{m}(z) \, \mathrm{d}\mu_s(y)$$
$$= 2^{p-1} \left(\int_X \mathsf{d}^p(y, \mathsf{o}) \, \mathrm{d}\mu_s(y) + \int_X \mathsf{d}^p(z, \mathsf{o}) \, \varrho(z) \, \mathrm{d}\mathfrak{m}(z) \right),$$

proving that $\tilde{\mu}_s \in \mathscr{P}_p(X)$ for all $s \in I$. Now for all $k \in \mathbb{N}$ let $\varphi_k \in \mathrm{C}_b(X \times X)$ be defined as $\varphi_k(x, y) = \mathsf{d}^p(x, y) \wedge k$ for all $x, y \in X$. Let $s_0, s_1 \in I$ and let $\pi_{s_0, s_1} \in \mathrm{OptPlan}(\mu_{s_0}, \mu_{s_1})$ be an optimal plan between μ_{s_0} and μ_{s_1}. Let $\tilde{\pi}_{s_0, s_1} \in \mathsf{Plan}(\tilde{\mu}_{s_0}, \tilde{\mu}_{s_1})$ be given by Lemma 5.12 accordingly. Since $\varphi_k(zx, zy) = \varphi_k(x, y)$ for all $x, y, z \in X$ and $k \in \mathbb{N}$, we can estimate

$$\int_{X \times X} \varphi_k(x, y) \, \mathrm{d}\tilde{\pi}_{s_0, s_1}(x, y) = \int_{X \times X} \int_X \varrho(z) \, \varphi_k(zx, zy) \, \mathrm{d}\mathfrak{m}(z) \, \mathrm{d}\pi_{s_0, s_1}(x, y)$$
$$= \int_{X \times X} \int_X \varrho(z) \, \varphi_k(x, y) \, \mathrm{d}\mathfrak{m}(z) \, \mathrm{d}\pi_{s_0, s_1}(x, y)$$
$$\leq \int_{X \times X} \mathsf{d}^p(x, y) \, \mathrm{d}\pi_{s_0, s_1}(x, y)$$
$$= W_p^p(\mu_{s_0}, \mu_{s_1}).$$

By the Monotone Convergence Theorem, we can pass to the limit as $k \to +\infty$ and get

$$W_p(\tilde{\mu}_{s_0}, \tilde{\mu}_{s_1}) \leq W_p(\mu_{s_0}, \mu_{s_1})$$

whenever $s_0, s_1 \in I$, concluding the proof. $\qquad\square$

Remark 5.14 (Right-convoluted measures and velocity)**.** It is not difficult to see that a statement similar to that of Lemma 5.12 holds for the right-convoluted measures $\hat{\mu}_1 = (\mu_1 \star \varrho)\mathfrak{m}$ and $\hat{\mu}_2 = (\mu_2 \star \varrho)\mathfrak{m}$. However, since d is not necessarily right-invariant, it is not clear how to prove a statement similar to that of Lemma 5.13 for the right-convoluted curve $s \mapsto \hat{\mu}_s = (\mu_s \star \varrho)\mathfrak{m}$. Since in admissible groups the heat semigroup acts on measures as the right-convolution with the heat kernel as seen in Lemma 5.10, the lack of an estimate on the Wasserstein velocity of right-convoluted curves is a central obstacle for the use of the heat-regularization techniques (which, in the standard $\mathsf{BE}(K, N)$ framework, inevitably rely on the crucial fact that $\underline{\mathsf{c}}(0^+) = 1$, recall the discussion made in Section 5.1). This also explains why, in Theorem 5.15 below, we need to assume that the ambient space is an admissible group and rely on left-convolution of measures.

We now prove the following crucial approximation result. The line of the proof is close to that of [20, Theorem 4.8]. For the approximation of curves under the standard $\mathsf{BE}(K, N)$ condition, we refer the reader to [16, Proposition 4.11] and [72, Lemma 4.11].

Theorem 5.15 (Approximation by strongly regular curves in $\mathscr{P}_2(X)$)**.** *Let* $(X, \mathsf{d}, \mathfrak{m})$ *be an admissible group. If* $s \mapsto \mu_s \in \mathrm{AC}^2([0, 1]; \mathscr{P}_2(X))$, *then there exist strongly regular curves* $s \mapsto \mu_s^n \in \mathrm{AC}^2([0, 1]; \mathscr{P}_2(X))$, $n \in \mathbb{N}$, *in the sense of Definition 5.1, such that:*

(i) $\mu_s^n \xrightarrow{W_2} \mu_s$ *as* $n \to +\infty$ *for all* $s \in [0, 1]$;

(ii) $\limsup\limits_{n \to +\infty} \int_0^1 |\dot{\mu}_s^n|^2 \, \mathrm{d}s \leq \int_0^1 |\dot{\mu}_s|^2 \, \mathrm{d}s$;

(iii) $\lim\limits_{n \to +\infty} \mathsf{Ent}_\mathfrak{m}(\mathsf{H}_t \mu_s^n) = \mathsf{Ent}_\mathfrak{m}(\mathsf{H}_t \mu_s)$ *for all* $s \in [0, 1]$ *and* $t \geq 0$.

Proof. We divide the proof in four steps.

Step 1: time-extension to $\mathbb{R}$. We define $\mathbb{R} \ni s \mapsto \nu_s \in \mathscr{P}_2(X)$ by extending $[0, 1] \ni s \mapsto \mu_s \in \mathscr{P}_2(X)$ by continuity with constant values in $(-\infty, 0) \cup (1, +\infty)$. Clearly, we have that $s \mapsto \nu_s \in \mathrm{AC}^2(\mathbb{R}; \mathscr{P}_2(X))$.

Step 2: smoothing in the space variable. For all $r > 0$, let $\varrho_r \in \mathrm{L}^1(X, \mathfrak{m}) \cap \mathrm{L}^\infty(X, \mathfrak{m})$ be defined as in (5.25). We thus define $\nu_s^r = f_s^r \mathfrak{m}$, where

$$f_s^r(x) = (\varrho_r \star \nu_s)(x) = \int_X \varrho_r(xy^{-1}) \, \mathrm{d}\nu_s(y), \quad x \in X,$$

for all $s \in \mathbb{R}$ and $r > 0$. By Lemma 5.13, we have $s \mapsto \nu_s^r \in \mathrm{AC}^2(\mathbb{R}; \mathscr{P}_2(X))$, with $|\dot{\nu}_s^r| \leq |\dot{\nu}_s|$ for $\mathscr{L}^1$-a.e. $s \in \mathbb{R}$, for all $r > 0$. Since the family $(\varrho_r)_{r>0}$ is a symmetric approximation of the identity, we have

$$\lim_{r \to 0^+} \int_X \varphi \, \mathrm{d}\nu_s^r = \lim_{r \to 0^+} \int_X (\varrho_r \star \varphi) \, \mathrm{d}\nu_s = \int_X \varphi \, \mathrm{d}\nu_s \tag{5.27}$$

for all $\varphi \in \mathrm{C}_b(X)$ by the Dominated Convergence Theorem for all $s \in \mathbb{R}$, so that $\nu_s^r \rightharpoonup \nu_s$ as $r \to 0^+$ for all $s \in \mathbb{R}$. In addition, we can write

$$\int_X \mathsf{d}^2(x, \mathsf{o}) \, \mathrm{d}\nu_s^r(x) = \int_X (\varrho_r \star \mathsf{d}^2(\cdot, \mathsf{o}))(x) \, \mathrm{d}\nu_s(x)$$
$$= \int_X \int_X \varrho_r(y) \, \mathsf{d}^2(x, y) \, \mathrm{d}\mathfrak{m}(y) \, \mathrm{d}\nu_s(x)$$
$$= \int_X \fint_{B_r(\mathsf{o})} \mathsf{d}^2(x, y) \, \mathrm{d}\mathfrak{m}(y) \, \mathrm{d}\nu_s(x)$$

for all $s \in \mathbb{R}$ and $r > 0$, so that

$$\left| \int_X \mathsf{d}^2(x, \mathsf{o}) \, \mathrm{d}\nu_s^r(x) - \int_X \mathsf{d}^2(x, \mathsf{o}) \, \mathrm{d}\nu_s(x) \right| \leq \int_X \fint_{B_r(\mathsf{o})} \left| \mathsf{d}^2(x, y) - \mathsf{d}^2(x, \mathsf{o}) \right| \, \mathrm{d}\mathfrak{m}(y) \, \mathrm{d}\nu_s(x)$$
$$\leq \int_X r(r + 2\mathsf{d}(x, \mathsf{o})) \, \mathrm{d}\nu_s(x)$$

for all $s \in \mathbb{R}$ and $r > 0$. Hence

$$\lim_{r \to 0^+} \int_X \mathsf{d}^2(x, \mathsf{o}) \, \mathrm{d}\nu_s^r(x) = \int_X \mathsf{d}^2(x, \mathsf{o}) \, \mathrm{d}\nu_s(x) \tag{5.28}$$

for all $s \in \mathbb{R}$. Consequently, from (5.27) and (5.28) we infer that $\nu_s^r \xrightarrow{W_2} \nu_s$ for all $s \in \mathbb{R}$ and, in particular,

$$\liminf_{r \to 0^+} \mathsf{Ent}_\mathfrak{m}(\nu_s^r) \geq \mathsf{Ent}_\mathfrak{m}(\nu_s)$$

for all $s \in \mathbb{R}$. We can also write

$$\nu_s^r = \int_X (L_y)_\sharp \nu_s \, \varrho_r(y) \, \mathrm{d}\mathfrak{m}(y)$$

for all $s \in \mathbb{R}$, where $L_y(x) = y^{-1}x$, $x, y \in X$, denotes the left-translation map. If $\mathfrak{n} = e^{-c\,\mathsf{d}^2(\cdot, \mathsf{o})} \mathfrak{m}$ is as in (2.18), then by Jensen inequality we can estimate

$$\mathsf{Ent}_\mathfrak{n}(\nu_s^r) \leq \int_X \mathsf{Ent}_\mathfrak{n}((L_y)_\sharp \nu_s) \, \varrho_r(y) \, \mathrm{d}\mathfrak{m}(y)$$

for all $s \in \mathbb{R}$, so that by (2.17) and (2.18) we can write

$$\mathsf{Ent}_\mathfrak{n}((L_y)_\sharp \nu_s) = \mathsf{Ent}_{\mathfrak{n}_{y^{-1}}}(\nu_s) = \mathsf{Ent}_\mathfrak{m}(\nu_s) + c \int_X \mathsf{d}^2(yx, \mathsf{o}) \, \mathrm{d}\nu_s(x)$$

for all $y \in X$ and $s \in \mathbb{R}$, where $\mathfrak{n}_y = (L_y)_\sharp \mathfrak{n} = e^{-c\,\mathsf{d}^2(\cdot, y)} \mathfrak{m}$. Since

$$\int_X \int_X \mathsf{d}^2(yx, \mathsf{o}) \, \mathrm{d}\nu_s(x) \, \varrho_r(y) \, \mathrm{d}\mathfrak{m}(y) = \int_X \int_X \mathsf{d}^2(x, y^{-1}) \, \varrho_r(y) \, \mathrm{d}\mathfrak{m}(y) \, \mathrm{d}\nu_s(x)$$
$$= \int_X \int_X \mathsf{d}^2(x, y) \, \varrho_r(y) \, \mathrm{d}\mathfrak{m}(y) \, \mathrm{d}\nu_s(x)$$

$$= \int_X \mathsf{d}^2(x, \mathsf{o}) \, \mathrm{d}\nu_s^r(x)$$

by the symmetry of ϱ_r, again by (2.18) we can estimate

$$\begin{aligned}
\mathsf{Ent}_{\mathfrak{m}}(\nu_s^r) &= \mathsf{Ent}_\nu(\nu_s^r) - c \int_X \mathsf{d}^2(x, \mathsf{o}) \, \mathrm{d}\nu_s^r(x) \\
&\leq \int_X \left(\mathsf{Ent}_{\mathfrak{m}}(\nu_s) + c \int_X \mathsf{d}^2(yx, \mathsf{o}) \, \mathrm{d}\nu_s(x) \right) \varrho_r(y) \, \mathrm{d}\mathfrak{m}(y) - c \int_X \mathsf{d}^2(x, \mathsf{o}) \, \mathrm{d}\nu_s^r(x) \\
&= \mathsf{Ent}_{\mathfrak{m}}(\nu_s) + c \left(\int_X \int_X \mathsf{d}^2(x, y) \, \varrho_r(y) \, \mathrm{d}\mathfrak{m}(y) \, \mathrm{d}\nu_s(x) - \int_X \mathsf{d}^2(x, \mathsf{o}) \, \mathrm{d}\nu_s^r(x) \right) \\
&= \mathsf{Ent}_{\mathfrak{m}}(\nu_s)
\end{aligned}$$

for all $s \in \mathbb{R}$ and $r > 0$, so that

$$\lim_{r \to 0^+} \mathsf{Ent}_{\mathfrak{m}}(\nu_s^r) = \mathsf{Ent}_{\mathfrak{m}}(\nu_s).$$

Step 3: smoothing in the time variable. Now let $r > 0$ be fixed. Let $\zeta \colon \mathbb{R} \to \mathbb{R}$ be a symmetric smooth mollifier in $\mathbb{R}$, i.e.

$$\zeta \in \mathrm{C}_c^\infty(\mathbb{R}), \quad \mathrm{supp}\,\zeta \subset [-1, 1], \quad 0 \leq \zeta \leq 1, \quad \int_{\mathbb{R}} \zeta \, \mathrm{d}\tau = 1.$$

We define $\zeta_j(\tau) = j\,\zeta(j\tau)$ for all $\tau \in \mathbb{R}$ and $j \in \mathbb{N}$ and consider $\nu_s^{j,r} = f_s^{j,r} \mathfrak{m}$ with

$$f_s^{j,r} = (\zeta_j * f_\cdot^r)(s) = \int_{\mathbb{R}} f_\tau^r \, \zeta_j(\tau - s) \, \mathrm{d}\tau$$

for all $s \in \mathbb{R}$ and all $j \in \mathbb{N}$. If $s, s' \in \mathbb{R}$ and $\pi_{s,s'}^r \in \mathsf{Plan}(\nu_s^r, \nu_{s'}^r)$, then the measure $\pi_s^{j,r} \in \mathscr{P}(X \times X)$ given by

$$\int_{X \times X} \varphi(x, y) \, \mathrm{d}\pi_s^{j,r}(x, y) = \int_{\mathbb{R}} \zeta_j(s - \tau) \int_{X \times X} \varphi(x, y) \, \mathrm{d}\pi_{s,\tau}^r(x, y) \, \mathrm{d}\tau,$$

for any $\varphi \in \mathrm{C}_b(X \times X)$, satisfies $\pi_s^{j,r} \in \mathsf{Plan}(\nu_s^r, \nu_s^{j,r})$. Thus, by the convexity properties of the squared 2-Wasserstein distance and Jensen inequality, we get

$$W_2^2(\nu_s^{j,r}, \nu_s^r) \leq \int_{\mathbb{R}} \zeta_j(s - \tau) \, W_2^2(\nu_s^r, \nu_\tau^r) \, \mathrm{d}\tau$$

for all $s \in \mathbb{R}$ and $j \in \mathbb{N}$, so that $\nu_s^{j,r} \xrightarrow{W_2} \nu_s^r$ as $j \to +\infty$ and, in particular,

$$\liminf_{j \to +\infty} \mathsf{Ent}_{\mathfrak{m}}(\nu_s^{j,r}) \geq \mathsf{Ent}_{\mathfrak{m}}(\nu_s^r)$$

for all $s \in \mathbb{R}$. In a similar fashion, we can estimate

$$W_2^2(\nu_s^{j,r}, \nu_{s'}^{j,r}) = W_2^2\left(\int_{\mathbb{R}} \nu_{s-\tau}^r \, \zeta_j(\tau) \, \mathrm{d}\tau, \int_{\mathbb{R}} \nu_{s'-\tau}^r \, \zeta_j(\tau) \, \mathrm{d}\tau \right) \leq \int_{\mathbb{R}} W_2^2\left(\nu_{s-\tau}^r, \nu_{s'-\tau}^r \right) \zeta_j(\tau) \, \mathrm{d}\tau$$

for all $s, s' \in \mathbb{R}$ and $j \in \mathbb{N}$, so that $s \mapsto \nu_s^{j,r} \in \mathrm{AC}^2(\mathbb{R}; \mathscr{P}_2(X))$ with $|\dot\nu_s^{j,r}| \leq (\zeta_j * |\dot\nu_\cdot^r|)(s)$ for $\mathscr{L}^1$-a.e. $s \in \mathbb{R}$ and all $j \in \mathbb{N}$. As in Step 2, let $\mathfrak{n} = e^{-c\,\mathsf{d}^2(\cdot, \mathsf{o})} \mathfrak{m}$ be as in (2.18). Since the function $H(u) = u \log u + (1 - u)$, defined for all $u \geq 0$, is convex and non-negative, by Jensen inequality we can estimate

$$\begin{aligned}
\mathsf{Ent}_{\mathfrak{n}}(\nu_s^{j,r}) &= \int_X H\left(f_s^{j,r} \tfrac{\mathrm{d}\mathfrak{m}}{\mathrm{d}\mathfrak{n}} \right) \mathrm{d}\mathfrak{n} \\
&= \int_X H\left((\zeta_j * f_\cdot^r)(s) \tfrac{\mathrm{d}\mathfrak{m}}{\mathrm{d}\mathfrak{n}} \right) \mathrm{d}\mathfrak{n} \\
&= \int_X H\left((\zeta_j * (f_\cdot^r \tfrac{\mathrm{d}\mathfrak{m}}{\mathrm{d}\mathfrak{n}}))(s) \right) \mathrm{d}\mathfrak{n}
\end{aligned}$$

$$\leq \int_X \left(\zeta_j * H \left(f^r \tfrac{\mathrm{d}\mathfrak{m}}{\mathrm{d}\mathfrak{n}} \right) \right) (s) \, \mathrm{d}\mathfrak{n}$$

$$= (\zeta_j * \mathsf{Ent}_\mathfrak{n}(\nu^r_{\cdot}))(s)$$

for all $s \in \mathbb{R}$ and $j \in \mathbb{N}$. Arguing as in Step 2, we immediately deduce that $\mathsf{Ent}_\mathfrak{m}(\nu^{j,r}_s) \leq (\zeta_j * \mathsf{Ent}_\mathfrak{m}(\nu^r_{\cdot}))(s)$ for all $s \in \mathbb{R}$ and $j \in \mathbb{N}$, so that

$$\lim_{j \to +\infty} \mathsf{Ent}_\mathfrak{m}(\nu^{j,r}_s) = \mathsf{Ent}_\mathfrak{m}(\nu^r_s)$$

for all $s \in \mathbb{R}$.

Step 4: time-restriction to $[0,1]$ and conclusion. Define the curve $s \mapsto \mu^{j,k}_s$ as the restriction of the curve $s \mapsto \nu^{j,k}_s$ to the interval $[0,1]$. Note that the regularization map $\mathcal{R}_{j,r}$ sending the original curve $s \mapsto \mu_s$ to the regularized curve $s \mapsto \mu^{j,k}_s$ is linear (with respect to convex combinations) and thus commutes with the dual heat flow map, i.e. $\mathcal{R}_{j,r} \circ \mathsf{H}_t = \mathsf{H}_t \circ \mathcal{R}_{j,r}$ for all $t \geq 0$. Since by construction

$$\lim_{j,r} \mathsf{Ent}_\mathfrak{m}(\mu^{j,r}_s) = \mathsf{Ent}_\mathfrak{m}(\mu_s)$$

for all $s \in [0,1]$, we thus have

$$\limsup_{j,r} \mathsf{Ent}_\mathfrak{m}(\mathsf{H}_t \mu^{j,r}_s) = \limsup_{j,r} \mathsf{Ent}_\mathfrak{m}(\mathcal{R}^{j,r}(\mathsf{H}_t \mu_s)) = \mathsf{Ent}_\mathfrak{m}(\mathsf{H}_t \mu_s) \leq \mathsf{Ent}_\mathfrak{m}(\mu_s)$$

for all $t \geq 0$ and $s \in [0,1]$ by Lemma 4.1(i). The conclusion thus follows by a standard diagonalization argument. $\qquad\square$

5.8. **Entropic inequalities in admissible groups.** We can now state and prove the main result of this paper. We refer the reader to [16, Theorem 4.17] and to [72, Theorem 4.19] for the analogous results in the standard $\mathsf{BE}(K, N)$ framework.

Theorem 5.16 (Entropic inequalities). *Let $(X, \mathsf{d}, \mathfrak{m})$ be an admissible group satisfying the heat-smoothing property as in Definition 3.27. The following are equivalent.*

(i) $(X, \mathsf{d}, \mathfrak{m})$ satisfies $\mathsf{BE}_w(\mathsf{c}, \infty)$.

(ii) The dual heat semigroup $(\mathsf{H}_t)_{t \geq 0}$ in (3.21) satisfies

$$\frac{1}{2} W_2^2(\mathsf{H}_{t_1} \mu_1, \mathsf{H}_{t_0} \mu_0) - \frac{1}{2\,\mathrm{R}(t_0, t_1)} W_2^2(\mu_1, \mu_0) \leq (t_1 - t_0) \Big(\mathsf{Ent}_\mathfrak{m}(\mathsf{H}_{t_0} \mu_0) - \mathsf{Ent}_\mathfrak{m}(\mathsf{H}_{t_1} \mu_1) \Big) \tag{5.29}$$

for all $\mu_0 \in \mathrm{Dom}(\mathsf{Ent})$, $\mu_1 \in \mathscr{P}_2(X)$ and $0 \leq t_0 \leq t_1$, with also $\mu_1 \in \mathrm{Dom}(\mathsf{Ent}_\mathfrak{m})$ in the particular case $t_1 = t_0 = 0$, where

$$\mathrm{R}(t_0, t_1) = \int_0^1 \mathsf{c}^{-2}((1-s)t_0 + st_1) \, \mathrm{d}s. \tag{5.30}$$

(iii) The dual heat semigroup $(\mathsf{H}_t)_{t \geq 0}$ in (3.21) uniquely extends to a map on $\mathscr{P}_2(X)$ satisfying (3.36) and such that

$$\mathsf{Ent}_\mathfrak{m}(\mathsf{H}_{t+h} \mu_s) \leq (1-s) \, \mathsf{Ent}_m(\mathsf{H}_t \mu_0) + s \, \mathsf{Ent}_\mathfrak{m}(\mathsf{H}_t \mu_1)$$
$$+ \frac{s(1-s)}{2h} \left(\frac{1}{\mathrm{R}(t, t+h)} W_2^2(\mu_0, \mu_1) - W_2^2(\mathsf{H}_t \mu_0, \mathsf{H}_t \mu_1) \right) \tag{5.31}$$

for all $t \geq 0$ and $h > 0$, whenever $s \mapsto \mu_s \in \mathrm{Geo}([0,1]; \mathscr{P}_2(X))$ is a W_2-geodesic joining $\mu_0, \mu_1 \in \mathrm{Dom}(\mathsf{Ent}_\mathfrak{m})$ and R is as in (5.30).

Mimicking the standard framework, thanks to Theorem 5.16 we can introducing the following notation.

Definition 5.17 ($\mathsf{EVI}_w(\mathsf{c})$ and $\mathsf{RCD}_w(\mathsf{c}, \infty)$ conditions)**.** An admissible metric-measure space $(X, \mathsf{d}, \mathfrak{m})$ is said to satisfy the *weak Evolution Variation Inequality* with respect to the function $\mathsf{c} \colon [0, +\infty) \to (0, +\infty)$ in (3.8), $\mathsf{EVI}_w(\mathsf{c})$ for short, if inequality (5.29) in Theorem 5.16(ii) holds. Analogously, $(X, \mathsf{d}, \mathfrak{m})$ is said to satisfy the (*dimension-free*) *weak Riemannian Curvature-Dimension Condition* with respect to the function c, $\mathsf{RCD}_w(\mathsf{c}, \infty)$ for short, if inequality (5.31) in Theorem 5.16(iii) holds.

With this terminology, one can rephrase Theorem 5.16 simply writing that, for an admissible heat-smoothing group $(X, \mathsf{d}, \mathfrak{m})$, it holds

$$\mathsf{BE}_w(\mathsf{c}, \infty) \iff \mathsf{EVI}_w(\mathsf{c}) \iff \mathsf{RCD}_w(\mathsf{c}, \infty).$$

Proof of Theorem 5.16. We prove each implication separately.

Proof of (i)$\Rightarrow$(ii). Let $0 \le t_0 \le t_1$ be fixed. Let $s \mapsto \mu_s \in \mathrm{AC}^2([0, 1]; \mathscr{P}_2(X))$ be a curve joining $\mu_0 \in \mathrm{Dom}(\mathsf{Ent}_\mathfrak{m})$ and $\mu_1 \in \mathscr{P}_2(X)$ (with also $\mu_1 \in \mathrm{Dom}(\mathsf{Ent}_\mathfrak{m})$ in the particular cas $t_1 = t_0 = 0$). We can find strongly regular curves $s \mapsto \mu_s^n \in \mathrm{AC}^2([0, 1]; \mathscr{P}_2(X))$, $n \in \mathbb{N}$, as in Definition 5.1, approximating the curve $s \mapsto \mu_s$ as stated in Theorem 5.15. By Theorem 5.6 applied to each $s \mapsto \mu_s^n$ with $\eta(s) = (1 - s)t_0 + st_1$ for all $s \in [0, 1]$, we get

$$
\begin{aligned}
\frac{1}{2} W_2^2(\mathsf{H}_{t_1}\mu_1^n, \mathsf{H}_{t_0}\mu_0^n) &+ (t_1 - t_0)\, E_\varepsilon(\mathsf{H}_{t_1}\mu_1^n) \\
&\le (t_1 - t_0)\, E_\varepsilon(\mathsf{H}_{t_0}\mu_0^n) + \frac{1}{2\, \mathrm{I}_{-2,\eta}(1)} \int_0^1 |\dot{\mu}_s^n|^2 \, \mathrm{d}s
\end{aligned}
\tag{5.32}
$$

for all $n \in \mathbb{N}$ and $\varepsilon > 0$. On the one hand, recalling (5.10), we have

$$E_\varepsilon(\mathsf{H}_{t_1}\mu_1^n) \ge \mathsf{Ent}_\mathfrak{m}(\mathsf{H}_{t_1}\mu_1^n)$$

for all $\varepsilon > 0$ and $n \in \mathbb{N}$. On the other hand, by (5.13) we have that

$$\lim_{\varepsilon \to 0^+} E_\varepsilon(\mathsf{H}_{t_0}\mu_0^n) = \mathsf{Ent}_\mathfrak{m}(\mathsf{H}_{t_0}\mu_0^n)$$

for all $n \in \mathbb{N}$, since $\mu_0^n = f_0^n \mathfrak{m}$ with $f_0^n \in \mathrm{L}^2(X, \mathfrak{m})$. Thus we can pass to the limit as $\varepsilon \to 0^+$ in (5.32) and get

$$
\begin{aligned}
\frac{1}{2} W_2^2(\mathsf{H}_{t_1}\mu_1^n, \mathsf{H}_{t_0}\mu_0^n) &+ (t_1 - t_0)\, \mathsf{Ent}_\mathfrak{m}(\mathsf{H}_{t_1}\mu_1^n) \\
&\le (t_1 - t_0)\, \mathsf{Ent}_\mathfrak{m}(\mathsf{H}_{t_0}\mu_0^n) + \frac{1}{2\, \mathrm{I}_{-2,\eta}(1)} \int_0^1 |\dot{\mu}_s^n|^2 \, \mathrm{d}s
\end{aligned}
\tag{5.33}
$$

for all $n \in \mathbb{N}$. By (3.36) in Theorem 3.16 and Theorem 5.15(i), we have $\mathsf{H}_{t_i}\mu_i^n \xrightarrow{W_2} \mathsf{H}_{t_i}\mu_i$ as $n \to +\infty$ for $i = 0, 1$, so that

$$\lim_{n \to +\infty} W_2(\mathsf{H}_{t_1}\mu_1^n, \mathsf{H}_{t_0}\mu_0^n) = W_2(\mathsf{H}_{t_1}\mu_1, \mathsf{H}_{t_0}\mu_0). \tag{5.34}$$

Also, by the lower semicontinuity of the entropy, we have

$$\mathsf{Ent}_\mathfrak{m}(\mathsf{H}_{t_1}\mu_1) \le \liminf_{n \to +\infty} \mathsf{Ent}_\mathfrak{m}(\mathsf{H}_{t_1}\mu_1^n). \tag{5.35}$$

Finally, by Theorem 5.15(iii), we can estimate

$$\limsup_{n \to +\infty} \mathsf{Ent}_\mathfrak{m}(\mathsf{H}_{t_0}\mu_0^n) \le \mathsf{Ent}_\mathfrak{m}(\mathsf{H}_{t_0}\mu_0). \tag{5.36}$$

By (5.34), (5.35) and (5.36), we can thus pass to the limit as $n \to +\infty$ in (5.33) getting

$$
\frac{1}{2} W_2^2(\mathsf{H}_{t_1}\mu_1, \mathsf{H}_{t_0}\mu_0) + (t_1 - t_0)\, \mathsf{Ent}_{\mathfrak{m}}(\mathsf{H}_{t_1}\mu_1)
$$
$$
\leq (t_1 - t_0)\, \mathsf{Ent}_{\mathfrak{m}}(\mathsf{H}_{t_0}\mu_0) + \frac{1}{2\, \mathrm{R}(t_0, t_1)} \int_0^1 |\dot{\mu}_s|^2 \, \mathrm{d}s,
\tag{5.37}
$$

so that (ii) follows by minimizing (5.37) with respect to the curves $\mu \in \mathrm{AC}^2([0,1]; \mathscr{P}_2(X))$ joining μ_0 and μ_1.

Proof of (ii) $\Rightarrow$ (i). Choosing $t_0 = t_1 = t > 0$ in (5.29), we get

$$
W_2(\mathsf{H}_t\mu_1, \mathsf{H}_t\mu_0) \leq \mathsf{c}(t)\, W_2(\mu_1, \mu_0)
$$

for all $\mu_0, \mu_1 \in \mathrm{Dom}(\mathsf{Ent}_{\mathfrak{m}})$. This proves the validity of (3.35) in Theorem 3.16(ii) for $\mathscr{D} = \mathrm{Dom}(\mathsf{Ent}_{\mathfrak{m}})$. Since $\mathrm{Dom}(\mathsf{Ent}_{\mathfrak{m}})$ is a W_2-dense subset of $\mathscr{P}_2^{\mathrm{ac}}(X)$, (i) immediately follows by Theorem 3.16.

Proof of (i) $\Rightarrow$ (iii). Since we already know that (i) $\Leftrightarrow$ (ii), by Theorem 3.16 we have that the dual heat semigroup uniquely extends to a map on $\mathscr{P}_2(X)$ satisfying (3.36) and we can thus argue as in the proof of [67, Theorem 3.2]. So let $t \geq 0$ and $h > 0$ be fixed and let $s \mapsto \mu_s \in \mathrm{Geo}([0,1]; \mathscr{P}_2(X))$ be a W_2-geodesic joining $\mu_0, \mu_1 \in \mathrm{Dom}(\mathsf{Ent}_{\mathfrak{m}})$. By (5.29) applied respectively to the couple $\mu_0 \in \mathrm{Dom}(\mathsf{Ent}_{\mathfrak{m}})$, $\mu_s \in \mathscr{P}_2(X)$, and to the couple $\mu_1 \in \mathrm{Dom}(\mathsf{Ent}_{\mathfrak{m}})$, $\mu_s \in \mathscr{P}_2(X)$, both with the choice $t_0 = t$ and $t_1 = t + h$ for all $s \in [0,1]$ (recall that $\mathsf{H}_t\mu \in \mathrm{Dom}(\mathsf{Ent}_{\mathfrak{m}})$ for all $\mu \in \mathscr{P}_2(X)$ and $t > 0$ by Theorem 4.4), we get

$$
\frac{1-s}{2} W_2^2(\mathsf{H}_{t+h}\mu_s, \mathsf{H}_t\mu_0) + \frac{s}{2} W_2^2(\mathsf{H}_{t+h}\mu_s, \mathsf{H}_t\mu_1)
$$
$$
- \frac{1}{2\, \mathrm{R}(t, t+h)} \Big((1-s)\, W_2^2(\mu_s, \mu_0) + s\, W_2^2(\mu_s, \mu_1) \Big)
\tag{5.38}
$$
$$
\leq h\Big((1-s)\, \mathsf{Ent}_{\mathfrak{m}}(\mathsf{H}_t\mu_0) + s\, \mathsf{Ent}_{\mathfrak{m}}(\mathsf{H}_t\mu_1) - \mathsf{Ent}_{\mathfrak{m}}(\mathsf{H}_{t+h}\mu_s) \Big)
$$

for all $s \in [0,1]$. Since $s \mapsto \mu_s$ is a W_2-geodesic, we can estimate

$$
(1-s)\, W_2^2(\mu_s, \mu_0) + s\, W_2^2(\mu_s, \mu_1) = s(1-s)\, W_2^2(\mu_1, \mu_0)
\tag{5.39}
$$

for all $s \in [0,1]$. Thanks to the elementary inequality

$$
(1-s)a^2 + sb^2 \geq s(1-s)(a+b)^2 \quad \text{for all } a, b \in \mathbb{R}, \ s \in [0,1],
$$

by the triangular inequality we can also estimate

$$
(1-s)\, W_2^2(\mathsf{H}_{t+h}\mu_s, \mathsf{H}_t\mu_0) + s\, W_2^2(\mathsf{H}_{t+h}\mu_s, \mathsf{H}_t\mu_1)
$$
$$
\geq s(1-s)\Big(W_2(\mathsf{H}_{t+h}\mu_s, \mathsf{H}_t\mu_0) + W_2(\mathsf{H}_{t+h}\mu_s, \mathsf{H}_t\mu_1) \Big)^2
\tag{5.40}
$$
$$
\geq s(1-s)\, W_2^2(\mathsf{H}_t\mu_1, \mathsf{H}_t\mu_0).
$$

By combining (5.39) and (5.40) with (5.38), we immediately deduce (iii).

Proof of (iii) $\Rightarrow$ (i). Since $(\mathsf{H}_t)_{t \geq 0}$ satisfies (3.36), (i) trivially follows by Theorem 3.16. $\square$

5.9. Application to Carnot groups and the $\mathbb{SU}(2)$ group.
We conclude this section with the application of Theorem 5.16 to Carnot groups and to the $\mathbb{SU}(2)$ group.

5.9.1. *Carnot groups.* We recall that a Carnot group $\mathbb{G}$ is a connected, simply connected and nilpotent Lie group whose Lie algebra $\mathfrak{g}$ of left-invariant vector fields has dimension $n \in \mathbb{N}$ and admits a stratification of step $\kappa \in \mathbb{N}$,

$$\mathfrak{g} = V_1 \oplus V_2 \oplus \cdots \oplus V_\kappa$$

with

$$V_i = [V_1, V_{i-1}] \quad \text{for } i = 1, \ldots, \kappa, \qquad [V_1, V_\kappa] = \{0\}.$$

We set $m_i = \dim(V_i)$ and $h_i = m_1 + \cdots + m_i$ for $i = 1, \ldots, \kappa$, with $h_0 = 0$ and $h_\kappa = n$. We fix an adapted basis of $\mathfrak{g}$, i.e. a basis $X_1, \ldots, X_n$ such that

$$X_{h_{i-1}+1}, \ldots, X_{h_i} \text{ is a basis of } V_i, \qquad i = 1, \ldots, \kappa.$$

Using the exponential coordinates, it is possible to identify $\mathbb{G}$ with $\mathbb{R}^n$ endowed with the group law determined by the Campbell–Hausdorff formula (in particular, the identity $\mathsf{o} \in \mathbb{G}$ corresponds to $0 \in \mathbb{R}^n$ and $x^{-1} = -x$ for $x \in \mathbb{G}$), and it is not restrictive to assume that $X_i(0) = \mathrm{e}_i$ for any $i = 1, \ldots, n$. In particular, by left-invariance, for any $x \in \mathbb{G}$ we get

$$X_i(x) = dL_x \mathrm{e}_i, \qquad i = 1, \ldots, n,$$

where $L_x \colon \mathbb{G} \to \mathbb{G}$ is the left-translation by $x \in \mathbb{G}$. We endow $\mathfrak{g}$ with the left-invariant Riemannian metric $\langle \cdot, \cdot \rangle_\mathbb{G}$ that makes the basis $X_1, \ldots, X_n$ orthonormal. We let $H\mathbb{G} \subset T\mathbb{G}$ be the *horizontal tangent bundle* of the group $\mathbb{G}$, i.e. the left-invariant sub-bundle of the tangent bundle $T\mathbb{G}$ such that $H_0\mathbb{G} = \{X(0) : X \in V_1\}$, and we let

$$\nabla_\mathbb{G} f = \sum_{j=1}^{m_1} (X_j f)\, X_j \in V_1$$

be the *horizontal gradient* of f.

For any $i = 1, \ldots, n$, we define the *degree* $d(i) \in \{1, \ldots, \kappa\}$ of the basis vector field X_i as $d(i) = j$ if and only if $X_i \in V_j$. Using this notation, the one-parameter family of *group dilations* $(\delta_\lambda)_{\lambda \geq 0} \colon \mathbb{G} \to \mathbb{G}$ is given by

$$\delta_\lambda(x) = \delta_\lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda^{d(i)} x_i, \ldots, \lambda^\kappa x_n), \quad \text{for all } x \in \mathbb{G}. \qquad (5.41)$$

The bi-invariant Haar measure of the group $\mathbb{G}$ coincides (up to a multiplicative constant) with the n-dimensional Lebesgue measure $\mathscr{L}^n$ and has the homogeneity property $\mathscr{L}^n(\delta_\lambda(E)) = \lambda^Q \mathscr{L}^n(E)$, where the integer $Q = \sum_{i=1}^\kappa i \dim(V_i)$ is called the *homogeneous dimension* of the group.

We endow the group $\mathbb{G}$ with the canonical *Carnot–Carathéodory metric structure* induced by $H\mathbb{G}$. More precisely, the *Carnot–Carathéodory distance* between $x, y \in \mathbb{G}$ is then defined as

$$\mathsf{d}_{\mathrm{cc}}(x, y) = \inf\left\{ \int_0^1 \|\dot\gamma(t)\|_\mathbb{G}\, dt : \gamma \text{ is horizontal}, \ \gamma(0) = x, \ \gamma(1) = y \right\}. \qquad (5.42)$$

Here and in the following, we say that a Lipschitz curve $\gamma \colon [0,1] \to \mathbb{G}$ is a *horizontal curve* if $\dot\gamma(t) \in H_{\gamma(t)}\mathbb{G}$ for a.e. $t \in [0,1]$. By Chow–Rashevskii's Theorem, the function d_{cc} is in fact a distance, which is also left-invariant and homogeneous with respect to the dilations defined in (5.41), in the sense that

$$\mathsf{d}_{\mathrm{cc}}(zx, zy) = \mathsf{d}_{\mathrm{cc}}(x, y), \quad \mathsf{d}_{\mathrm{cc}}(\delta_\lambda(x), \delta_\lambda(y)) = \lambda\, \mathsf{d}_{\mathrm{cc}}(x, y),$$

for all $x, y, z \in \mathbb{G}$ and $\lambda \geq 0$. The resulting metric space $(\mathbb{G}, \mathsf{d}_{\mathrm{cc}})$ is a complete, separable, locally compact and geodesic space. Note that

$$\mathscr{L}^n(B_r(x)) = c_n r^Q \quad \text{for all } x \in \mathbb{G} \text{ and } r \geq 0,$$

where $c_n = \mathscr{L}^n(B_1(0))$.

The standard *sub-Laplacian operator* is $\Delta_{\mathbb{G}} = \sum_{i=1}^{m_1} X_i^2$. Since the horizontal vector fields $X_1, \ldots, X_{h_1}$ satisfy the *Hörmander condition*, by Hörmander Theorem the *sub-elliptic heat operator* $\partial_t - \Delta_{\mathbb{G}}$ is *hypoelliptic*, meaning that its fundamental solution, the heat kernel $\mathsf{p} \colon (0, +\infty) \times \mathbb{G} \to (0, +\infty)$, is a smooth function. For the main properties of the heat kernel, we refer to [20, Theorem 2.3] and to the references therein. Here we only recall that, given a function $f \in \mathrm{L}^1(\mathbb{G}, \mathscr{L}^n)$, the function

$$\mathsf{P}_t f(x) = f_t(x) = (f \star \mathsf{p}_t)(x) = \int_{\mathbb{G}} f(y)\, \mathsf{p}_t(y^{-1}x)\, \mathrm{d}y, \quad (t,x) \in (0,+\infty) \times \mathbb{G},$$

is smooth and is a solution of the heat diffusion problem

$$\begin{cases} \partial_t f_t = \Delta_{\mathbb{G}} f_t & \text{in } (0,+\infty) \times \mathbb{G}, \\ f_0 = f, & \text{on } \{0\} \times \mathbb{G}, \end{cases}$$

where the initial datum is assumed in the L^1-sense, i.e. $\lim_{t \to 0} f_t = f$ in $\mathrm{L}^1(\mathbb{G}, \mathscr{L}^n)$. Accordingly, we can define

$$\mathsf{P}_t \mu(x) = (\mu \star \mathsf{p}_t)(x) = \int_{\mathbb{G}} \mathsf{p}_t(y^{-1}x)\, \mathrm{d}\mu(y), \quad (t,x) \in (0,+\infty) \times \mathbb{G},$$

whenever $\mu \in \mathscr{P}(X)$, so that we can identify $\mathsf{H}_t = \mathsf{P}_t$ for all $t \geq 0$.

It is not difficult to recognize that the space $\mathrm{W}^{1,2}(\mathbb{G}, \mathsf{d}_{\mathrm{cc}}, \mathscr{L}^n)$ induced by the metric-measure structure $(\mathbb{G}, \mathsf{d}_{\mathrm{cc}}, \mathscr{L}^n)$ actually coincides with the well-known *horizontal Sobolev space*

$$\mathrm{W}^{1,2}_{\mathbb{G}}(\mathbb{G}) = \left\{ f \in \mathrm{L}^2(\mathbb{G}, \mathscr{L}^n) : X_i f \in \mathrm{L}^2(\mathbb{G}, \mathscr{L}^n),\ i = 1, \ldots, m_1 \right\},$$

where $X_i f$ stands for the derivative of the function f in the direction X_i defined in the usual weak sense via integration by parts against test functions. In particular, the Sobolev space $\mathrm{W}^{1,2}(\mathbb{G}, \mathsf{d}_{\mathrm{cc}}, \mathscr{L}^n)$ is Hilbertian and the Cheeger energy coincides with the *horizontal Dirichlet energy*

$$\mathsf{Ch}(f) = \int_{\mathbb{G}} \|\nabla_{\mathbb{G}} f(x)\|_{\mathbb{G}}^2\, \mathrm{d}x \quad \text{for all } f \in \mathrm{W}^{1,2}_{\mathbb{G}}(\mathbb{G}),$$

so that $|\mathrm{D}f|_w = \|\nabla_{\mathbb{G}} f\|_{\mathbb{G}}$. We refer the reader to [95, Theorem 1.3], [105, Theorem 1.2] and [111, Theorem 6.3] for more general results in this direction (note that strictly related observations are made in [9, Section 3.2] for the *BV* space in the sub-Riemannian context).

By a standard regularization argument via group convolution, we thus immediately deduce that if $f \in \mathrm{W}^{1,2}_{\mathbb{G}}(\mathbb{G})$ with $\|\nabla_{\mathbb{G}} f\|_{\mathbb{G}} \leq L$, then f agrees $\mathscr{L}^n$-a.e. with a d_{cc}-Lipschitz function with Lipschitz constant not larger than L.

In conclusion, $(\mathbb{G}, \mathsf{d}_{\mathrm{cc}}, \mathscr{L}^n)$ is an admissible metric-measure space in the sense of Section 2.11 which is also a heat-smoothing admissible group as in Definitions 3.27 and 5.8. Therefore, combining [112, Theorem 1.8] with Theorem 3.16 and Theorem 5.16, we get the following result.

Theorem 5.18 (Equivalence in Carnot groups)**.** *Let* $(\mathbb{G}, \mathsf{d}_{\mathrm{cc}}, \mathscr{L}^n)$ *be a Carnot group. There exists an optimal constant* $C_{\mathbb{G}} \geq 1$ *(depending only on the group structure and such that* $C_{\mathbb{G}} = 1$ *if and only if* $\mathbb{G}$ *is commutative) satisfying the following four equivalent properties.*

(i) $[\mathsf{BE}_w(C_{\mathbb{G}}, \infty)]$ *If* $f \in \mathrm{C}^\infty(\mathbb{G})$*, then* $\Gamma^{\mathbb{G}}(\mathsf{P}_t f) \leq C_{\mathbb{G}}^2\, \mathsf{P}_t \Gamma^{\mathbb{G}}(f)$ *for all* $t \geq 0$*.*

(ii) [Kuwada] *If* $\mu, \nu \in \mathscr{P}_2(\mathbb{G})$*, then* $W_2(\mathsf{P}_t \mu, \mathsf{P}_t \nu) \leq C_{\mathbb{G}}\, W_2(\mu, \nu)$ *for all* $t \geq 0$*.*

(iii) $[\mathsf{EVI}_w(C_{\mathbb{G}})]$ *If* $\mu_0, \mu_1 \in \mathrm{Dom}(\mathsf{Ent}_{\mathscr{L}^n})$, *then*

$$W_2^2(\mathsf{P}_{t_1}\mu_1, \mathsf{P}_{t_0}\mu_0) - C_{\mathbb{G}}^2 \, W_2^2(\mu_1, \mu_0) \leq 2\,(t_1 - t_0)\left(\mathsf{Ent}_{\mathscr{L}^n}(\mathsf{P}_{t_0}\mu_0) - \mathsf{Ent}_{\mathscr{L}^n}(\mathsf{P}_{t_1}\mu_1)\right)$$

for all $0 \leq t_0 \leq t_1$.

(iv) $[\mathsf{RCD}_w(C_{\mathbb{G}}, \infty)]$ *If* $s \mapsto \mu_s \in \mathrm{Geo}([0,1]; \mathscr{P}_2(\mathbb{G}))$ *connects* $\mu_0, \mu_1 \in \mathrm{Dom}(\mathsf{Ent}_{\mathscr{L}^n})$, *then*

$$\mathsf{Ent}_{\mathscr{L}^n}(\mathsf{P}_{t+h}\mu_s) \leq (1-s)\,\mathsf{Ent}_{\mathscr{L}^n}(\mathsf{P}_t\mu_0) + s\,\mathsf{Ent}_{\mathscr{L}^n}(\mathsf{P}_t\mu_1)$$
$$+ \frac{s(1-s)}{2h}\left(C_{\mathbb{G}}^2\,W_2^2(\mu_0, \mu_1) - W_2^2(\mathsf{P}_t\mu_0, \mathsf{P}_t\mu_1)\right)$$

for all $s \in [0,1]$, $t \geq 0$ *and* $h > 0$.

Thanks to the entropy dissipation along the heat flow proved in [20, Proposition 4.2] and to the integrability property of the Fisher information along the heat flow given by Lemma 4.1(ii), from Theorem 5.18(iv) we deduce the following weak convexity property of the entropy along W_2-geodesics in Carnot groups. Here and in the following, we let

$$\mathsf{F}_{\mathbb{G}}(f) = \int_{\mathbb{G} \cap \{f > 0\}} \frac{\|\nabla_{\mathbb{G}} f\|_{\mathbb{G}}^2}{f} \, \mathrm{d}x$$

be the Fisher information in the Carnot group $\mathbb{G}$.

Corollary 5.19 (Weak convexity of $\mathsf{Ent}_{\mathscr{L}^n}$ in Carnot groups)**.** *Let* $(\mathbb{G}, \mathsf{d}_{\mathrm{cc}}, \mathscr{L}^n)$ *be a Carnot group and let* $C_{\mathbb{G}} \geq 1$ *be as in Theorem 5.18. Let* $s \mapsto \mu_s \in \mathrm{Geo}([0,1]; \mathscr{P}_2(\mathbb{G}))$ *be a* W_2-geodesic connecting $\mu_0, \mu_1 \in \mathrm{Dom}(\mathsf{Ent}_{\mathscr{L}^n})$. If $\mu_s \in \mathrm{Dom}(\mathsf{Ent}_{\mathscr{L}^n})$ for some $s \in (0,1)$, then*

$$\mathsf{Ent}_{\mathscr{L}^n}(\mu_s) \leq (1-s)\,\mathsf{Ent}_{\mathscr{L}^n}(\mathsf{P}_t\mu_0) + s\,\mathsf{Ent}_{\mathscr{L}^n}(\mathsf{P}_t\mu_1)$$
$$+ \frac{s(1-s)}{2h}\left(C_{\mathbb{G}}^2\,W_2^2(\mu_0, \mu_1) - W_2^2(\mathsf{P}_t\mu_0, \mathsf{P}_t\mu_1)\right) + \int_0^{t+h} \mathsf{F}_{\mathbb{G}}(\mathsf{P}_r\mu_s)\,\mathrm{d}r \tag{5.43}$$

for all $t \geq 0$ *and* $h > 0$.

It is interesting to compare inequality (5.43) when $t = 0$ with the entropic inequality obtained in [29, Corollary 3.4] when $\mathbb{G} = \mathbb{H}^n$, the n-dimensional Heisenberg group. Note that similar comparisons can be made for several others Carnot groups thanks to the results obtained in [30, 35].

The Heisenberg group $\mathbb{H}^n$, $n \in \mathbb{N}$, is the non-commutative Carnot group of step 2 whose Lie algebra satisfies $\mathfrak{g} = V_1 \oplus V_2$ with $m_1 = 2n$, $m_2 = 1$ and

$$X_i = \partial_{x_i} - \frac{x_{n+i}}{2}\,\partial_{x_{2n+1}}, \quad X_{n+i} = \partial_{x_{n+i}} + \frac{x_i}{2}\,\partial_{x_{2n+1}}, \quad X_{2n+1} = \partial_{x_{2n+1}},$$

for all $i = 1, \ldots, n$. Thanks to [29, Corollary 3.4], the unique W_2-geodesic $s \mapsto \mu_s \in \mathrm{Geo}([0,1]; \mathscr{P}_2(X))$ joining two compactly supported measures $\mu_0, \mu_1 \in \mathrm{Dom}(\mathsf{Ent}_{\mathscr{L}^{2n+1}})$ satisfies

$$\mathsf{Ent}_{\mathscr{L}^{2n+1}}(\mu_s) \leq (1-s)\,\mathsf{Ent}_{\mathscr{L}^{2n+1}}(\mu_0) + s\,\mathsf{Ent}_{\mathscr{L}^{2n+1}}(\mu_1) + w(s) \tag{5.44}$$

for all $s \in (0,1)$, where

$$w(s) = -2\log\left((1-s)^{(1-s)}s^s\right) \tag{5.45}$$

for all $s \in (0,1)$. Note that the function in (5.45) is concave and such that

$$\lim_{s \to 0^+} w(s) = \lim_{s \to 1^-} w(s) = 0$$

and it satisfies

$$0 < w(s) \leq w(\tfrac{1}{2}) = \log 4$$

for all $s \in (0,1)$. Therefore, as a consequence of (5.44), we get that $\mu_s \in \mathrm{Dom}(\mathsf{Ent}_{\mathscr{L}^{2n+1}})$ and hence, by Corollary 5.19, we can also estimate

$$\mathsf{Ent}_{\mathscr{L}^{2n+1}}(\mu_s) \leq (1-s)\,\mathsf{Ent}_{\mathscr{L}^{2n+1}}(\mu_0) + s\,\mathsf{Ent}_{\mathscr{L}^{2n+1}}(\mu_1) + \sigma(s) \tag{5.46}$$

where

$$\sigma(s) = \inf\left\{ h > 0 : s(1-s)\,\frac{C_{\mathbb{H}^n}^2 - 1}{2h}\,W_2^2(\mu_0, \mu_1) + \int_0^h \mathsf{F}_{\mathbb{H}^n}(\mathsf{P}_r \mu_s)\,\mathrm{d}r \right\} \tag{5.47}$$

for all $s \in [0,1]$.

Although we are not able to give a more explicit formula for the function in (5.47), it appears that, at least in some cases when μ_0 and μ_1 are very close to each other, inequality (5.46) is more precise than inequality (5.44) for intermediate times, that is, $\sigma(s) < w(s)$ for some $s \in (0,1)$. As a trivial example, if $\mu_0 = \mu_1$, then $\mu_s = \mu_0$ for all $s \in (0,1)$ and thus inequality (5.44) reduces to $0 \leq w(s)$ for all $s \in (0,1)$, while $\sigma(s) = 0$ for all $s \in [0,1]$. As a less trivial example, exploiting the fact that right translations are optimal transport maps in $\mathbb{H}^n$, see [18, Example 5.7] and [76, Section 2.1], we can prove the following result. Here we let

$$\tilde{\mathsf{F}}_{\mathbb{H}^n}(f) = \int_{\mathbb{H}^n \cap \{f > 0\}} \frac{\|\tilde{\nabla}_{\mathbb{H}^n} f\|_{\mathbb{H}^n}^2}{f}\,\mathrm{d}x$$

be the Fisher information in the Heisenberg group $\mathbb{H}^n$ relative to the *right-invariant horizontal gradient* $\tilde{\nabla}_{\mathbb{H}^n}$.

Proposition 5.20 (An estimate of σ for right translations). *Let $\mu_0 = f_0 \mathscr{L}^{2n+1} \in \mathrm{Dom}(\mathsf{Ent}_{\mathscr{L}^{2n+1}})$ be such that $f_0 \in \mathrm{C}_c^1(\mathbb{R}^{2n+1})$ with $\tilde{\mathsf{F}}_{\mathbb{H}^n}(f_0) < +\infty$. Let $u \in \mathbb{H}^n$ be a horizontal point in $\mathbb{H}^n$, i.e. $u_{2n+1} = 0$, and define $T_s(x) = x u_s$ for all $x \in \mathbb{H}^n$ and $s \in [0,1]$, where $u_s = su$. Then $s \mapsto \mu_s = (T_s)_\sharp \mu_0$ is the unique W_2-geodesic joining μ_0 with $\mu_1 = (T_1)_\sharp \mu_0$ and*

$$\sigma(s) \leq \mathsf{d}_{\mathrm{cc}}(u, \mathsf{o})\sqrt{2s(1-s)\,(C_{\mathbb{H}^n}^2 - 1)\,\tilde{\mathsf{F}}_{\mathbb{H}^n}(f)}$$

for all $s \in (0,1)$. In particular, for any $\varepsilon \in (0,1)$ there exists $\delta > 0$ such that

$$\mathsf{d}_{\mathrm{cc}}(u, \mathsf{o})\sqrt{\tilde{\mathsf{F}}_{\mathbb{H}^n}(f)} < \delta \implies \sigma(s) < w(s) \text{ for all } s \in (\varepsilon, 1 - \varepsilon).$$

Proof. The fact that $s \mapsto \mu_s = (T_s)_\sharp \mu_0$ is the unique W_2-geodesic joining μ_0 with $\mu_1 = (T_1)_\sharp \mu_0$ can be proved arguing as in [76, Section 2.1] since, up to a rotation fixing the vertical axis $\{x \in \mathbb{R}^{2n+1} : x_i = 0 \text{ for all } i = 1, \dots, 2n\}$, one can also assume $u_i = 0$ for all $i = 2, \dots, 2n$. We thus omit the details. By the optimality of right translations, we have

$$W_2^2(\mu_0, \mu_1) = \int_{\mathbb{H}^n} \mathsf{d}_{\mathrm{cc}}^2(x, T_1(x))\,\mathrm{d}\mu_0(x) = \int_{\mathbb{H}^n} \mathsf{d}_{\mathrm{cc}}^2(x, xu)\,\mathrm{d}\mu_0(x) = \mathsf{d}_{\mathrm{cc}}^2(u, \mathsf{o}).$$

Moreover, we can write $\mu_s = f_s \mathscr{L}^{2n+1}$, where $f_s = f_0 \circ T_s \in \mathrm{C}_c^1(\mathbb{R}^{2n+1})$, so that

$$\nabla_{\mathbb{H}^n}(\mathsf{P}_r f_s) = \nabla_{\mathbb{H}^n}(f_s \star \mathsf{p}_r) = (\tilde{\nabla}_{\mathbb{H}^n} f_s) \star \mathsf{p}_r$$

for all $r > 0$ and $s \in (0,1)$, and thus

$$\mathsf{F}_{\mathbb{H}^n}(\mathsf{P}_r f_s) = \int_{\mathbb{H}^n} \frac{|((\tilde{\nabla}_{\mathbb{H}^n} f_s) \star \mathsf{p}_r)(x)|_{\mathbb{H}^n}^2}{(f_s \star \mathsf{p}_r)(x)}\,\mathrm{d}x \leq \int_{\mathbb{H}^n} \frac{|\tilde{\nabla}_{\mathbb{H}^n} f_s(x)|_{\mathbb{H}^n}^2}{f_s(x)}\,\mathrm{d}x = \int_{\mathbb{H}^n} \frac{|\tilde{\nabla}_{\mathbb{H}^n} f_0(x)|_{\mathbb{H}^n}^2}{f_0(x)}\,\mathrm{d}x$$

for all $r > 0$ and $s \in (0, 1)$ by Jensen inequality, arguing as in the proof of [20, Lemma 4.5] (for right convolutions instead of left ones). Hence, from (5.47), we get

$$\sigma(s) \leq s(1 - s) \frac{C^2_{\mathbb{H}^n} - 1}{2h} W_2^2(\mu_0, \mu_1) + \int_0^h \mathsf{F}_{\mathbb{H}^n}(\mathsf{P}_r \mu_s) \, \mathrm{d}r$$

$$\leq s(1 - s) \frac{C^2_{\mathbb{H}^n} - 1}{2h} \mathsf{d}^2_{\mathrm{cc}}(u, \mathsf{o}) + h \tilde{\mathsf{F}}_{\mathbb{H}^n}(f_0)$$

for all $h > 0$ and $s \in (0, 1)$. The conclusion thus follows by optimizing with respect to $h > 0$ and by recalling (5.45). $\qquad \square$

5.9.2. *The* $\mathbb{SU}(2)$ *group.* The group $\mathbb{SU}(2)$ is the Lie group of 2×2 complex unitary matrices with determinant 1. Its Lie algebra $\mathfrak{su}(2)$ consists of all 2×2 complex unitary skew-Hermitian matrices with trace 0. Following the notation of [38], a basis of $\mathfrak{su}(2)$ is given by the *Pauli matrices*

$$X = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},$$

satisfying the relations

$$[X, Y] = 2Z, \quad [Y, Z] = 2X, \quad [Z, X] = 2Y.$$

We keep the same notation X, Y and Z for the left invariant vector fields on $\mathbb{SU}(2)$ corresponding to the Pauli matrices. Similarly as before, we let

$$\nabla_{\mathbb{SU}(2)} f = (Xf)X + (Yf)Y$$

be the *horizontal gradient* of f. Using the cylindric coordinates

$$(r, \vartheta, z) \mapsto \exp(r \cos \vartheta \, X + r \sin \vartheta \, Y) \exp(\zeta Z) = \begin{pmatrix} e^{i\zeta} \cos r & e^{i(\vartheta - \zeta)} \sin r \\ -e^{-i(\vartheta - \zeta)} \sin r & e^{-i\zeta} \cos r \end{pmatrix}$$

valid for $r \in [0, \frac{\pi}{2})$, $\vartheta \in [0, 2\pi]$ and $\zeta \in [-\pi, \pi]$ (originally introduced in [66]), the normalized bi-invariant Haar measure $\mathfrak{m} \in \mathscr{P}(\mathbb{SU}(2))$ can be written as

$$\mathrm{d}\mathfrak{m} = \frac{1}{4\pi^2} \sin(2r) \, \mathrm{d}r \, \mathrm{d}\vartheta \, \mathrm{d}\zeta.$$

Once the left-invariant Riemannian metric $\langle \cdot, \cdot \rangle_{\mathbb{SU}(2)}$ making the basis X, Y, Z orthonormal is introduced, we can endow the group $\mathbb{SU}(2)$ with the Carnot–Carathéodory distance d_{cc} defined analogously as in (5.42). The resulting metric space $(\mathbb{SU}(2), \mathsf{d}_{\mathrm{cc}})$ is a complete, separable, locally compact and geodesic space.

The standard sub-Laplacian operator is $\Delta_{\mathbb{SU}(2)} = X^2 + Y^2$ and, again by Hörmander Theorem, the heat operator $\partial_t - \Delta_{\mathbb{SU}(2)}$ has a smooth fundamental solution $\mathsf{p} \colon (0, +\infty) \times \mathbb{SU}(2) \to (0, +\infty)$ which induces the associated heat flow $(\mathsf{P}_t)_{t \geq 0}$ by right convolution (so that we can still identify $\mathsf{H}_t = \mathsf{P}_t$ for all $t \geq 0$).

Arguing similarly in the case of Carnot groups (recall the previously cited [95, 105, 111]), it is possible to identify the space $\mathrm{W}^{1,2}(\mathbb{SU}(2), \mathsf{d}_{\mathrm{cc}}, \mathfrak{m})$ with the horizontal Sobolev space

$$\mathrm{W}^{1,2}_{\mathbb{SU}(2)}(\mathbb{SU}(2)) = \left\{ f \in \mathrm{L}^2(\mathbb{SU}(2), \mathfrak{m}) : Xf, Yf \in \mathrm{L}^2(\mathbb{SU}(2), \mathfrak{m}) \right\}$$

defined using integration by parts against test functions, so that $\mathrm{W}^{1,2}(\mathbb{SU}(2), \mathsf{d}_{\mathrm{cc}}, \mu)$ is Hilbertian, the Cheeger energy coincides with the horizontal Dirichlet energy

$$\mathsf{Ch}(f) = \int_{\mathbb{SU}(2)} \|\nabla_{\mathbb{SU}(2)} f\|^2_{\mathbb{SU}(2)} \, \mathrm{d}\mathfrak{m}, \quad \text{for all } f \in \mathrm{W}^{1,2}_{\mathbb{SU}(2)}(\mathbb{SU}(2)),$$

and $|\mathrm{D}f|_w = \|\nabla_{\mathbb{SU}(2)}f\|_{\mathbb{SU}(2)}$. We again refer the reader to [105, Theorem 1.2] for a proof of these identifications.

Exploiting the group structure of $\mathbb{SU}(2)$ similarly to the case of Carnot groups, we get that if $f \in \mathrm{W}^{1,2}_{\mathbb{SU}(2)}(\mathbb{SU}(2))$ with $\|\nabla_{\mathbb{SU}(2)}f\|_{\mathbb{SU}(2)} \leq L$, then f agrees μ-a.e. with a d_{cc}-Lipschitz function with Lipschitz constant not larger than L.

Therefore, $(\mathbb{SU}(2), \mathsf{d}_{\mathrm{cc}}, \mathfrak{m})$ is a heat-smoothing admissible group and, combining [38, Theorem 4.10] with Theorem 3.16 and Theorem 5.16, we get the following result.

Theorem 5.21 (Equivalence in $\mathbb{SU}(2)$). *Let* $(\mathbb{SU}(2), \mathsf{d}_{\mathrm{cc}}, \mathfrak{m})$ *be as above. There exists a constant* $C_{\mathbb{SU}(2)} \geq \sqrt{2}$ *satisfying the following four equivalent properties.*

  (i) $[\mathsf{BE}_w(C_{\mathbb{SU}(2)}e^{-2t}, \infty)]$ *If* $f \in \mathrm{C}^\infty(\mathbb{SU}(2))$, *then* $\Gamma^{\mathbb{SU}(2)}(\mathsf{P}_t f) \leq C^2_{\mathbb{SU}(2)}e^{-4t}\,\mathsf{P}_t\Gamma^{\mathbb{SU}(2)}(f)$ *for all* $t \geq 0$.

  (ii) [Kuwada] *If* $\mu, \nu \in \mathscr{P}_2(\mathbb{SU}(2))$, *then* $W_2(\mathsf{P}_t\mu, \mathsf{P}_t\nu) \leq C_{\mathbb{SU}(2)}e^{-2t}\,W_2(\mu,\nu)$ *for all* $t \geq 0$.

  (iii) $[\mathsf{EVI}_w(C_{\mathbb{SU}(2)}e^{-2t})]$ *If* $\mu_0, \mu_1 \in \mathrm{Dom}(\mathsf{Ent}_\mathfrak{m})$, *then*

$$W_2^2(\mathsf{P}_{t_1}\mu_1, \mathsf{P}_{t_0}\mu_0) - C^2_{\mathbb{SU}(2)}\,\frac{4(t_1 - t_0)}{e^{4t_1} - e^{4t_0}}\,W_2^2(\mu_1, \mu_0) \leq 2\,(t_1 - t_0)\left(\mathsf{Ent}_\mathfrak{m}(\mathsf{P}_{t_0}\mu_0) - \mathsf{Ent}_\mathfrak{m}(\mathsf{P}_{t_1}\mu_1)\right)$$

  for all $0 \leq t_0 \leq t_1$.

  (iv) $[\mathsf{RCD}_w(C_{\mathbb{SU}(2)}e^{-2t}, \infty)]$ *If* $s \mapsto \mu_s \in \mathrm{Geo}([0,1]; \mathscr{P}_2(\mathbb{SU}(2)))$ *is a geodesic connecting* $\mu_0, \mu_1 \in \mathrm{Dom}(\mathsf{Ent}_\mathfrak{m})$, *then*

$$\mathsf{Ent}_\mathfrak{m}(\mathsf{P}_{t+h}\mu_s) \leq (1 - s)\,\mathsf{Ent}_\mathfrak{m}(\mathsf{P}_t\mu_0) + s\,\mathsf{Ent}_\mathfrak{m}(\mathsf{P}_t\mu_1)$$

$$+ \frac{s(1-s)}{2h}\left(\frac{4C^2_{\mathbb{SU}(2)}h}{e^{4(t+h)} - e^{4t}}\,W_2^2(\mu_0, \mu_1) - W_2^2(\mathsf{P}_t\mu_0, \mathsf{P}_t\mu_1)\right)$$

  for all $s \in [0,1]$, $t \geq 0$ *and* $h > 0$.

Since $(\mathbb{SU}(2), \mathsf{d}_{\mathrm{cc}}, \mathfrak{m})$ is a *Sasakian manifold*, the resulting sub-Riemannian structure on $\mathbb{SU}(2)$ is *ideal*, see [36, Definition 13 and Section 7.4] for the precise definitions. Thus, according to [36, Theorem 39], if $\mu_0, \mu_1 \in \mathrm{Dom}(\mathsf{Ent}_\mathfrak{m})$ are two compactly supported probability measures, then the unique Wasserstein geodesic $s \mapsto \mu_s \in \mathrm{Geo}([0,1]; \mathscr{P}_2(\mathbb{SU}(2)))$ joining them satisfies $\mu_s \ll \mathfrak{m}$ for all $s \in [0,1]$. Thanks to [36, Theorem 9 and Corollary 67] (see also [5,106]), it actually holds that $\mu_s \in \mathrm{Dom}(\mathsf{Ent}_\mathfrak{m})$ for all $s \in (0,1)$ and the function $s \mapsto \mathsf{Ent}_\mathfrak{m}(\mu_s)$ satisfies an inequality similar to (5.44).

Up to our knowledge, there is no analogue of the entropy dissipation for L^1-densities proved in [20, Proposition 4.2] for the $\mathbb{SU}(2)$ group and we can only rely on the general result for $\mathrm{L}^1 \cap \mathrm{L}^2$-densities obtained in [14, Proposition 4.22]. Thus, at the present moment, an inequality for the function $s \mapsto \mathsf{Ent}_\mathfrak{m}(\mu_s)$ similar to (5.43) holds in the $\mathbb{SU}(2)$ group under the additional assumption that $\frac{\mathrm{d}\mu_s}{\mathrm{d}\mathfrak{m}} \in \mathrm{L}^2(X, \mathfrak{m})$ for some $s \in (0,1)$. Also, up to our knowledge, it is not known if right translations in the $\mathbb{SU}(2)$ group are optimal transport maps. For this reason, a comparison of the entropic inequalities in the $\mathbb{SU}(2)$ group analogous to the one done above for Carnot groups is not easily reachable at the present moment. We will hopefully come back to this topic in a future work.

References

[1] A. Agrachev, D. Barilari, and U. Boscain, *A Comprehensive Introduction to sub-Riemannian Geometry*, Cambridge Studies in Advanced Mathematics, vol. 181, Cambridge University Press, Cambridge, 2020.

[2] A. Agrachev, D. Barilari, and L. Rizzi, *Sub-Riemannian curvature in contact geometry*, J. Geom. Anal. **27** (2017), no. 1, 366–408.

[3] A. Agrachev, D. Barilari, and L. Rizzi, *Curvature: a variational approach*, Mem. Amer. Math. Soc. **256** (2018), no. 1225, v+142.

[4] A. Agrachev and P. Lee, *Optimal transportation under nonholonomic constraints*, Trans. Amer. Math. Soc. **361** (2009), no. 11, 6019–6047.

[5] ———, *Generalized Ricci curvature bounds for three dimensional contact subriemannian manifolds*, Math. Ann. **360** (2014), no. 1-2, 209–253.

[6] A. Agrachev and I. Zelenko, *Geometry of Jacobi curves. I*, J. Dynam. Control Systems **8** (2002), no. 1, 93–140.

[7] ———, *Geometry of Jacobi curves. II*, J. Dynam. Control Systems **8** (2002), no. 2, 167–215.

[8] L. Ambrosio, M. Colombo, and S. Di Marino, *Sobolev spaces in metric measure spaces: reflexivity and lower semicontinuity of slope*, Variational methods for evolving objects, Adv. Stud. Pure Math., vol. 67, Math. Soc. Japan, Tokyo, 2015, pp. 1–58.

[9] L. Ambrosio, R. Ghezzi, and V. Magnani, *BV functions and sets of finite perimeter in sub-Riemannian manifolds*, Ann. Inst. H. Poincaré Anal. Non Linéaire **32** (2015), no. 3, 489–517.

[10] L. Ambrosio and N. Gigli, *A user's guide to optimal transport*, Modelling and optimisation of flows on networks, Lecture Notes in Math., vol. 2062, Springer, Heidelberg, 2013, pp. 1–155.

[11] L. Ambrosio, N. Gigli, A. Mondino, and T. Rajala, *Riemannian Ricci curvature lower bounds in metric measure spaces with σ-finite measure*, Trans. Amer. Math. Soc. **367** (2015), no. 7, 4661–4701.

[12] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, 2nd ed., Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2008.

[13] ———, *Density of Lipschitz functions and equivalence of weak gradients in metric measure spaces*, Rev. Mat. Iberoam. **29** (2013), no. 3, 969–996.

[14] ———, *Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below*, Invent. Math. **195** (2014), no. 2, 289–391.

[15] ———, *Metric measure spaces with Riemannian Ricci curvature bounded from below*, Duke Math. J. **163** (2014), no. 7, 1405–1490.

[16] ———, *Bakry-Émery curvature-dimension condition and Riemannian Ricci curvature bounds*, Ann. Probab. **43** (2015), no. 1, 339–404.

[17] L. Ambrosio, A. Mondino, and G. Savaré, *Nonlinear diffusion equations and curvature conditions in metric measure spaces*, Mem. Amer. Math. Soc. **262** (2019), no. 1270.

[18] L. Ambrosio and S. Rigot, *Optimal mass transportation in the Heisenberg group*, J. Funct. Anal. **208** (2004), no. 2, 261–301.

[19] L. Ambrosio, G. Savaré, and L. Zambotti, *Existence and stability for Fokker-Planck equations with log-concave reference measure*, Probab. Theory Related Fields **145** (2009), no. 3-4, 517–564.

[20] L. Ambrosio and G. Stefani, *Heat and entropy flows in Carnot groups*, Rev. Mat. Iberoam. **36** (2020), no. 1, 257–290.

[21] Z. Badreddine, *Mass transportation on sub-Riemannian structures of rank two in dimension four*, Ann. Inst. H. Poincaré Anal. Non Linéaire **36** (2019), no. 3, 837–860.

[22] D. Bakry, *L'hypercontractivité et son utilisation en théorie des semigroupes*, Lectures on probability theory (Saint-Flour, 1992), Lecture Notes in Math., vol. 1581, Springer, Berlin, 1994, pp. 1–114 (French).

[23] ———, *Functional inequalities for Markov semigroups*, Probability measures on groups: recent directions and trends, Tata Inst. Fund. Res., Mumbai, 2006, pp. 91–147.

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

[25] ———, *On Harnack inequalities and optimal transportation*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **14** (2015), no. 3, 705–727.

[26] D. Bakry, F. Baudoin, M. Bonnefont, and D. Chafaï, *On gradient bounds for the heat kernel on the Heisenberg group*, J. Funct. Anal. **255** (2008), no. 8, 1905–1938.

[27] D. Bakry and M. Émery, *Diffusions hypercontractives*, Séminaire de probabilités, XIX, 1983/84, Lecture Notes in Math., vol. 1123, Springer, Berlin, 1985, pp. 177–206.

[28] D. Bakry and M. Ledoux, *A logarithmic Sobolev form of the Li-Yau parabolic inequality*, Rev. Mat. Iberoam. **22** (2006), no. 2, 683–702.

[29] Z. M. Balogh, A. Kristály, and K. Sipos, *Geometric inequalities on Heisenberg groups*, Calc. Var. Partial Differential Equations **57** (2018), no. 2, Paper No. 61, 41.

[30] _____, *Jacobian determinant inequality on corank 1 Carnot groups with applications*, J. Funct. Anal. **277** (2019), no. 12, 108293, 36.

[31] D. Barilari and S. Ivanov, *A Bonnet-Myers type theorem for quaternionic contact structures*, Calc. Var. Partial Differential Equations **58** (2019), no. 1, Paper No. 37, 26.

[32] D. Barilari and L. Rizzi, *Comparison theorems for conjugate points in sub-Riemannian geometry*, ESAIM Control Optim. Calc. Var. **22** (2016), no. 2, 439–472.

[33] _____, *On Jacobi fields and a canonical connection in sub-Riemannian geometry*, Arch. Math. (Brno) **53** (2017), no. 2, 77–92.

[34] _____, *Sharp measure contraction property for generalized H-type Carnot groups*, Commun. Contemp. Math. **20** (2018), no. 6, 1750081, 24.

[35] _____, *Sub-Riemannian interpolation inequalities*, Invent. Math. **215** (2019), no. 3, 977–1038.

[36] _____, *Bakry-Émery curvature and model spaces in sub-Riemannian geometry*, Math. Ann. **377** (2020), no. 1-2, 435–482.

[37] F. Baudoin, *Bakry-Émery meet Villani*, J. Funct. Anal. **273** (2017), no. 7, 2275–2291.

[38] F. Baudoin and M. Bonnefont, *The subelliptic heat kernel on* SU(2)*: representations, asymptotics and gradient bounds*, Math. Z. **263** (2009), no. 3, 647–672.

[39] _____, *Curvature-dimension estimates for the Laplace-Beltrami operator of a totally geodesic foliation*, Nonlinear Anal. **126** (2015), 159–169.

[40] _____, *Reverse Poincaré inequalities, isoperimetry, and Riesz transforms in Carnot groups*, Nonlinear Anal. **131** (2016), 48–59.

[41] F. Baudoin, M. Bonnefont, and N. Garofalo, *A sub-Riemannian curvature-dimension inequality, volume doubling property and the Poincaré inequality*, Math. Ann. **358** (2014), no. 3-4, 833–860.

[42] F. Baudoin, M. Bonnefont, N. Garofalo, and I. H. Munive, *Volume and distance comparison theorems for sub-Riemannian manifolds*, J. Funct. Anal. **267** (2014), no. 7, 2005–2027.

[43] F. Baudoin and M. Cecil, *The subelliptic heat kernel on the three-dimensional solvable Lie groups*, Forum Math. **27** (2015), no. 4, 2051–2086.

[44] F. Baudoin and N. Garofalo, *Curvature-dimension inequalities and Ricci lower bounds for sub-Riemannian manifolds with transverse symmetries*, J. Eur. Math. Soc. (JEMS) **19** (2017), no. 1, 151–219.

[45] F. Baudoin and D. J. Kelleher, *Differential one-forms on Dirichlet spaces and Bakry-Émery estimates on metric graphs*, Trans. Amer. Math. Soc. **371** (2019), no. 5, 3145–3178.

[46] F. Baudoin and B. Kim, *Sobolev, Poincaré, and isoperimetric inequalities for subelliptic diffusion operators satisfying a generalized curvature dimension inequality*, Rev. Mat. Iberoam. **30** (2014), no. 1, 109–131.

[47] F. Baudoin and J. Wang, *Curvature dimension inequalities and subelliptic heat kernel gradient bounds on contact manifolds*, Potential Anal. **40** (2014), no. 2, 163–193.

[48] A. Bejancu, *Curvature in sub-Riemannian geometry*, J. Math. Phys. **53** (2012), no. 2, 023513, 25.

[49] A. Bejancu and H. R. Farran, *Curvature of CR manifolds*, An. Ştiinţ. Univ. Al. I. Cuza Iaşi. Mat. (N.S.) **59** (2013), no. 1, 43–72.

[50] V. N. Berestovskii, *Curvatures of homogeneous sub-Riemannian manifolds*, Eur. J. Math. **3** (2017), no. 4, 788–807.

[51] M. Biroli and U. Mosco, *A Saint-Venant type principle for Dirichlet forms on discontinuous media*, Ann. Mat. Pura Appl. (4) **169** (1995), 125–181.

[52] V. I. Bogachev, *Measure theory. Vol. I, II*, Springer-Verlag, Berlin, 2007.

[53] F. Bolley, I. Gentil, and A. Guillin, *Dimensional contraction via Markov transportation distance*, J. Lond. Math. Soc. (2) **90** (2014), no. 1, 309–332.

[54] F. Bolley, I. Gentil, A. Guillin, and K. Kuwada, *Equivalence between dimensional contractions in Wasserstein distance and the curvature-dimension condition*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **18** (2018), no. 3.

[55] A. Bonfiglioli, E. Lanconelli, and F. Uguzzoni, *Stratified Lie groups and potential theory for their sub-Laplacians*, Springer Monographs in Mathematics, Springer, Berlin, 2007.

[56] N. Bouleau and F. Hirsch, *Dirichlet forms and analysis on Wiener space*, De Gruyter Studies in Mathematics, vol. 14, Walter de Gruyter & Co., Berlin, 1991.

[57] H. Brézis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, 1973 (French). North-Holland Mathematics Studies, No. 5. Notas de Matemática (50).

[58] _____, *Functional analysis, Sobolev spaces and partial differential equations*, Universitext, Springer, New York, 2011.

[59] E. A. Carlen, S. Kusuoka, and D. W. Stroock, *Upper bounds for symmetric Markov transition functions*, Ann. Inst. H. Poincaré Probab. Statist. **23** (1987), no. 2, suppl., 245–287.

[60] F. Cavalletti and A. Mondino, *Measure rigidity of Ricci curvature lower bounds*, Adv. Math. **286** (2016), 430–480.

[61] _____, *Optimal maps in essentially non-branching spaces*, Commun. Contemp. Math. **19** (2017), no. 6, 1750007, 27.

[62] J. Cheeger, *Differentiability of Lipschitz functions on metric measure spaces*, Geom. Funct. Anal. **9** (1999), no. 3, 428–517.

[63] L.-J. Cheng, A. Thalmaier, and S.-Q. Zhang, *Exponential contraction in Wasserstein distance on static and evolving manifolds*, Rev. Roumaine Math. Pures Appl. **66** (2021), no. 1, 107–129.

[64] D. Cordero-Erausquin, R. J. McCann, and M. Schmuckenschläger, *A Riemannian interpolation inequality à la Borell, Brascamp and Lieb*, Invent. Math. **146** (2001), no. 2, 219–257.

[65] T. Coulhon, R. Jiang, P. Koskela, and A. Sikora, *Gradient estimates for heat kernels and harmonic functions*, J. Funct. Anal. **278** (2020), no. 8, 108398, 67.

[66] M. Cowling and A. Sikora, *A spectral multiplier theorem for a sublaplacian on* SU(2), Math. Z. **238** (2001), no. 1, 1–36.

[67] S. Daneri and G. Savaré, *Eulerian calculus for the displacement convexity in the Wasserstein distance*, SIAM J. Math. Anal. **40** (2008), no. 3, 1104–1122.

[68] L. De Pascale and S. Rigot, *Monge's transport problem in the Heisenberg group*, Adv. Calc. Var. **4** (2011), no. 2, 195–227.

[69] B. K. Driver and T. Melcher, *Hypoelliptic heat kernel inequalities on the Heisenberg group*, J. Funct. Anal. **221** (2005), no. 2, 340–365.

[70] R. Durrett, *Probability: Theory and Examples*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 49, Cambridge University Press, Cambridge, 2019.

[71] M. Erbar, *The heat equation on manifolds as a gradient flow in the Wasserstein space*, Ann. Inst. Henri Poincaré Probab. Stat. **46** (2010), no. 1, 1–23.

[72] M. Erbar, K. Kuwada, and K.-T. Sturm, *On the equivalence of the entropic curvature-dimension condition and Bochner's inequality on metric measure spaces*, Invent. Math. **201** (2015), no. 3, 993–1071.

[73] H. Federer, *Geometric measure theory*, Die Grundlehren der mathematischen Wissenschaften, Band 153, Springer-Verlag New York Inc., New York, 1969.

[74] Q. Feng and W. Li, *Sub-Riemannian Ricci curvature via generalized Gamma z calculus* (2020), preprint, available at `https://arxiv.org/abs/2004.01863`.

[75] _____, *Generalized Gamma z calculus via sub-Riemannian density manifold* (2020), preprint, available at `https://arxiv.org/abs/1910.07480`.

[76] A. Figalli and N. Juillet, *Absolute continuity of Wasserstein geodesics in the Heisenberg group*, J. Funct. Anal. **255** (2008), no. 1, 133–141.

[77] A. Figalli and L. Rifford, *Mass transportation on sub-Riemannian manifolds*, Geom. Funct. Anal. **20** (2010), no. 1, 124–159.

[78] G. B. Folland, *Real analysis. Modern Techniques and Their Applications*, 2nd ed., Pure and Applied Mathematics (New York), John Wiley & Sons, Inc., New York, 1999.

[79] _____, *A Course in Abstract Harmonic Analysis*, 2nd ed., Textbooks in Mathematics, CRC Press, Boca Raton, FL, 2016.

[80] I. Gentil, *Dimensional contraction in Wasserstein distance for diffusion semigroups on a Riemannian manifold*, Potential Anal. **42** (2015), no. 4, 861–873.

[81] N. Gigli, *On the heat flow on metric measure spaces: existence, uniqueness and stability*, Calc. Var. Partial Differential Equations **39** (2010), no. 1-2, 101–120.

[82] _____, *An overview of the proof of the splitting theorem in spaces with non-negative Ricci curvature*, Anal. Geom. Metr. Spaces **2** (2014), no. 1, 169–213.

[83] _____, *On the differential structure of metric measure spaces and applications*, Mem. Amer. Math. Soc. **236** (2015), no. 1113, vi+91.

[84] N. Gigli, K. Kuwada, and S.-I. Ohta, *Heat flow on Alexandrov spaces*, Comm. Pure Appl. Math. **66** (2013), no. 3, 307–331.

[85] N. Gigli and C. Mantegazza, *A flow tangent to the Ricci flow via heat kernels and mass transport*, Adv. Math. **250** (2014), 74–104.

[86] E. Grong, *Model spaces in sub-Riemannian geometry*, Comm. Anal. Geom. **29** (2021), no. 1, 77–113.

[87] _____, *Affine connections and curvature in sub-Riemannian Geometry* (2020), preprint, available at `https://arxiv.org/abs/2001.03817v1`.

[88] E. Grong and A. Thalmaier, *Curvature-dimension inequalities on sub-Riemannian manifolds obtained from Riemannian foliations: part I*, Math. Z. **282** (2016), no. 1-2, 99–130.

[89] _____, *Curvature-dimension inequalities on sub-Riemannian manifolds obtained from Riemannian foliations: part II*, Math. Z. **282** (2016), no. 1-2, 131–164.

[90] _____, *Stochastic completeness and gradient representations for sub-Riemannian manifolds*, Potential Anal. **51** (2019), no. 2, 219–254.

[91] J. Heinonen, P. Koskela, N. Shanmugalingam, and J. T. Tyson, *Sobolev spaces on metric measure spaces. An approach based on upper gradients*, New Mathematical Monographs, vol. 27, Cambridge University Press, Cambridge, 2015.

[92] E. Hille and R. S. Phillips, *Functional Analysis and Semi-Groups*, American Mathematical Society, Providence, R. I., 1974. Third printing of the revised edition of 1957; American Mathematical Society Colloquium Publications, Vol. XXXI.

[93] R. K. Hladky, *Connections and curvature in sub-Riemannian geometry*, Houston J. Math. **38** (2012), no. 4, 1107–1134.

[94] J.-Q. Hu and H.-Q. Li, *Gradient estimates for the heat semigroup on H-type groups*, Potential Anal. **33** (2010), no. 4, 355–386.

[95] Y. Huang and S. Sun, *Non-embedding theorems of nilpotent Lie groups and sub-Riemannian manifolds*, Front. Math. China **15** (2020), no. 1, 91–114.

[96] B. Khesin and P. Lee, *A nonholonomic Moser theorem and optimal transport*, J. Symplectic Geom. **7** (2009), no. 4, 381–414.

[97] K. Kuwada, *Duality on gradient estimates and Wasserstein controls*, J. Funct. Anal. **258** (2010), no. 11, 3758–3774.

[98] _____, *Gradient estimate for Markov kernels, Wasserstein control and Hopf-Lax formula*, Potential theory and its related fields, RIMS Kôkyûroku Bessatsu, B43, Res. Inst. Math. Sci. (RIMS), Kyoto, 2013, pp. 61–80.

[99] _____, *Space-time Wasserstein controls and Bakry-Ledoux type gradient estimates*, Calc. Var. Partial Differential Equations **54** (2015), no. 1, 127–161.

[100] R. Jordan, D. Kinderlehrer, and F. Otto, *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal. **29** (1998), no. 1, 1–17.

[101] N. Juillet, *Geometric inequalities and generalized Ricci bounds in the Heisenberg group*, Int. Math. Res. Not. IMRN **13** (2009), 2347–2373.

[102] _____, *Diffusion by optimal transport in Heisenberg groups*, Calc. Var. Partial Differential Equations **50** (2014), no. 3-4, 693–721.

[103] _____, *Sub-Riemannian structures do not satisfy Riemannian Brunn-Minkowski inequalities*, Rev. Mat. Iberoam. **37** (2021), no. 1, 177–188.

[104] E. Le Donne, *A primer on Carnot groups: homogenous groups, Carnot-Carathéodory spaces, and regularity of their isometries*, Anal. Geom. Metr. Spaces **5** (2017), 116–137.

[105] E. Le Donne, D. Lučić, and E. Pasqualetto, *Universal infinitesimal Hilbertianity property of sub-Riemannian manifolds* (2019), preprint, available at `https://arxiv.org/abs/1910.05962`.

[106] P. Lee, C. Li, and I. Zelenko, *Ricci curvature type lower bounds for sub-Riemannian structures on Sasakian manifolds*, Discrete Contin. Dyn. Syst. **36** (2016), no. 1, 303–321.

[107] H.-Q. Li, *Estimation optimale du gradient du semi-groupe de la chaleur sur le groupe de Heisenberg*, J. Funct. Anal. **236** (2006), no. 2, 369–394.

[108] S. Lisini, *Characterization of absolutely continuous curves in Wasserstein spaces*, Calc. Var. Partial Differential Equations **28** (2007), no. 1, 85–120.

[109] J. Lott and C. Villani, *Weak curvature conditions and functional inequalities*, J. Funct. Anal. **245** (2007), no. 1, 311–333.

[110] ———, *Ricci curvature for metric-measure spaces via optimal transport*, Ann. of Math. (2) **169** (2009), no. 3, 903–991.

[111] D. Lučić and E. Pasqualetto, *Infinitesimal Hilbertianity of weighted Riemannian manifolds*, Canad. Math. Bull. **63** (2020), no. 1, 118–140.

[112] T. Melcher, *Hypoelliptic heat kernel inequalities on Lie groups*, Stochastic Process. Appl. **118** (2008), no. 3, 368–388.

[113] T. Mietton and L. Rizzi, *Branching geodesics in sub-Riemannian Geometry*, Geom. Funct. Anal. (2020), in press.

[114] E. Milman, *The Quasi Curvature-Dimension Condition with applications to sub-Riemannian manifolds*, Comm. Pure Appl. Math. (2019), to appear, available at `https://arxiv.org/abs/1908.01513`.

[115] J. Mitchell, *On Carnot-Carathéodory metrics*, J. Differential Geom. **21** (1985), no. 1, 35–45.

[116] R. Montgomery, *A tour of subriemannian geometries, their geodesics and applications*, Mathematical Surveys and Monographs, vol. 91, American Mathematical Society, Providence, RI, 2002.

[117] I. H. Munive, *Sub-Riemannian curvature of Carnot groups with rank-two distributions*, J. Dyn. Control Syst. **23** (2017), no. 4, 779–814.

[118] S.-I. Ohta, *On the measure contraction property of metric measure spaces*, Comment. Math. Helv. **82** (2007), no. 4, 805–828.

[119] ———, *(K, N)-convexity and the curvature-dimension condition for negative N*, J. Geom. Anal. **26** (2016), no. 3, 2067–2096.

[120] S.-I. Ohta and K.-T. Sturm, *Heat flow on Finsler manifolds*, Comm. Pure Appl. Math. **62** (2009), no. 10, 1386–1433.

[121] ———, *Bochner-Weitzenböck formula and Li-Yau estimates on Finsler manifolds*, Adv. Math. **252** (2014), 429–448.

[122] F. Otto, *The geometry of dissipative evolution equations: the porous medium equation*, Comm. Partial Differential Equations **26** (2001), no. 1-2, 101–174.

[123] F. Otto and C. Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal. **173** (2000), no. 2, 361–400.

[124] F. Otto and M. Westdickenberg, *Eulerian calculus for the contraction in the Wasserstein distance*, SIAM J. Math. Anal. **37** (2005), no. 4, 1227–1255.

[125] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Applied Mathematical Sciences, vol. 44, Springer-Verlag, New York, 1983.

[126] T. Rajala, *Local Poincaré inequalities from stable curvature conditions on metric spaces*, Calc. Var. Partial Differential Equations **44** (2012), no. 3-4, 477–494.

[127] M.-K. von Renesse and K.-T. Sturm, *Transport inequalities, gradient estimates, entropy, and Ricci curvature*, Comm. Pure Appl. Math. **58** (2005), no. 7, 923–940.

[128] S. Rigot, *Mass transportation in groups of type H*, Commun. Contemp. Math. **7** (2005), no. 4, 509–537.

[129] L. Rizzi, *Measure contraction properties of Carnot groups*, Calc. Var. Partial Differential Equations **55** (2016), no. 3, Art. 60, 20.

[130] L. Rizzi and P. Silveira, *Sub-Riemannian Ricci curvatures and universal diameter bounds for 3-Sasakian manifolds*, J. Inst. Math. Jussieu **18** (2019), no. 4, 783–827.

[131] G. Savaré, *Self-improvement of the Bakry-Émery condition and Wasserstein contraction of the heat flow in* RCD(K, ∞) *metric measure spaces*, Discrete Contin. Dyn. Syst. **34** (2014), no. 4, 1641–1661.

[132] P. Stollmann, *A dual characterization of length spaces with application to Dirichlet metric spaces*, Studia Math. **198** (2010), no. 3, 221–233.

[133] R. S. Strichartz, *Sub-Riemannian geometry*, J. Differential Geom. **24** (1986), no. 2, 221–263.

[134] ———, *Corrections to: "Sub-Riemannian geometry" [J. Differential Geom. **24** (1986), no. 2, 221–263]*, J. Differential Geom. **30** (1989), no. 2, 595–596.

[135] K.-T. Sturm, *Analysis on local Dirichlet spaces. II. Upper Gaussian estimates for the fundamental solutions of parabolic equations*, Osaka J. Math. **32** (1995), no. 2, 275–312.

[136] _____ , *The geometric aspect of Dirichlet forms*, New directions in Dirichlet forms, AMS/IP Stud. Adv. Math., vol. 8, Amer. Math. Soc., Providence, RI, 1998, pp. 233–277.

[137] _____ , *On the geometry of metric measure spaces. I*, Acta Math. **196** (2006), no. 1, 65–131.

[138] _____ , *On the geometry of metric measure spaces. II*, Acta Math. **196** (2006), no. 1, 133–177.

[139] _____ , *Distribution-valued Ricci bounds for metric measure spaces, singular time changes, and gradient estimates for Neumann heat flows*, Geom. Funct. Anal. **30** (2020), no. 6, 1648–1711.

[140] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003.

[141] _____ , *Optimal transport, old and new*, Fundamental Principles of Mathematical Sciences, vol. 338, Springer-Verlag, Berlin, 2009.

[142] _____ , *Inégalités isopérimétriques dans les espaces métriques mesurés [d'après F. Cavalletti & A. Mondino]*, Astérisque **407** (2019), Exp. No. 1127, 213–265 (French). Séminaire Bourbaki. Vol. 2016/2017. Exposés 1120–1135.

[143] F.-Y. Wang, *Equivalent semigroup properties for the curvature-dimension condition*, Bull. Sci. Math. **135** (2011), no. 6-7, 803–815.

[144] _____ , *Exponential contraction in Wasserstein distances for diffusion semigroups with negative curvature*, Potential Anal. **53** (2020), no. 3, 1123–1144.

[145] _____ , *Derivative formula and gradient estimates for Gruschin type semigroups*, J. Theoret. Probab. **27** (2014), no. 1, 80–95.

[146] I. Zelenko and C. Li, *Differential geometry of curves in Lagrange Grassmannians with given Young diagram*, Differential Geom. Appl. **27** (2009), no. 6, 723–742.

[147] H.-C. Zhang and X.-P. Zhu, *Yau's gradient estimates on Alexandrov spaces*, J. Differential Geom. **91** (2012), no. 3, 445–522.

Department Mathematik und Informatik, Universität Basel, Spiegelgasse 1, CH-4051 Basel, Switzerland

Email address: giorgio.stefani@unibas.ch