J. Math. Pures Appl. 147 (2021) 60–97

Contents lists available at ScienceDirect

Journal de Mathématiques Pures et Appliquées

www.elsevier.com/locate/matpur

Optimal transport, gradient estimates, and pathwise Brownian coupling on spaces with variable Ricci bounds

Mathias Braun [a,1], Karen Habermann [b,2], Karl-Theodor Sturm [a,*,3]

[a] *University of Bonn, Institute for Applied Mathematics, Endenicher Allee 60, 53115 Bonn, Germany*
[b] *Department of Statistics, University of Warwick, Coventry, CV4 7AL, United Kingdom*

ARTICLE INFO

Article history:
Received 11 November 2019
Available online 15 January 2021

MSC:
58J65
49Q22
60J60
35K05
58J35
53C23
31C25
60J46

Keywords:
Ricci curvature
Bakry–Émery estimate
Gradient estimate
Optimal transport
Coupling

ABSTRACT

Given a metric measure space $(X, \mathsf{d}, \mathfrak{m})$ and a lower semicontinuous, lower bounded function $k\colon X \to \mathbb{R}$, we prove the equivalence of the synthetic approaches to *Ricci curvature at $x \in X$ being bounded from below by $k(x)$* in terms of

- the Bakry–Émery estimate $\Delta\Gamma(f)/2 - \Gamma(f, \Delta f) \geq k\,\Gamma(f)$ in an appropriate weak formulation, and
- the curvature-dimension condition $\mathrm{CD}(k, \infty)$ in the sense of Lott–Sturm–Villani with variable k.

Moreover, for all $p \in (1, \infty)$, these properties hold if and only if the perturbed p-transport cost

$$W_p^{\underline{k}}(\mu_1, \mu_2, t) := \inf_{(\mathsf{b}^1, \mathsf{b}^2)} \mathbb{E}\Big[\mathrm{e}^{\int_0^{2t} p\underline{k}(\mathsf{b}_r^1, \mathsf{b}_r^2)/2\,\mathrm{d}r}\, \mathsf{d}^p\big(\mathsf{b}_{2t}^1, \mathsf{b}_{2t}^2\big)\Big]^{1/p}$$

is nonincreasing in t. The infimum here is taken over pairs of coupled Brownian motions b^1 and b^2 on X with given initial distributions μ_1 and μ_2, respectively, and $\underline{k}(x, y) := \inf_\gamma \int_0^1 k(\gamma_s)\,\mathrm{d}s$ denotes the "average" of k along geodesics γ connecting x and y.

Furthermore, for any pair of initial distributions μ_1 and μ_2 on X, we prove the existence of a pair of coupled Brownian motions b^1 and b^2 such that a.s. for every $s, t \in [0, \infty)$ with $s \leq t$, we have

$$\mathsf{d}\big(\mathsf{b}_t^1, \mathsf{b}_t^2\big) \leq \mathrm{e}^{-\int_s^t \underline{k}(\mathsf{b}_r^1, \mathsf{b}_r^2)/2\,\mathrm{d}r}\, \mathsf{d}\big(\mathsf{b}_s^1, \mathsf{b}_s^2\big).$$

* Corresponding author.
E-mail addresses: braun@iam.uni-bonn.de (M. Braun), karen.habermann@warwick.ac.uk (K. Habermann), sturm@uni-bonn.de (K.-T. Sturm).

[1] Funded by the European Research Council through the ERC-AdG "RicciBounds" ERC project 10760021.
[2] Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - GZ 2047/1, Projekt-ID 390685813.
[3] Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - GZ 2047/1, Projekt-ID 390685813 as well as through the Collaborative Research Center 1060, and funded by the European Research Council through the ERC-AdG "RicciBounds", ERC project 10760021.

https://doi.org/10.1016/j.matpur.2021.01.002
0021-7824/

R É S U M É

Étant donnés un espace métrique mesuré $(X, \mathsf{d}, \mathfrak{m})$ et une fonction $k\colon X \to \mathbb{R}$ semi-continue inférieurement et minorée, nous prouvons l'équivalence des approches synthétiques pour la *courbure de Ricci en $x \in X$ étant minorée par $k(x)$* exprimées en termes

- de l'estimée de Bakry–Émery $\Delta\Gamma(f)/2 - \Gamma(f, \Delta f) \geq k\,\Gamma(f)$ formulée en un sens faible approprié et
- de la condition $\mathrm{CD}(k, \infty)$ au sens de Lott–Sturm–Villani avec k variable.

De plus, pour tout $p \in (1, \infty)$, ces propriétés sont vérifiées si et seulement si le p-coût de transport perturbé

$$W_p^{\underline{k}}(\mu_1, \mu_2, t) := \inf_{(\mathsf{b}^1, \mathsf{b}^2)} \mathbb{E}\Big[\mathrm{e}^{\int_0^{2t} p\underline{k}(\mathsf{b}_r^1, \mathsf{b}_r^2)/2\,\mathrm{d}r}\,\mathsf{d}^p\big(\mathsf{b}_{2t}^1, \mathsf{b}_{2t}^2\big)\Big]^{1/p}$$

est décroissant en t. Ici, l'infimum est pris sur toutes les paires de mouvements browniens couplés b^1 et b^2 sur X avec lois initiales respectives μ_1 et μ_2 et $\underline{k}(x, y) := \inf_\gamma \int_0^1 k(\gamma_s)\,\mathrm{d}s$ désigne la "moyenne" de k le long des géodésiques γ reliant x et y. En outre, pour toute paire de lois initiales μ_1 et μ_2 sur X, nous prouvons l'existence d'une paire de mouvements browniens couplés b^1 et b^2 telle que presque sûrement pour tous $s, t \in [0, \infty)$ avec $s \leq t$, on a

$$\mathsf{d}\big(\mathsf{b}_t^1, \mathsf{b}_t^2\big) \leq \mathrm{e}^{-\int_s^t \underline{k}(\mathsf{b}_r^1, \mathsf{b}_r^2)/2\,\mathrm{d}r}\,\mathsf{d}\big(\mathsf{b}_s^1, \mathsf{b}_s^2\big).$$

Contents

1. Introduction

Throughout this paper, the triple $(X, \mathsf{d}, \mathfrak{m})$ is a *metric measure space*, that is, a complete and separable metric space (X, d) equipped with a locally finite measure $\mathfrak{m}$ defined on the Borel σ-field $\mathscr{B}(X)$, and

$k\colon X \to \mathbb{R}$ is a lower semicontinuous function which is bounded from below. For simplicity, we assume that $\mathfrak{m}$ has full topological support. We say that $(X, \mathsf{d}, \mathfrak{m})$ is an RCD space if it satisfies the $\mathrm{RCD}(K, \infty)$ condition for some $K \in \mathbb{R}$. This will be our standing assumption throughout.

Denote by $\mathscr{P}(X)$ the space of Borel probability measures on (X, d). For $p \in [1, \infty)$, $\mathscr{P}_p(X)$ is the set of $\mu \in \mathscr{P}(X)$ with $\int_X \mathsf{d}^p(x, y)\,\mathrm{d}\mu(y) < \infty$ for some $x \in X$. As usual, W_p denotes the p-Kantorovich–Wasserstein distance defined through

$$W_p(\mu, \nu) := \inf_{\pi} \Big(\int_{X \times X} \mathsf{d}^p(x, y)\,\mathrm{d}\pi(x, y) \Big)^{1/p},$$

where the infimum is taken over all $\pi \in \mathscr{P}(X \times X)$ with marginals μ and ν. If it exists, the limit $|\dot{\gamma}_t| := \lim_{h \to 0} \mathsf{d}(\gamma_{t+h}, \gamma_t)/|h|$ is called *metric speed* of the curve $\gamma \in \mathrm{C}([0, 1]; X)$ at $t \in [0, 1]$, and we write $|\dot{\gamma}|$ if $|\dot{\gamma}_t| = |\dot{\gamma}_s|$ for every $s, t \in [0, 1]$. Moreover, $\mathrm{Geo}(X)$ denotes the space of *geodesics* on X, i.e. the set of $\gamma \in \mathrm{C}([0, 1]; X)$ with $\mathsf{d}(\gamma_t, \gamma_s) = |t - s|\ \mathsf{d}(\gamma_0, \gamma_1)$ for all $s, t \in [0, 1]$. Similarly, we define $\mathrm{Geo}(\mathscr{P}_p(X))$ as the space of W_p-geodesics in the space of probability measures. We say that $\boldsymbol{\pi} \in \mathscr{P}(\mathrm{Geo}(X))$ *represents the* W_p*-geodesic* $(\mu_t)_{t \in [0,1]}$ if $\mu_t = (\mathsf{e}_t)_\sharp \boldsymbol{\pi}$ for all $t \in [0, 1]$, where $\mathsf{e}_t\colon \mathrm{C}([0, 1]; X) \to X$ is the evaluation map defined by $\mathsf{e}_t(\gamma) := \gamma_t$. By [23], every W_p-geodesic can be represented by some $\boldsymbol{\pi} \in \mathscr{P}(\mathrm{Geo}(X))$.

We present various synthetic approaches to the definition of *Ricci curvature at* $x \in X$ *bounded from below by* $k(x)$ and prove their equivalence. These characterizations are suitable extensions of the curvature-dimension condition, the evolution variational inequality, Bochner's inequality, gradient estimates and transport estimates to nonconstant curvature bounds. To this list, we add a description in terms of pathwise coupling of Brownian motions. In total, our main result is the following.

Theorem 1.1. *Let* $(X, \mathsf{d}, \mathfrak{m})$ *be an* RCD *space, and let* $k\colon X \to \mathbb{R}$ *be a lower semicontinuous, lower bounded function. For all exponents* $p \in (1, \infty)$ *and* $q \in [1, \infty)$*, the following properties are equivalent:*

- (i) *the curvature-dimension condition* $\mathrm{CD}(k, \infty)$,
- (ii) *the evolution variational inequality* $\mathrm{EVI}(k)$,
- (iii) *the* q*-Bochner inequality* $\mathrm{BE}_q(k, \infty)$,
- (iv) *the* q*-gradient estimate* $\mathrm{GE}_q(k)$,
- (v) *the* p*-transport estimate* $\mathrm{TE}_p(k)$, *and*
- (vi) *the pathwise coupling property* $\mathrm{PCP}(k)$.

Moreover, any of these properties yields (iii), (iv) *and* (v) *for all exponents* $p, q \in [1, \infty)$.

Let us now introduce each of these extensions and give an overview of the organization of our reasoning. Throughout, we assume the reader to be familiar with the theory of $\mathrm{RCD}(K, \infty)$ spaces and basic properties of these. An account on this will be collected in Section 2 which can be read independently of the rest of this paper.

1.1. Lagrangian formulation of synthetic variable Ricci bounds

Here and in the sequel, $\mathrm{g}(s, t) := \min\{s(1 - t), t(1 - s)\}$ denotes the Green's function of the unit interval $[0, 1]$. Define the *Boltzmann entropy* $\mathrm{Ent}_{\mathfrak{m}}\colon \mathscr{P}_2(X) \to (-\infty, \infty]$ as

$$\mathrm{Ent}_{\mathfrak{m}}(\mu) := \int_X \rho \log \rho \,\mathrm{d}\mathfrak{m} \quad \text{if } \mu \ll \mathfrak{m} \text{ with } \mu = \rho\,\mathfrak{m}, \quad \mathrm{Ent}_{\mathfrak{m}}(\mu) := \infty \quad \text{otherwise.}$$

We put $\mathrm{Dom}(\mathrm{Ent}_\mathfrak{m}) := \{\mu \in \mathscr{P}_2(X) : \mathrm{Ent}_\mathfrak{m}(\mu) \in \mathbb{R}\}$. Convexity properties of the Boltzmann entropy are at the center of the curvature-dimension condition $\mathrm{CD}(K,\infty)$ introduced in [29,24] (as well as of its enforcements in [30,8,14]) which we are extend now from fixed K to variable k.

Definition 1.2 [31, Definition 3.2]. *An* RCD *space* $(X,\mathsf{d},\mathfrak{m})$ *is said to satisfy the* curvature-dimension condition *with variable curvature bound k, briefly* $\mathrm{CD}(k,\infty)$*, if for every $\mu_0,\mu_1 \in \mathrm{Dom}(\mathrm{Ent}_\mathfrak{m})$ there exists a measure $\boldsymbol{\pi} \in \mathscr{P}(\mathrm{Geo}(X))$ representing some W_2-geodesic $(\mu_t)_{t\in[0,1]}$ connecting μ_0 and μ_1 such that, for all $t \in [0,1]$,*

$$\mathrm{Ent}_\mathfrak{m}(\mu_t) \leq (1-t)\,\mathrm{Ent}_\mathfrak{m}(\mu_0) + t\,\mathrm{Ent}_\mathfrak{m}(\mu_1) - \int_0^1 \int_{\mathrm{Geo}(X)} \mathrm{g}(s,t)\,k(\gamma_s)\,|\dot\gamma|^2\,\mathrm{d}\boldsymbol{\pi}(\gamma)\,\mathrm{d}s.$$

Definition 1.3 [31, Definition 3.3]. *An* RCD *space* $(X,\mathsf{d},\mathfrak{m})$ *is said to satisfy the* evolution variational inequality *with variable curvature bound k, briefly* $\mathrm{EVI}(k)$*, if for every $\mu_0 \in \mathscr{P}_2(X)$ there exists a locally absolutely continuous curve $(\mu_t)_{t>0}$ in $\mathrm{Dom}(\mathrm{Ent}_\mathfrak{m})$ with $W_2(\mu_t,\mu_0) \to 0$ as $t \to 0$, and for every $t > 0$ and $\nu \in \mathscr{P}_2(X)$ there exists a measure $\boldsymbol{\pi}_t \in \mathscr{P}(\mathrm{Geo}(X))$ representing some W_2-geodesic connecting μ_t and ν such that*

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\frac{1}{2}W_2^2(\mu_t,\nu) + \int_0^1 \int_{\mathrm{Geo}(X)} (1-s)\,k(\gamma_s)\,|\dot\gamma|^2\,\mathrm{d}\boldsymbol{\pi}_t(\gamma)\,\mathrm{d}s \leq \mathrm{Ent}_\mathfrak{m}(\nu) - \mathrm{Ent}_\mathfrak{m}(\mu_t).$$

From [31, Theorem 3.4], it is already known that $\mathrm{CD}(k,\infty)$ is equivalent to $\mathrm{EVI}(k)$ on RCD spaces, which establishes the equivalence of (i) and (ii) in Theorem 1.1.

1.2. Eulerian formulation of synthetic variable Ricci bounds

Let us now switch to the Eulerian picture which, to shorten the presentation, is directly presented for arbitrary exponents. Define the *Cheeger energy* $\mathscr{E}\colon L^2(X,\mathfrak{m}) \to [0,\infty]$ as

$$\mathscr{E}(f) := \inf\Big\{\liminf_{n\to\infty} \int_X \mathrm{lip}(f_n)^2\,\mathrm{d}\mathfrak{m} : f_n \in \mathrm{Lip}_\mathrm{b}(X),\ f_n \to f \text{ in } L^2(X,\mathfrak{m})\Big\},$$

where $\mathrm{lip}(f)(x) := \limsup_{y\to x} |f(x)-f(y)|/\mathsf{d}(x,y)$ denotes the *local Lipschitz slope* at $x \in X$. We put $\mathrm{Dom}(\mathscr{E}) := \big\{f \in L^2(X,\mathfrak{m}) : \mathscr{E}(f) < \infty\big\}$.

Definition 1.4. Given $q \in [1,\infty)$, we say that an RCD space $(X,\mathsf{d},\mathfrak{m})$ satisfies the *q-Bochner inequality* or *q-Bakry–Émery estimate* with variable curvature bound k, briefly $\mathrm{BE}_q(k,\infty)$, if

$$\int_X \Big(\frac{1}{q}\Gamma(f)^{q/2}\,\Delta\phi - \Gamma(f)^{q/2-1}\,\Gamma(f,\Delta f)\,\phi\Big)\,\mathrm{d}\mathfrak{m} \geq \int_X k\,\Gamma(f)^{q/2}\,\phi\,\mathrm{d}\mathfrak{m}$$

holds for all $f \in \mathrm{Dom}(\Delta)$ with $\Delta f \in \mathrm{Dom}(\mathscr{E})$ as well as $\Gamma(f) \in L^\infty(X,\mathfrak{m})$ and for every nonnegative $\phi \in \mathrm{Dom}(\Delta) \cap L^\infty(X,\mathfrak{m})$ with $\Delta\phi \in L^\infty(X,\mathfrak{m})$.

The equivalence of (i) and (iii) for $q = 2$ in our major Theorem 1.1 above states that the variable Eulerian and Lagrangian approaches to synthetic lower Ricci bounds coincide, i.e. $\mathrm{CD}(k,\infty)$ is equivalent to $\mathrm{BE}_2(k,\infty)$. If k is constant, this has been proved by Ambrosio, Gigli and Savaré in their groundbreaking works, see [4], which follows [15], for (i) implying (iii), and [5] for (iii) implying (i). In the nonconstant case, this remained open in previous contributions [20,21,31].

The implication from $\mathrm{BE}_2(k,\infty)$ to $\mathrm{CD}(k,\infty)$ follows from Theorem 3.4 and Theorem 4.5. The proof of the converse is a consequence of Proposition 4.6, Proposition 5.6, Theorem 5.19 and eventually Theorem 3.4. This requires a detailed heat flow analysis, both at the level of functions and measures, and in particular an extension of Kuwada's duality [22, Theorem 2.2] between q-gradient estimates and p-transport estimates for dual p and q. This is quite demanding – indeed, until now not even a formulation of an appropriate p-transport estimate with nonconstant curvature bound existed.

The "self-improvement property" of the q-Bochner inequality will be another key result. Indeed, the $\mathrm{BE}_q(k,\infty)$ condition is *independent* of q, see Theorem 3.6, which provides the equivalence of (i) and (iii) in Theorem 1.1 for general q.

1.3. Improved gradient estimates

Following [31], let $(\mathsf{P}_t^{qk})_{t\geq 0}$ be the Schrödinger semigroup on $L^2(X,\mathfrak{m})$ associated to the generator $\Delta - qk$ for $q\in[1,\infty)$. It extends to a strongly continuous semigroup on $L^r(X,\mathfrak{m})$ for each $r\in[1,\infty)$. In terms of the Brownian motion $(\mathbb{P}_x,\mathsf{b})$ on X starting in $x\in X$, it can be expressed through the Feynman–Kac formula

$$\mathsf{P}_t^{qk} f(x) = \mathbb{E}_x\Big[\mathrm{e}^{-\int_0^{2t} qk(\mathsf{b}_r)/2\,\mathrm{d}r}\, f(\mathsf{b}_{2t})\Big] \quad \text{for every } f\in L^r(X,\mathfrak{m}). \tag{1.1}$$

Definition 1.5. We say that a *q-gradient estimate* with variable curvature bound k, briefly $\mathrm{GE}_q(k)$, holds whenever

$$\Gamma(\mathsf{P}_t f)^{q/2} \leq \mathsf{P}_t^{qk}\big(\Gamma(f)^{q/2}\big) \quad \mathfrak{m}\text{-a.e.}$$

is satisfied for every $f\in\mathrm{Dom}(\mathscr{E})$ and every $t\geq 0$.

Adapting the well-known arguments for constant Ricci curvature bounds from [10,27], we establish, as stated in Theorem 3.4, that $\mathrm{BE}_q(k,\infty)$ holds if and only if $\mathrm{GE}_q(k)$ is satisfied. This yields the equivalence of (iii) and (iv) in Theorem 1.1 for general $q\in[1,\infty)$.

1.4. Variable transport estimates

In order to formulate a dual p-transport estimate for $p\in[1,\infty)$, we consider evolutions on the product space $X\times X$. Denoting by $\mathrm{G}_\varepsilon(x,y)$ the set of $\gamma\in\mathrm{Geo}(X)$ with $\gamma_0\in\overline{B}_\varepsilon(x)$ and $\gamma_1\in\overline{B}_\varepsilon(y)$, we introduce the function $\underline{k}\colon X\times X\to\mathbb{R}$ defined by

$$\underline{k}(x,y) := \lim_{\varepsilon\to 0}\ \inf_{\gamma\in\mathrm{G}_\varepsilon(x,y)} \int_0^1 k(\gamma_s)\,\mathrm{d}s. \tag{1.2}$$

Its basic properties are summarized in Section 2. As we will see in Remark 5.12, Theorem 6.1 and Theorem 5.17, it turns out that $\underline{k}$ can indeed equivalently be replaced in all relevant quantities by the larger function $\overline{k}\colon X\times X\to\mathbb{R}$ defined by

$$\overline{k}(x,y) := \liminf_{(x_n,y_n)\to(x,y)}\ \sup_{\gamma\in\mathrm{G}_0(x_n,y_n)} \int_0^1 k(\gamma_s)\,\mathrm{d}s. \tag{1.3}$$

Definition 1.6. A pair $\big((\mathsf{b}^1_t)_{t\geq 0}, (\mathsf{b}^2_t)_{t\geq 0}\big)$ of stochastic processes on X is called *coupling of Brownian motions* if it is defined on a common probability space $(\Omega, \mathbb{P})$ and each of the processes $(\mathsf{b}^1_t)_{t\geq 0}$ and $(\mathsf{b}^2_t)_{t\geq 0}$ is a Brownian motion on the RCD space $(X, \mathsf{d}, \mathfrak{m})$.

Given $\mu_1, \mu_2 \in \mathscr{P}_p(X)$, we define the *perturbed p-transport cost* at time $t \geq 0$ by

$$W_p^{\underline{k}}(\mu_1, \mu_2, t) := \inf_{(\mathbb{P}, \mathsf{b}^1, \mathsf{b}^2)} \mathbb{E}\Big[\mathrm{e}^{\int_0^{2t} p\underline{k}(\mathsf{b}^1_r, \mathsf{b}^2_r)/2\,\mathrm{d}r}\, \mathsf{d}^p\big(\mathsf{b}^1_{2t}, \mathsf{b}^2_{2t}\big)\Big]^{1/p},$$

where the infimum is taken over all pairs of coupled Brownian motions $\big(\mathbb{P}, \mathsf{b}^1\big)$ and $\big(\mathbb{P}, \mathsf{b}^2\big)$ on X, restricted to $[0, 2t]$ and modeled on a common probability space, with initial distributions μ_1 and μ_2, respectively. Note that $W_p^{\underline{k}}(\mu_1, \mu_2, 0) = W_p(\mu_1, \mu_2)$ and that for general $t \geq 0$, if k is constant, say $k = K$, the perturbed p-transport cost can be expressed in terms of the usual p-transport cost via

$$W_p^{\underline{k}}(\mu_1, \mu_2, t) = \mathrm{e}^{Kt}\, W_p(\mathsf{H}_t\mu_1, \mathsf{H}_t\mu_2).$$

Definition 1.7. Given any $p \in [1, \infty)$, we say that a *p-transport estimate* with variable curvature bound k, briefly $\mathrm{TE}_p(k)$, holds if the map $t \mapsto W_p^{\underline{k}}(\mu_1, \mu_2, t)$ is nonincreasing on $[0, \infty)$ for every pair $\mu_1, \mu_2 \in \mathscr{P}_p(X)$.

Having at our disposal appropriate replacements for the expressions $\mathrm{e}^{-qKt}\, \mathsf{P}_t\big(\Gamma(f)^{q/2}\big)$ and $\mathrm{e}^{Kt}\, W_p(\mathsf{H}_t\mu_1, \mathsf{H}_t\mu_2)$ in terms of Feynman–Kac formulas with potentials qk for the Brownian motion on X and $-p\underline{k}$ for pairs of coupled Brownian motions on $X \times X$, respectively, we are in a position to formulate and prove a generalization of the fundamental Kuwada duality in the case of nonconstant k. This addresses the equivalence of (iv) and (v) in Theorem 1.1.

Theorem 1.8. *For every $p, q \in (1, \infty)$ with $1/p + 1/q = 1$, the following are equivalent:*

(iv) *the q-gradient estimate $\mathrm{GE}_q(k)$, and*
(v) *the p-transport estimate $\mathrm{TE}_p(k)$.*

This result is a consequence of Theorem 5.16 and Theorem 5.19. For both results, it is crucial to use a localization argument in regions where k or $\underline{k}$ are "approximately constant" and then use tail estimates for Brownian paths to control the remainder terms.

Suitable extensions to the case $q = 1$ and $p = \infty$ will be discussed, and eventually shown to be equivalent, in Theorem 5.10, Theorem 5.17 and Theorem 6.1. Therefore, making sense of an appropriate $\mathrm{TE}_p(k)$ condition for $p = \infty$ is the content of the subsequent Section 1.5.

Remark 1.9. It is often convenient to use the characterization of $\mathrm{TE}_p(k)$, which is zeroth-order in nature, through a first-order condition via the *differential p-transport inequality*

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0} W_p^p(\mathsf{H}_t\delta_x, \mathsf{H}_t\delta_y) \leq -p\,\underline{k}(x, y)\, \mathsf{d}^p(x, y) \quad \text{for every } x, y \in X,$$

very much in the spirit of the connection between $\mathrm{BE}_q(k, \infty)$ and $\mathrm{GE}_q(k)$. The equivalence of $\mathrm{TE}_p(k)$ and the foregoing estimate, which for constant k is essentially Gronwall's lemma and a standard coupling technique, is treated in Theorem 5.7.

A posteriori, for every $p \in (1, \infty)$, any of the conditions (i) to (vi) from Theorem 1.1 will indeed give the stronger estimate

$$\frac{\mathrm{d}^+}{\mathrm{d}t} W_p^p(\mathsf{H}_t\mu_1, \mathsf{H}_t\mu_2) \leq -p \int_0^1 \int_{\mathrm{Geo}(X)} k(\gamma_s)\, |\dot\gamma|^p \,\mathrm{d}\boldsymbol{\pi}_t(\gamma)\,\mathrm{d}s \quad \text{for every } t \geq 0,$$

where $\mu_1, \mu_2 \in \mathscr{P}(X)$ have finite W_p-distance to each other, and $\boldsymbol{\pi}_t \in \mathscr{P}(\mathrm{Geo}(X))$ is an *arbitrary* measure representing a W_p-geodesic from $\mathsf{H}_t\mu_1$ to $\mathsf{H}_t\mu_2$, see Corollary 5.11. ■

1.5. Pathwise coupling of Brownian motions

Finally, we reinforce the p-transport estimate by passing to the limit $p \to \infty$ and by replacing the mean value estimates by a pathwise one.

Definition 1.10. We say that the *pathwise coupling property* with variable curvature bound k, briefly $\mathrm{PCP}(k)$, holds if for every pair $\mu_1, \mu_2 \in \mathscr{P}(X)$ there exists a pair $(\mathbb{P}, \mathsf{b}^1)$ and $(\mathbb{P}, \mathsf{b}^2)$ of coupled Brownian motions on X with initial distributions μ_1 and μ_2, respectively, such that $\mathbb{P}$-a.s., we have

$$\mathsf{d}\big(\mathsf{b}_t^1, \mathsf{b}_t^2\big) \leq \mathrm{e}^{-\int_s^t \underline{k}(\mathsf{b}_r^1, \mathsf{b}_r^2)/2\,\mathrm{d}r}\, \mathsf{d}\big(\mathsf{b}_s^1, \mathsf{b}_s^2\big) \quad \text{for every } s, t \in [0, \infty) \text{ with } s \leq t.$$

It is proved in [7, Theorem 4.1] that complete Riemannian manifolds with Ricci curvature bounded from below by $K \in \mathbb{R}$ satisfy $\mathrm{PCP}(k)$ with constant $k = K$. The work [31, Theorem 2.9] extended this to general $\mathrm{RCD}(K, \infty)$ spaces. A first result into the nonconstant direction is due to [32, Theorem 6]. Again on Riemannian manifolds with a uniform lower bound on the Ricci curvature, it deduces the existence of a pair $(\mathsf{b}^1, \mathsf{b}^2)$ of coupled Brownian motions starting in (x, y) obeying for every $t \geq 0$, on the event that $(\mathsf{b}_r^1, \mathsf{b}_r^2)$ does not belong to the cut-locus of X for all $r \in [0, t]$, the estimate

$$\mathsf{d}\big(\mathsf{b}_t^1, \mathsf{b}_t^2\big) \leq \mathrm{e}^{-\int_0^t \kappa(\mathsf{b}_r^1, \mathsf{b}_r^2)/2\,\mathrm{d}r}\, \mathsf{d}(x, y),$$

where $\kappa(x, y) := -\frac{\mathrm{d}^+}{\mathrm{d}t}\big|_{t=0} \log W_1(\mathsf{H}_t\delta_x, \mathsf{H}_t\delta_y)$ denotes the *coarse curvature* at $x, y \in X$, $x \neq y$. For x, y close to each other, say $y = \exp_x(\varepsilon v)$ with $\varepsilon > 0$, $v \in T_xX$, we have

$$\kappa(x, y) = \mathrm{Ric}_x(v, v) + \mathrm{o}(1),$$

see [32, Theorem 19 and Remark 20]. The construction of this process deeply relies on smooth calculus tools, which are unavailable in our setting and thus cannot be adopted.

Our main theorem extends these results in terms of $\underline{k}$ and circumvents regularity issues involving the variable curvature bound. The existence of a process satisfying the $\mathrm{PCP}(k)$ condition is even equivalent to $\mathrm{CD}(k, \infty)$. Indeed, given $\mathrm{TE}_p(k)$ for every large enough $p \in (1, \infty)$, we deduce $\mathrm{PCP}(k)$ by means of Theorem 6.1, the content of which is the implication from (v) to (vi) in Theorem 1.1. Note that according to the previous Theorem 1.8 and nestedness of q-gradient estimates, see Lemma 3.3, the 1-gradient estimate $\mathrm{GE}_1(k)$ implies $\mathrm{TE}_p(k)$ for all $p \in (1, \infty)$ and thus $\mathrm{PCP}(k)$. The converse of this, i.e. the implication from $\mathrm{PCP}(k)$ to $\mathrm{GE}_1(k)$, is addressed in Theorem 5.17.

Acknowledgments The authors warmly thank Matthias Erbar for a number of fruitful and enlightening discussions.

2. Preliminaries

Notations We write $\mathrm{C}(X)$ and $\mathrm{Lip}(X)$ for the spaces of continuous and Lipschitz functions $f\colon X \to \mathbb{R}$, respectively. We set $\mathrm{Lip}(f) := \sup_{x\neq y} |f(x) - f(y)|/\mathsf{d}(x, y)$ for $f \in \mathrm{Lip}(X)$. The space of *bounded* continuous functions on X is denoted by $\mathrm{C_b}(X)$, and the space of functions in $\mathrm{C}(X)$ with *bounded support* is called $\mathrm{C_{bs}}(X)$, and similarly for $\mathrm{Lip_b}(X)$ and $\mathrm{Lip_{bs}}(X)$.

The Riemannian curvature-dimension condition We say the metric measure space $(X, \mathsf{d}, \mathfrak{m})$ is *infinitesimally Hilbertian* if the Cheeger energy $\mathscr{E}$ is a quadratic form (in other words, if it satisfies the parallelogram identity). Furthermore, we say that $(X, \mathsf{d}, \mathfrak{m})$ satisfies the *Riemannian curvature-dimension condition* $\mathrm{RCD}(k, \infty)$ if it is infinitesimally Hilbertian and satisfies the curvature-dimension condition $\mathrm{CD}(k, \infty)$ according to Definition 1.2. As said, we always assume that $(X, \mathsf{d}, \mathfrak{m})$ is an $\mathrm{RCD}(K, \infty)$ space for some constant $K \in \mathbb{R}$. The value of K does not enter any of our results. Without restriction $k \geq K$ on X. Indeed, one should think of k as being much larger than K everywhere on X.

The $\mathrm{RCD}(K, \infty)$ assumption carries numerous important consequences for $(X, \mathsf{d}, \mathfrak{m})$. Further details on the subsequent results can be found in [3,4,17,26,27].

a. **Volume growth.** For each $z \in X$ there exists a nonnegative constant C such that $\mathfrak{m}[B_r(z)] \leq \mathrm{e}^{Cr^2}$ for every $r > 0$.
b. **Nondegeneracy of entropy.** $\mathrm{Ent}_{\mathfrak{m}}$ is well-defined and does not attain the value $-\infty$ on $\mathscr{P}_2(X)$.
c. **Uniqueness of W_2-geodesics.** For each pair of $\mathfrak{m}$-absolutely continuous measures $\mu_0, \mu_1 \in \mathscr{P}_2(X)$, there exists a unique W_2-geodesic connecting them.
d. **Dirichlet form.** By polarization, $\mathscr{E}$ defines a quasi-regular, strongly local, conservative Dirichlet form, unambiguously denoted by $\mathscr{E}$, on $L^2(X, \mathfrak{m})$ with dense domain $W^{1,2}(X) := \mathrm{Dom}(\mathscr{E})$. The latter is a Hilbert space w.r.t. $\big[\|f\|^2_{L^2(X,\mathfrak{m})} + \mathscr{E}(f)\big]^{1/2}$. The generator of $\mathscr{E}$, i.e. the self-adjoint operator Δ on $L^2(X, \mathfrak{m})$ defined by putting $f \in \mathrm{Dom}(\Delta)$ and $h = \Delta f$ if and only if

$$\mathscr{E}(f, g) = -\int_X h\, g \,\mathrm{d}\mathfrak{m} \quad \text{for every } g \in W^{1,2}(X),$$

is called *Laplacian.*
e. **Heat flow.** The Dirichlet form $\mathscr{E}$ defines the heat semigroup $(\mathsf{P}_t)_{t\geq 0}$ as its gradient flow in $L^2(X, \mathfrak{m})$, or alternatively via spectral calculus as $\mathsf{P}_t = \mathrm{e}^{\Delta t}$, $t \geq 0$. This semigroup is $\mathfrak{m}$-symmetric and extends to a strongly continuous contraction semigroup on $L^r(X, \mathfrak{m})$ for any $r \in [1, \infty)$. It can be chosen to be strong Feller, more precisely, P_t maps $L^\infty(X, \mathfrak{m})$ to $\mathrm{Lip}(X)$ for $t > 0$ with $\mathrm{Lip}(\mathsf{P}_t f) \leq \|f\|_{L^\infty(X,\mathfrak{m})}/\sqrt{t}$ if $K = 0$, while if $K \neq 0$, then

$$\mathrm{Lip}(\mathsf{P}_t f)^2 \leq \frac{K}{\mathrm{e}^{2Kt} - 1} \|f\|^2_{L^\infty(X,\mathfrak{m})} \quad \text{for every } f \in L^\infty(X, \mathfrak{m}). \tag{2.1}$$

The semigroup $(\mathsf{P}_t)_{t\geq 0}$ is in duality with the semigroup $(\mathsf{H}_t)_{t\geq 0}$ defined as the gradient flow of $\mathrm{Ent}_{\mathfrak{m}}$ in $\mathscr{P}_2(X)$ and extended to $\mathscr{P}(X)$ by continuity, i.e.

$$\int_X f \,\mathrm{d}\mathsf{H}_t\mu = \int_X \mathsf{P}_t f \,\mathrm{d}\mu \quad \text{for every } f \in \mathrm{C_b}(X) \text{ and } \mu \in \mathscr{P}(X).$$

In particular, $\mathsf{H}_t(g\, \mathfrak{m}) = (\mathsf{P}_t g)\, \mathfrak{m}$ for every $g \in L^1(X, \mathfrak{m})$.
f. **Uniqueness of EVI curves.** Every curve $(\mu_t)_{t\geq 0}$ in $\mathscr{P}_2(X)$ satisfying the obstructions from Definition 1.3 with arbitrary choice of $k \geq K$ necessarily coincides with the heat flow $(\mathsf{H}_t\mu_0)_{t\geq 0}$ starting at μ_0.
g. **Brownian motion.** For each $\mu \in \mathscr{P}(X)$, there exists a conservative Markov process $(\mathbb{P}, (\mathsf{b}_t)_{t\geq 0})$ on X, or $(\mathbb{P}, \mathsf{b})$ for short, unique in law, with continuous sample paths and transition semigroup given by

$$\mathbb{E}\big[f(\mathsf{b}_{t+s}) \mid \mathsf{b}_s\big] = \mathsf{P}_{t/2}f(\mathsf{b}_s) \quad \text{for every } s,t \in [0,\infty) \text{ and } f \in \mathrm{C_b}(X),$$

and with $(\mathsf{b}_0)_\sharp\mathbb{P} = \mu$. This process is called *the* Brownian motion on X with initial distribution μ. If we want to stress the dependence on the initial distribution, we write $\mathbb{P}_\mu$ instead of $\mathbb{P}$, where we abbreviate $\mathbb{P}_{\delta_x}$ by $\mathbb{P}_x$ for $x \in X$.

h. **Carré du champ.** The set $\mathrm{Lip}(X) \cap L^2(X,\mathfrak{m})$ is a core for $\mathscr{E}$. A quadratic functional $\Gamma\colon W^{1,2}(X) \to L^1(X,\mathfrak{m})$ can be defined by requiring

$$\int_X \Gamma(f)\, g \,\mathrm{d}\mathfrak{m} = \mathscr{E}(f, f\, g) - \frac{1}{2}\mathscr{E}\big(f^2, g\big) \quad \text{for every } g \in \mathrm{Lip_b}(X).$$

Indeed, $\Gamma(f)^{1/2}$ coincides $\mathfrak{m}$-a.e. with the *minimal weak upper gradient* $|\mathrm{D}f|$.

i. **Test functions.** The set

$$\mathrm{TestF}(X) := \big\{f \in \mathrm{Dom}(\Delta) \cap L^\infty(X,\mathfrak{m}) : \Gamma(f) \in L^\infty(X,\mathfrak{m}),\ \Delta f \in W^{1,2}(X)\big\} \tag{2.2}$$

is a core for $\mathscr{E}$ and an algebra w.r.t. pointwise multiplication.

j. **Twice differentiability.** We have $\Gamma(f)^{1/2} \in \mathrm{Dom}(\mathscr{E})$ for all $f \in \mathcal{D}(\Delta)$ and

$$\mathscr{E}\big(\Gamma(f)^{1/2}\big) \leq \|\Delta f\|^2_{L^2(X,\mathfrak{m})} - K\mathscr{E}(f).$$

k. **Sobolev-to-Lipschitz property.** Every $f \in W^{1,2}(X,\mathfrak{m})$ with $|\mathrm{D}f| \in L^\infty(X,\mathfrak{m})$ has a Lipschitz representative $\overline{f}$ with $\mathrm{Lip}(\overline{f}) \leq \||\mathrm{D}f|\|_{L^\infty(X,\mathfrak{m})}$.

Hopf–Lax evolution For later use, we summarize the main properties of the general p-Hopf–Lax (or Hamilton–Jacobi) semigroup $(Q_s)_{s\geq 0}$, $p \in (1,\infty)$. A detailed account on this topic in general metric spaces can be found in [2,3,18].

Fix a Lipschitz function f on X. Its p-Hopf–Lax evolution $(Q_s f)_{s\geq 0}$ is defined by

$$Q_0 f := f \quad \text{and} \quad Q_s f(x) := \inf_{y\in X}\Big\{f(y) + \frac{\mathsf{d}^p(x,y)}{p s^{p-1}}\Big\} \quad \text{for every } s \in (0,\infty) \text{ and } x \in X.$$

The map $s \mapsto Q_s f$ belongs to $\mathrm{Lip}([0,\infty);\mathrm{C}(X))$, where $\mathrm{C}(X)$ is endowed with the usual supremum metric. We also have $Q_s f \in \mathrm{Lip}(X)$ with $\mathrm{Lip}(Q_s f) \leq p\,\mathrm{Lip}(f)$ for all $s \in (0,\infty)$. Denoting by $q \in (1,\infty)$ the dual exponent to p, for every $x \in X$, we have

$$\frac{\mathrm{d}}{\mathrm{d}s}Q_s f(x) + \frac{1}{q}\mathrm{lip}(Q_s f)^q(x) \leq 0$$

for all but at most countably many $s \in (0,\infty)$, and equality holds e.g. if (X,d) is geodesic.

Using the p-Hopf–Lax semigroup gives a nice duality formula for the p-Kantorovich–Wasserstein distance, see [22,33] for details: for all $\mu,\nu \in \mathscr{P}(X)$, one has

$$\frac{1}{p}W_p^p(\mu,\nu) = \sup\Big\{\int_X Q_1 f\,\mathrm{d}\mu - \int_X f\,\mathrm{d}\nu : f \in \mathrm{Lip_b}(X)\Big\}. \tag{2.3}$$

The function $\underline{k}$ and Lipschitz approximation Recall that k is lower semicontinuous and bounded from below by K, and so is $\underline{k}$ by construction. If k is also bounded from above, say by $C \in \mathbb{R}$, then so is $\underline{k}$. By reparameterization of geodesics, we get $\underline{k}(x,y) = \underline{k}(y,x)$ for every $x,y \in X$. Note that k can be reconstructed from $\underline{k}$, since $k(x) = \underline{k}(x,x)$. Lastly, the function $\underline{k}$ defined in (1.2) is the pointwise monotone limit from below of bounded Lipschitz functions $\underline{k}_n$, and so is the function k by considering $\underline{k}_n$ on the diagonal. We intend Lipschitz continuity on $X \times X$ w.r.t. the product metric $\mathsf{d}_{X\times X}$ given by $\mathsf{d}_{X\times X}\big((x,y),(x',y')\big) := \big[\mathsf{d}^2(x,x') + \mathsf{d}^2(y,y')\big]^{1/2}$. The former fact will be used frequently. Following [1], we can, for instance, define $\underline{k}_n \colon X \times X \to \mathbb{R}$ for $n \in \mathbb{N}$ by

$$\underline{k}_n(x,y) := \inf\big\{\min\{\underline{k}(x',y'), n\} + n\,\mathsf{d}_{X\times X}\big((x,y),(x',y')\big) : x',y' \in X\big\}.$$

Lemma 2.1. *The above functions $\underline{k}_n$, $n \in \mathbb{N}$, have the following properties:*

(i) *for every $n \in \mathbb{N}$, the function $\underline{k}_n$ is Lipschitz on $X \times X$ with $\mathrm{Lip}(\underline{k}_n) \leq n$,*
(ii) *for all $x \in X$ and each $n \in \mathbb{N}$, we have $K \leq \underline{k}_n(x) \leq \underline{k}_{n+1}(x) \leq n+1$, and*
(iii) *the sequence $(\underline{k}_n)_{n\in\mathbb{N}}$ converges pointwise from below to $\underline{k}$.*

3. Gradient estimates, Bochner's inequality, and their self-improvements

In this section, we adapt the well-known arguments of [9,10,27] for constant curvature lower bounds to derive the equivalence of the q-Bochner inequality with the q-gradient estimate with exponent $q \in [1,\infty)$. Moreover, we prove that these properties are independent of q.

Up to replacing k by $k_n := \min\{k, n\}$, $n \in \mathbb{N}$, we may assume throughout this chapter that k is bounded. In the general case, each of the subsequent results still holds for k since $\mathrm{BE}_q(k,\infty)$ and $\mathrm{GE}_q(k)$ trivially imply $\mathrm{BE}_q(k_n,\infty)$ and $\mathrm{GE}_q(k_n)$ for every $n \in \mathbb{N}$, respectively, and conversely, if $\mathrm{BE}_q(k_n,\infty)$ and $\mathrm{GE}_q(k_n)$ hold for each $n \in \mathbb{N}$, the monotone convergence theorem implies $\mathrm{BE}_q(k,\infty)$ and $\mathrm{GE}_q(k)$, respectively.

In this chapter, we denote an $\mathscr{E}$-quasi-continuous representative of a given function $h \in W^{1,2}(X)$ by $\widetilde{h}$. The function $\widetilde{h}$ is uniquely determined $\mathscr{E}$-quasi-everywhere.

3.1. Equivalence of Bochner and gradient estimate

First, we review the measure-valued Laplacian $\mathbf{\Delta}$ and the measure-valued Γ_2-operator $\mathbf{\Gamma}_2$ as introduced and analyzed in [16,27], defined by means of

$$\int_X g \,\mathrm{d}\mathbf{\Delta} f = -\int_X \Gamma(g,f)\,\mathrm{d}\mathfrak{m} \quad \text{for every } g \in \mathrm{Lip}_{\mathrm{bs}}(X) \quad \text{and} \tag{3.1}$$

$$\mathbf{\Gamma}_2(f) := \frac{1}{2}\mathbf{\Delta}\Gamma(f) - \Gamma(f,\Delta f)\,\mathfrak{m}$$

for suitable $f \in W^{1,2}(X)$. We write $f \in \mathrm{Dom}(\mathbf{\Delta})$ if the signed measure $\mathbf{\Delta} f$ exists, which is then uniquely determined by (3.1) and does not charge sets of zero capacity. We denote the density of the $\mathfrak{m}$-absolutely continuous part of $\mathbf{\Gamma}_2(f)$ by $\gamma_2(f)$. The singular part of $\mathbf{\Gamma}_2(f)$ w.r.t. $\mathfrak{m}$ is a nonnegative measure. Both $\mathbf{\Delta} f$ and $\mathbf{\Gamma}_2(f)$ are well-defined for $f \in \mathrm{TestF}(X)$. Lastly, a consequence of the generic calculus rules of Γ is the following chain rule for $\mathbf{\Delta}$.

Lemma 3.1. *Fix $f \in \mathrm{Dom}(\mathbf{\Delta}) \cap L^\infty(X,\mathfrak{m})$, an interval $I \subset \mathbb{R}$ with $0 \in I$ containing the image of f, and a function $\Phi \in \mathrm{C}^2(I)$ such that $\Phi(0) = 0$. Then $\Phi(f) \in \mathrm{Dom}(\mathbf{\Delta})$ and*

$$\mathbf{\Delta}\Phi(f) = \Phi'\big(\widetilde{f}\big)\,\mathbf{\Delta} f + \Phi''(f)\,\Gamma(f)\,\mathfrak{m}. \tag{3.2}$$

Proof. Given any $f \in \mathrm{Dom}(\mathbf{\Delta}) \cap L^\infty(X,\mathfrak{m})$, first observe that $\Phi(f) \in W^{1,2}(X)$. Furthermore, by density of Lipschitz functions in $W^{1,2}(X)$ and a simple truncation argument, we find a sequence $(f_n)_{n\in\mathbb{N}}$ of functions in $W^{1,2}(X)\cap\mathrm{Lip}(X)$ which is uniformly bounded in $L^\infty(X,\mathfrak{m})$ such that $f_n\to f$ in $W^{1,2}(X)$. Up to taking a subsequence, we assume without restriction that $(f_n)_{n\in\mathbb{N}}$ converges $\mathscr{E}$-quasi-uniformly to an $\mathscr{E}$-quasi-continuous representative $\widetilde{f}$ of f, cf. [12, Theorem 1.3.3]. Since $\mathbf{\Delta} f$ has finite total variation on any bounded subset of X, the dominated convergence theorem yields, for every $g\in\mathrm{Lip}_{\mathrm{bs}}(X)$ and up to possibly taking further subsequences,

$$\begin{aligned}
&\int_X g\,\Phi'\big(\widetilde{f}\big)\,\mathrm{d}\mathbf{\Delta} f+\int_X g\,\Phi''(f)\,\Gamma(f)\,\mathrm{d}\mathfrak{m}\\
&\qquad=\lim_{n\to\infty}\int_X g\,\Phi'(f_n)\,\mathrm{d}\mathbf{\Delta} f+\int_X g\,\Phi''(f)\,\Gamma(f)\,\mathrm{d}\mathfrak{m}\\
&\qquad=-\lim_{n\to\infty}\int_X \Gamma\big(g\,\Phi'(f_n),f\big)\,\mathrm{d}\mathfrak{m}+\int_X g\,\Phi''(f)\,\Gamma(f)\,\mathrm{d}\mathfrak{m}\\
&\qquad=-\lim_{n\to\infty}\Big(\int_X \Phi'(f_n)\,\Gamma(g,f)\,\mathrm{d}\mathfrak{m}+\int_X g\,\Phi''(f_n)\,\Gamma(f_n,f)\,\mathrm{d}\mathfrak{m}\Big)+\int_X g\,\Phi''(f)\,\Gamma(f)\,\mathrm{d}\mathfrak{m}\\
&\qquad=-\int_X \Phi'(f)\,\Gamma(g,f)\,\mathrm{d}\mathfrak{m}=-\int_X \Gamma\big(g,\Phi(f)\big)\,\mathrm{d}\mathfrak{m}.
\end{aligned}$$

The claim follows from definition (3.1) of $\mathbf{\Delta}$. □

Once $\mathrm{BE}_2(k,\infty)$ holds, as in the proof of [27, Lemma 3.2] we get

$$\begin{aligned}
&\mathscr{E}\big(\Gamma(f)\big)\le-\int_X 2k\,\Gamma(f)^2+\Gamma(f)\,\Gamma(f,\Delta f)\,\mathrm{d}\mathfrak{m}\quad\text{and}\\
&k\,\Gamma(f)\,\mathfrak{m}\le\frac{1}{2}\mathbf{\Delta}\Gamma(f)-\Gamma(f,\Delta f)\,\mathfrak{m}
\end{aligned}$$

for every $f\in\mathrm{TestF}(X)$. Taking these estimates into account, one can argue exactly as in [9, Proposition 1], which has also been leaned on [27, Theorem 3.4], to obtain that, for every $f\in\mathrm{TestF}(X)$,

$$\Gamma(\Gamma(f))\le 4\big(\gamma_2(f)-k\,\Gamma(f)\big)\,\Gamma(f)\quad\mathfrak{m}\text{-a.e.}\tag{3.3}$$

Using this, we deduce the whole range of q-Bochner inequalities from $\mathrm{BE}_2(k,\infty)$.

Proposition 3.2. *The condition* $\mathrm{BE}_2(k,\infty)$ *implies* $\mathrm{BE}_q(k,\infty)$ *for every* $q\in[1,\infty)$.

Proof. Fix $f\in\mathrm{TestF}(X)$ and a nonnegative $\phi\in\mathrm{Dom}(\Delta)\cap L^\infty(X,\mathfrak{m})$ with $\Delta\phi\in L^\infty(X,\mathfrak{m})$. Given $\varepsilon>0$, consider the smooth function $\Phi_\varepsilon(r):=(r+\varepsilon)^{q/2}-\varepsilon^{q/2}$ defined for $r\ge0$. Since $2-q\le1$, we obtain the $\mathfrak{m}$-a.e. inequalities

$$-\Gamma(\Gamma(f))\,\Phi_\varepsilon''(\Gamma(f))\le\frac{q}{4}\Gamma(\Gamma(f))\,\big(\Gamma(f)+\varepsilon\big)^{q/2-2}\le2\big(\gamma_2(f)-k\,\Gamma(f)\big)\,\Phi_\varepsilon'(\Gamma(f))$$

by means of (3.3). Multiplying this by ϕ and integrating, one gets

$$-\int_X \Gamma(\Gamma(f))\,\Phi_\varepsilon''(\Gamma(f))\,\phi\,\mathrm{d}\mathfrak{m}$$

$$\leq 2\int_X \Phi'_\varepsilon\big(\widetilde{\Gamma(f)}\big)\,\widetilde{\phi}\,\mathrm{d}\boldsymbol{\Gamma}_2(f) - 2\int_X k\,\Gamma(f)\,\Phi'_\varepsilon(\Gamma(f))\,\phi\,\mathrm{d}\mathfrak{m}$$

$$= \int_X \Phi'_\varepsilon\big(\widetilde{\Gamma(f)}\big)\,\widetilde{\phi}\,\mathrm{d}\boldsymbol{\Delta}\Gamma(f) - 2\int_X \Phi'_\varepsilon(\Gamma(f))\,\big(\Gamma(f,\Delta f) + k\,\Gamma(f)\big)\,\phi\,\mathrm{d}\mathfrak{m}.$$

Invoking Lemma 3.1, this amounts to

$$2\int_X \Phi'_\varepsilon(\Gamma(f))\,\big(\Gamma(f,\Delta f) + k\,\Gamma(f)\big)\,\phi\,\mathrm{d}\mathfrak{m} \leq \int_X \widetilde{\phi}\,\mathrm{d}\boldsymbol{\Delta}\Phi_\varepsilon(\Gamma(f)) = \int_X \Phi_\varepsilon(\Gamma(f))\,\Delta\phi\,\mathrm{d}\mathfrak{m}.$$

By the dominated convergence theorem, letting $\varepsilon \downarrow 0$ in the preceding inequality gives the $\mathrm{BE}_q(k,\infty)$ inequality for $f \in \mathrm{TestF}(X)$.

To extend this to general $f \in \mathrm{Dom}(\Delta)$ with $\Delta f \in W^{1,2}(X)$ and $\Gamma(f) \in L^\infty(X,\mathfrak{m})$, we approximate it in $W^{1,2}(X)$ by means of its heat flow regularizations $\mathsf{P}_t f \in \mathrm{TestF}(X)$ as $t \downarrow 0$. Since $\Gamma(\mathsf{P}_t f) \to \Gamma(f)$ and $\Gamma(\mathsf{P}_t f, \Delta\mathsf{P}_t f) \to \Gamma(f,\Delta f)$ in $L^1(X,\mathfrak{m})$ as $t \downarrow 0$, $\Gamma(\mathsf{P}_t f)$ is uniformly bounded in $L^\infty(X,\mathfrak{m})$ for small enough t, and $\Gamma(\Delta\mathsf{P}_t f)^{1/2}$ is uniformly bounded in $L^2(X,\mathfrak{m})$ for small enough t, we easily get

$$\lim_{t\downarrow 0}\Gamma(\mathsf{P}_t f)^{q/2} = \Gamma(f)^{q/2} \quad\text{and}\quad \lim_{t\downarrow 0}\Gamma(\mathsf{P}_t f)^{q/2-1}\,\Gamma(\mathsf{P}_t f,\Delta\mathsf{P}_t f) = \Gamma(f)^{q/2-1}\,\Gamma(f,\Delta f)$$

in $L^1(X,\mathfrak{m})$. This yields the claim. □

By the Feynman–Kac representation (1.1) of P_t^{qk} and Jensen's inequality, the following hierarchy between gradient estimates is immediate. This and the above self-improvement property of $\mathrm{BE}_2(k,\infty)$ will be used in the proof of Theorem 3.4 below.

Lemma 3.3. *If $\mathrm{GE}_q(k)$ holds for some $q \in [1,\infty)$, then $\mathrm{GE}_{q'}(k)$ is satisfied for all $q' \in [q,\infty)$.*

Theorem 3.4. *For every $q \in [1,\infty)$, the properties $\mathrm{BE}_q(k,\infty)$ and $\mathrm{GE}_q(k)$ are equivalent to each other.*

Proof. By density of $\mathrm{TestF}(X)$ in $W^{1,2}(X)$ and an argument as in the proof of Proposition 3.2, the function f under consideration may be assumed to belong to $\mathrm{TestF}(X)$.

Suppose that $\mathrm{BE}_q(k,\infty)$ is satisfied. Fix any $t > 0$, f as above and a nonnegative $\phi \in \mathrm{Dom}(\Delta)\cap L^\infty(X,\mathfrak{m})$ with $\Delta\phi \in L^\infty(X,\mathfrak{m})$. Given any $\varepsilon > 0$, consider the function Φ_ε as defined in the proof of Proposition 3.2 above. Define $F_\varepsilon\colon [0,t] \to \mathbb{R}$ by

$$F_\varepsilon(s) := \int_X \mathsf{P}_s^{qk}\big(\Phi_\varepsilon\big(\Gamma(\mathsf{P}_{t-s}f)\big)\big)\,\phi\,\mathrm{d}\mathfrak{m} = \int_X \Phi_\varepsilon\big(\Gamma(\mathsf{P}_{t-s}f)\big)\,\mathsf{P}_s^{qk}\phi\,\mathrm{d}\mathfrak{m}.$$

This function belongs to $\mathrm{C}^1([0,t])$ since the functions $s \mapsto \mathsf{P}_s^{qk}\phi$ and $s \mapsto \Phi_\varepsilon\big(\Gamma(\mathsf{P}_{t-s}f)\big)$ as well as their derivatives in $L^2(X,\mathfrak{m})$ are bounded on $[0,t]$, see also [5, Lemma 2.1] for a similar argument. Thus

$$\begin{aligned}\liminf_{\varepsilon\downarrow 0} F'_\varepsilon(s) \geq \liminf_{\varepsilon\downarrow 0}\int_X \Phi_\varepsilon\big(\Gamma(\mathsf{P}_{t-s}f)\big)\,(\Delta - qk)\mathsf{P}_s^{qk}\phi\,\mathrm{d}\mathfrak{m}\\ - 2\limsup_{\varepsilon\downarrow 0}\int_X \Phi'_\varepsilon\big(\Gamma(\mathsf{P}_{t-s}f)\big)\,\Gamma(\mathsf{P}_{t-s}f,\Delta\mathsf{P}_{t-s}f)\,\mathsf{P}_s^{qk}\phi\,\mathrm{d}\mathfrak{m},\end{aligned}$$

which is nonnegative by $\mathrm{BE}_q(k,\infty)$. Fatou's lemma gives

$$F_0(t) - F_0(0) = \liminf_{\varepsilon\downarrow 0} \big(F_\varepsilon(t) - F_\varepsilon(0)\big) \geq \int_0^t \liminf_{\varepsilon\downarrow 0} F'_\varepsilon(s)\,\mathrm{d}s \geq 0,$$

which establishes $\mathrm{GE}_q(k)$ for $f \in \mathrm{TestF}(X)$ by the arbitrariness of ϕ.

Conversely, assume $\mathrm{GE}_q(k)$ for $q \in [2,\infty)$. As $\Phi_0 \in \mathrm{C}^1([0,\infty))$ for such q, we deduce $F'_0(0) \geq 0$, which is a reformulation of the $\mathrm{BE}_q(k,\infty)$ inequality with $\mathsf{P}_t f$ in place of f. Letting $t \downarrow 0$ gives the desired conclusion. If, on the other hand, we have $q \in [1,2)$, we cannot rely on the above regularity of Φ_0. However, Lemma 3.3 ensures $\mathrm{GE}_2(k)$, which implies $\mathrm{BE}_2(k,\infty)$ by the previous discussion. Therefore, $\mathrm{BE}_q(k,\infty)$ holds by Proposition 3.2. □

3.2. Independence of the q-Bochner inequality of q

In this section, we prove the independence of the q-Bochner inequality of $q \in [1,\infty)$. See Theorem 3.6 below for the precise statement.

We start with the following result. For its proof, we adapt the arguments of [19]. A crucial point in this argument is that our a priori RCD assumption guarantees $\Gamma(f)^{q/2} \in \mathrm{Dom}(\mathbf{\Delta})$ for all $f \in \mathrm{TestF}(X)$ and *every* $q \in [1,\infty)$, and that $\mathrm{TestF}(X)$ is dense in $W^{1,2}(X)$.

Proposition 3.5. *The condition* $\mathrm{BE}_q(k,\infty)$ *implies* $\mathrm{BE}_2(k,\infty)$ *for every* $q \in (2,\infty)$.

Proof. As discussed above, it suffices to show the claimed implication starting from $\mathrm{GE}_q(k)$ with $q \in (2,\infty)$.

Arguing exactly as in the constant situation in [19, Lemma 3.2] (see also [27, Theorem 3.4]), one can show that for every $r \in (2,\infty)$, $\mathrm{BE}_r(k,\infty)$ holds if and only if the inequalities

$$\frac{1}{2}\Gamma(f)\,\delta(\Gamma(f)) + \frac{r-2}{4}\Gamma(\Gamma(f)) \geq \Gamma(f)\,\Gamma(f,\Delta f) + k\,\Gamma(f)^2 \quad \mathfrak{m}\text{-a.e.} \quad \text{and} \quad \widetilde{\Gamma(f)}\,\mathbf{\Delta}_\perp\Gamma(f) \geq 0 \tag{3.4}$$

are valid for every $f \in \mathrm{TestF}(X)$. Here, $\delta(\Gamma(f))$ is the density of the $\mathfrak{m}$-absolutely continuous part of $\mathbf{\Delta}\Gamma(f)$ w.r.t. $\mathfrak{m}$, $\mathbf{\Delta}_\perp\Gamma(f)$ stands for the corresponding $\mathfrak{m}$-singular part. In particular, note that $\mathrm{GE}_q(k)$ already yields $\widetilde{\Gamma(f)}\,\mathbf{\Delta}_\perp\Gamma(f) \geq 0$ by (3.4) which is independent of q.

The crucial point is to show that

$$\frac{1}{2}\Gamma(f)\,\delta(\Gamma(f)) + \varepsilon\Gamma(\Gamma(f)) \geq \Gamma(f)\,\Gamma(f,\Delta f) + k\,\Gamma(f)^2 \quad \mathfrak{m}\text{-a.e.} \tag{3.5}$$

for every $\varepsilon > 0$. Given the observation (3.4), this will imply $\mathrm{BE}_{2+4\varepsilon}(k,\infty)$ for each $\varepsilon > 0$, and eventually letting $\varepsilon \downarrow 0$ and applying the monotone convergence theorem, we get the claimed $\mathrm{BE}_2(k,\infty)$ condition.

Given $\mathrm{BE}_{q'}(k,\infty)$ for arbitrary $q' \geq q$, it is straightforward to follow the proof of [19, Theorem 3.6], which relies on generic calculus rules for $\mathbf{\Gamma}_2$ and closely follows the strategy presented in [27], to prove (3.5) with ε replaced by $q' - \frac{1}{4(q'+1)}$. Now, according to [19, Lemma 3.3], given any $\varepsilon > 0$ there exist $n \in \mathbb{N}$ and $q' \geq q$ so that $P^n(q') = \varepsilon$, where $P(r) := r - \frac{1}{4(r+1)}$ and P^n is the n-fold composition of P. Since $\mathrm{BE}_q(k,\infty)$ yields $\mathrm{BE}_{q'}(k,\infty)$, iterating the foregoing reasoning allows us to finally reach the inequality (3.5). □

Combining Proposition 3.2 and Lemma 3.3 with Proposition 3.5 yields the following.

Theorem 3.6. *If the q-Bakry–Émery estimate* $\mathrm{BE}_q(k,\infty)$ *holds for some* $q \in [1,\infty)$*, then it holds for every* $q \in [1,\infty)$.

It is also possible to obtain an equivalent characterization of $\mathrm{BE}_2(k,\infty)$ in terms of a lower bound on the measure-valued *Ricci tensor*

$$\mathbf{Ric}(\nabla f, \nabla f) := \mathbf{\Gamma}_2(f) - \big|\mathrm{Hess}\, f\big|^2_{\mathsf{HS}}\, \mathfrak{m} \quad \text{for every } f \in \mathrm{TestF}(X)$$

introduced in [17]. As for the measure-valued Laplacian $\mathbf{\Delta}$, we denote by $\mathrm{ric}(\nabla f, \nabla f)$ the density of the $\mathfrak{m}$-absolutely continuous part and by $\mathbf{Ric}_\perp(\nabla f, \nabla f)$ the $\mathfrak{m}$-singular part of $\mathbf{Ric}(\nabla f, \nabla f)$, respectively.

Corollary 3.7. *The* RCD *space* $(X, \mathsf{d}, \mathfrak{m})$ *satisfies* $\mathrm{BE}_2(k, \infty)$ *if and only if for every* $f \in \mathrm{TestF}(X)$*, we have*

$$\mathrm{ric}(\nabla f, \nabla f) \geq k\, \Gamma(f) \quad \mathfrak{m}\text{-}a.e. \quad and \quad \mathbf{Ric}_\perp(\nabla f, \nabla f) \geq 0.$$

3.3. Localization of Bochner's inequality

To study a suitable local-to-global behavior of the q-Bochner inequality, we present a reformulation of it where we enlarge the class of functions ϕ. Recall that our standing assumption $\mathrm{RCD}(K, \infty)$ implies $\Gamma(f)^{q/2} \in W^{1,2}(X)$ for every $f \in \mathrm{TestF}(X)$ and $q \in [1, \infty)$.

Lemma 3.8. *Given* $q \in [1, \infty)$*, the* $\mathrm{BE}_q(k, \infty)$ *property holds if and only if for all* $f \in \mathrm{TestF}(X)$ *and all nonnegative* $\phi \in W^{1,2}(X) \cap L^\infty(X, \mathfrak{m})$*,*

$$-\int_X \Big(\frac{1}{q}\Gamma\big(\Gamma(f)^{q/2}, \phi\big) + \Gamma(f)^{q/2-1}\, \Gamma(f, \Delta f)\, \phi\Big)\, \mathrm{d}\mathfrak{m} \geq \int_X k\, \Gamma(f)^{q/2}\, \phi\, \mathrm{d}\mathfrak{m}. \tag{3.6}$$

Proof. Obtaining $\mathrm{BE}_q(k, \infty)$ from (3.6) through integration by parts and the density of $\mathrm{TestF}(X)$ in $W^{1,2}(X)$ is easy, thus we focus on the converse. Trivially, the inequality (3.6) holds for all $\phi \in \mathrm{Dom}(\Delta) \cap L^\infty(X, \mathfrak{m})$ with $\Delta\phi \in L^\infty(X, \mathfrak{m})$. Recall now, e.g. from [17,27], that any function $\phi \in W^{1,2}(X) \cap L^\infty(X, \mathfrak{m})$ can be approximated in $W^{1,2}(X)$ by means of a mollified heat flow

$$\mathfrak{P}_\varepsilon \phi := \int_0^\infty \eta(s)\, \mathsf{P}_{\varepsilon s}\phi\, \mathrm{d}s, \quad \text{where} \quad \eta \in \mathrm{C}_\mathrm{c}^\infty((0, \infty); [0, \infty)) \quad \text{with} \quad \int_0^\infty \eta(s)\, \mathrm{d}s = 1,$$

as $\varepsilon \downarrow 0$. Since $\mathfrak{P}_\varepsilon\phi \in \mathrm{Dom}(\Delta) \cap L^\infty(X, \mathfrak{m})$ and $\Delta\mathfrak{P}_\varepsilon\phi = -\int_0^\infty \eta'(s)\, \mathsf{P}_{\varepsilon s}\phi\, \mathrm{d}s/\varepsilon \in L^\infty(X, \mathfrak{m})$ for every $\varepsilon > 0$, this allows us to extend the class of admissible ϕ. □

Definition 3.9. We say that the *local q-Bakry–Émery condition* with variable curvature bound k, briefly $\mathrm{BE}_{q,\mathrm{loc}}(k, \infty)$, with $q \in [1, \infty)$ holds if for every $z \in X$ there exists $\delta > 0$ such that

$$-\int_X \Big(\frac{1}{q}\Gamma\big(\Gamma(f)^{q/2}, \phi\big) + \Gamma(f)^{q/2-1}\, \Gamma(f, \Delta f)\, \phi\Big)\, \mathrm{d}\mathfrak{m} \geq \int_X k\, \Gamma(f)^{q/2}\, \phi\, \mathrm{d}\mathfrak{m}$$

for all $f \in \mathrm{TestF}(X)$ and every nonnegative $\phi \in W^{1,2}(X) \cap L^\infty(X, \mathfrak{m})$ with $\mathrm{supp}\, \phi \subset B_\delta(z)$.

It is elementary to pass from the global $\mathrm{BE}_q(k, \infty)$ condition to $\mathrm{BE}_{q,\mathrm{loc}}(k, \infty)$. The converse is more involved. The proof of the following result is similar to the one of [6, Theorem 6.12], but uses a more elementary partition of unity and does not require local compactness or upper dimension bounds of the base space.

Theorem 3.10. *For* $q \in [1, \infty)$*, the property* $\mathrm{BE}_{q,\mathrm{loc}}(k, \infty)$ *implies the* $\mathrm{BE}_q(k, \infty)$ *condition.*

Proof. Let $\{z_i : i \in \mathbb{N}\}$ be a countable dense subset of X and consider the collection of metric balls $B_{\delta_i}(z_i)$ with $\delta_i > 0$ chosen in such a way that the local q-Bakry–Émery inequality is satisfied around z_i. For $i \in \mathbb{N}$, define functions on X by

$$\eta_i^0 := \frac{2}{\delta_i}\mathsf{d}\big(\cdot, X \setminus B_{\delta_i}(z_i)\big), \quad \eta_i^* := \min\Big\{\sum_{j=1}^{i}\eta_j^0, 1\Big\} \quad \text{and} \quad \eta_i := \eta_i^* - \eta_{i-1}^*.$$

Then $\eta_i \in \mathrm{Lip}_{\mathrm{b}}(X)$ with support in $B_{\delta_i}(z_i)$ and $\sum_{i=1}^{\infty}\eta_i = 1$ on X. Thus, for arbitrary nonnegative $\phi \in W^{1,2}(X) \cap L^\infty(X, \mathfrak{m})$, the assumption $\mathrm{BE}_{q,\mathrm{loc}}(k,\infty)$ allows us to deduce

$$\begin{aligned}
&-\int_X \Big(\frac{1}{q}\Gamma\big(\Gamma(f)^{q/2}, \phi\big) + \Gamma(f)^{q/2-1}\,\Gamma(f, \Delta f)\,\phi\Big)\,\mathrm{d}\mathfrak{m} \\
&\qquad = -\sum_{i=1}^{\infty}\int_X \Big(\frac{1}{q}\Gamma\big(\Gamma(f)^{q/2}, \phi\,\eta_i\big) + \Gamma(f)^{q/2-1}\,\Gamma(f, \Delta f)\,\phi\,\eta_i\Big)\,\mathrm{d}\mathfrak{m} \\
&\qquad \geq \sum_{i=1}^{\infty}\int_X k\,\Gamma(f)^{q/2}\,\phi\,\eta_i\,\mathrm{d}\mathfrak{m} = \int_X k\,\Gamma(f)^{q/2}\,\phi\,\mathrm{d}\mathfrak{m}.
\end{aligned}$$

We conclude the assertion using Lemma 3.8 above. □

4. From 2-gradient estimates to CD and differential 2-transport estimates

Our goal now is to derive the evolution variational inequality $\mathrm{EVI}(k)$ with variable curvature bound k from the 2-gradient estimate $\mathrm{GE}_2(k)$. In [31] there is a first part of the proof for this implication. With some extra arguments, we complete it.

The key point is a localization argument. Indeed, it suffices to prove the $\mathrm{EVI}(k)$ "locally", that is, for measures in a given small neighborhood. The heat flow will neither stay within this neighborhood nor in any other bounded region. We thus modify it by truncating its tails. Due to the Gaussian behavior of the heat flow, the difference is of arbitrary polynomial order for small times. This will imply the $\mathrm{CD}(k,\infty)$ inequality locally. However, the latter is already known to give the $\mathrm{CD}(k,\infty)$ inequality globally, and this in turn yields the global version of the $\mathrm{EVI}(k)$.

4.1. Tail estimates for the heat flow

Given any ball $B_\delta(z) \subset X$ with $\delta > 0$ and $z \in X$, and $\rho \in \mathscr{P}(X)$, we put

$$\mathsf{H}_t^*\rho := \mathbb{1}_{B_{2\delta}(z)}\,\mathsf{H}_t\rho + \mathsf{H}_t\rho[X \setminus B_{2\delta}(z)]\,\delta_z.$$

Lemma 4.1. *Assume that $\rho \in \mathscr{P}(X)$ is $\mathfrak{m}$-absolutely continuous with density $f \in L^2(X, \mathfrak{m})$ and $\operatorname{supp}\rho \subset B_\delta(z)$. Then for every $a > 0$ there exists $t_* > 0$ such that for all $t \in [0, t_*]$ and all bounded Borel functions ϕ, we have*

$$W_2^2(\mathsf{H}_t^*\rho, \mathsf{H}_t\rho) \leq t^a \quad \textit{and} \quad \Big|\int_X \phi\,\mathrm{d}\mathsf{H}_t^*\rho - \int_X \phi\,\mathrm{d}\mathsf{H}_t\rho\Big| \leq t^a\,\sup|\phi|(X).$$

Proof. To see the first assertion for $t > 0$, the case $t = 0$ being trivial, observe that

$$W_2^2(\mathsf{H}_t^*\rho, \mathsf{H}_t\rho) \leq \int_{X\setminus B_{2\delta}(z)} \mathsf{d}^2(z,x)\,\mathrm{d}\mathsf{H}_t\rho(x)$$

$$\leq \sum_{n=3}^{\infty} (n\delta)^2 \int\limits_{B_{n\delta}(z)\setminus B_{(n-1)\delta}(z)} \mathsf{P}_t f \,\mathrm{d}\mathfrak{m}$$

$$\leq \|f\|_{L^2(X,\mathfrak{m})} \sum_{n=3}^{\infty} (n\delta)^2 \Big(\mathfrak{m}\big[B_{n\delta}(z) \setminus B_{(n-1)\delta}(z)\big]\Big)^{1/2} \mathrm{e}^{-(n-2)^2\delta^2/4t}$$

where the last inequality comes from the integrated Gaussian heat kernel estimate of [28, Theorem 1.8] applied with $K = k = 1$, $\gamma = \lambda = 0$, $A = B_{n\delta}(z) \setminus B_{(n-1)\delta}(z)$ and $B = B_\delta(z)$ and replacing $\mathbb{1}_B$ and $\mathfrak{m}[B]^{1/2}$ by f and $\|f\|_{L^2(X,\mathfrak{m})}$ therein, respectively. Therefore, by the volume growth property in $\mathrm{RCD}(K,\infty)$ spaces and finally assuming that t is small enough, we obtain

$$W_2^2(\mathsf{H}_t^*\rho, \mathsf{H}_t\rho) \leq \|f\|_{L^2(X,\mathfrak{m})} \Big(\sum_{n=3}^{\infty} \mathfrak{m}\big[B_{n\delta}(z) \setminus B_{(n-1)\delta}(z)\big]\, \mathrm{e}^{-n^2\delta^2/72t}\Big)^{1/2} \mathrm{e}^{-\delta^2/8t}$$

$$\leq \|f\|_{L^2(X,\mathfrak{m})} \Big(\int\limits_X \mathrm{e}^{-\,\mathsf{d}^2(z,x)/72t}\,\mathrm{d}\mathfrak{m}(x)\Big)^{1/2} \mathrm{e}^{-\delta^2/8t} \leq t^a.$$

The second assertion follows from the first one, since

$$\Big|\int\limits_X \phi\,\mathrm{d}\mathsf{H}_t^*\rho - \int\limits_X \phi\,\mathrm{d}\mathsf{H}_t\rho\Big| \leq \sup|\phi|(X)\,\mathsf{H}_t\rho[X \setminus B_{2\delta}(z)] \leq \frac{\sup|\phi|(X)}{\delta^2}\,W_2^2(\mathsf{H}_t^*\rho, \mathsf{H}_t\rho). \qquad \square$$

In Chapter 5, we need the following result, which is a consequence of Lemma 4.1.

Lemma 4.2. *For each $z \in X$, $\delta > 0$ and $a > 0$ there exists $t_* > 0$ such that*

$$\mathbb{P}_x\big[\mathsf{b}_t^x \notin B_{3\delta}(z)\big] \leq t^a \quad \textit{for every } x \in B_\delta(z) \textit{ and } t \in [0, t_*],$$

where $(\mathbb{P}_x, \mathsf{b}^x)$ denotes Brownian motion on X starting in x.

Proof. Let ρ be the uniform distribution of $B_{\delta/2}(z)$. Choose a pair $(\mathbb{P}, \mathsf{b}^x)$ and $(\mathbb{P}, \mathsf{b})$ of coupled Brownian motions with initial distributions δ_x and ρ, respectively, such that $\mathsf{d}(\mathsf{b}_t^x, \mathsf{b}_t) \leq \mathrm{e}^{-Kt}\,\mathsf{d}(x, \mathsf{b}_0)$ $\mathbb{P}$-a.s. for every $t \geq 0$, see [31, Theorem 2.9] for the construction. Thus in particular, $\mathbb{P}$-a.s. we have

$$\mathsf{d}(\mathsf{b}_t^x, \mathsf{b}_t) \leq \delta$$

for every $t \in [0, t_*']$ and a suitable $t_*' > 0$. According to the previous Lemma 4.1,

$$\mathbb{P}\big[\mathsf{b}_t \notin B_{2\delta}(z)\big] \leq t^a$$

for all $t \in [0, t_*]$ and some $t_* > 0$ depending only on $\mathfrak{m}[B_{\delta/2}(z)]$ and a. Combining both estimates yields that

$$\mathbb{P}\big[\mathsf{b}_t^x \notin B_{3\delta}(z)\big] \leq \mathbb{P}\big[\mathsf{b}_t \notin B_{2\delta}(z)\big] \leq t^a.$$

uniformly in $x \in B_\delta(z)$ for small enough times. $\square$

4.2. From 2-gradient estimates to CD

In this section, we assume that k is Lipschitz and bounded. The general case follows using the approximation scheme via the sequence $(k_n)_{n\in\mathbb{N}}$ with $k_n(x) := \underline{k}_n(x,x)$ for $x \in X$ derived from Lemma 2.1. Indeed, $\mathrm{GE}_2(k)$ trivially implies $\mathrm{GE}_2(k_n)$ for every $n \in \mathbb{N}$, which will imply both $\mathrm{CD}(k_n,\infty)$ and $\mathrm{EVI}(k_n)$. Since W_2-geodesics between $\mathfrak{m}$-absolutely continuous measures and $\mathrm{EVI}(k)$-curves are unique, we may then pass to the limit $n \to \infty$ by monotone convergence.

We present a modification of [31, Lemma 3.5] which is proved in exactly the same way as the previous version subject to the choice of parameterization from [5, Theorem 4.16] involving the additional parameter κ. Throughout this section, we denote by $(Q_s)_{s\geq 0}$ the 2-Hopf–Lax semigroup.

Lemma 4.3. *Assume the 2-gradient estimate* $\mathrm{GE}_2(k)$ *with variable curvature bound k, and let $\kappa \in \mathbb{R}$ be an arbitrary constant. Let $(\rho_s)_{s\in[0,1]}$ with $\rho_s = f_s\,\mathfrak{m}$ be a regular curve in the sense of [5, Definition 4.10], and for $t > 0$, define $\vartheta_{\kappa,t}(s) := \frac{e^{\kappa st}-1}{e^{\kappa t}-1}$ if $\kappa \neq 0$ and $\vartheta_{0,t}(s) := s$ as well as $\mathrm{R}_\kappa(t) := \frac{\kappa t}{e^{\kappa t}-1}$ if $\kappa \neq 0$ and* $\mathrm{R}_0(t) := 1$. *Then*

$$\int_X Q_1\phi\,\mathrm{d}\mathsf{H}_t\rho_1 - \int_X \phi\,\mathrm{d}\rho_0 - \frac{1}{2}\mathrm{R}_\kappa^2(t)\int_0^1 \left|\dot\rho_{\vartheta_{\kappa,t}(s)}\right|^2\mathrm{d}s + t\left(\mathrm{Ent}_\mathfrak{m}(\mathsf{H}_t\rho_1) - \mathrm{Ent}_\mathfrak{m}(\rho_0)\right)$$
$$\leq -\int_0^1\int_0^{st}\int_X \mathsf{P}_r\Big((k-\kappa)\,\mathsf{P}_{st-r}^{2(k-\kappa)}\Gamma(Q_s\phi)\Big)\,\mathrm{d}\rho_{\vartheta_{\kappa,t}(s)}\,\mathrm{d}r\,\mathrm{d}s$$

is satisfied for every $\phi \in \mathrm{Lip}_{\mathrm{bs}}(X)$ and all $t > 0$. The term $\left|\dot\rho_{\vartheta_t(s)}\right|$ has to be understood as the metric speed of the original curve $(\rho_s)_{s\in[0,1]}$ evaluated at $\vartheta_t(s)$.

The same estimate is satisfied for every W_2-geodesic $(\rho_s)_{s\in[0,1]}$ with $\mathfrak{m}$-absolutely continuous measures, in which case $\int_0^1 |\dot\rho_{\vartheta_{\kappa,t}(s)}|^2\,\mathrm{d}s = W_2^2(\rho_0,\rho_1)$, independently of κ and t.

Lemma 4.4. *Assume the 2-gradient estimate* $\mathrm{GE}_2(k)$ *with variable curvature bound k. Suppose that $k \geq K_z$ in $B_{2\delta}(z)$ for some $z \in X$, $K_z \in \mathbb{R}$ and $\delta > 0$. Then for all $\rho_0, \rho_1 \in \mathscr{P}_2(X) \cap \mathrm{Dom}(\mathrm{Ent}_\mathfrak{m})$ with support in $B_\delta(z)$ and bounded densities w.r.t. $\mathfrak{m}$, we have*

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0}\frac{1}{2}W_2^2(\mathsf{H}_t\rho_1,\rho_0) + \frac{K_z}{2}W_2^2(\rho_0,\rho_1) \leq \mathrm{Ent}_\mathfrak{m}(\rho_0) - \mathrm{Ent}_\mathfrak{m}(\rho_1).$$

Proof. The proof follows the reasoning for [31, Lemma 3.6] and [5, Theorem 4.16], but with a subtle modification. Fix $t > 0$. While the curve $(\mathsf{H}_{ts}\rho_{\vartheta_t(s)})_{s\in[0,1]}$ connects ρ_0 and $\mathsf{H}_t\rho_1$, the potentials $Q_s\phi_t$, $s \in [0,1]$, are Hopf–Lax interpolations of optimal Kantorovich potentials for the transport from ρ_0 to $\mathsf{H}_t^*\rho_1$. Thus, we have to match these two different situations and then use the nice behavior of the remainder terms.

We know by [3, Proposition 3.9] that for any W_2-optimal coupling $\pi_t \in \mathscr{P}(X\times X)$ of ρ_0 and $\mathsf{H}_t^*\rho_1$, and any Kantorovich potential φ_t relative to π_t, we have $|\mathrm{D}\varphi_t| \leq \mathsf{d}(x,y) \leq 4\delta$ for π_t-a.e. $(x,y) \in X\times X$. Taking (2.3) and the bounded support of ρ_0 into account,

$$\frac{1}{2}W_2^2(\mathsf{H}_t^*\rho_1,\rho_0) = \sup\Big\{\int_X Q_1 f\,\mathrm{d}\mathsf{H}_t^*\rho_1 - \int_X f\,\mathrm{d}\rho_0 : f \in \mathrm{Lip}_{\mathrm{bs}}(X),\ \mathrm{Lip}(f) \leq 4\delta\Big\}.$$

The latter supremum is attained, see [3, Proposition 2.12], at some $\phi_t \in \mathrm{Lip}_{\mathrm{bs}}(X)$. Possibly adding constants and invoking a cutoff argument, we may assume that $|\phi_t| \leq C$ everywhere on X for some $C > 0$ independent of t. Thus, $|Q_s\phi_t|$ is bounded on X and $\mathrm{Lip}(Q_s\phi_t) \leq 8\delta$, uniformly in $s \in [0,1]$.

Let $(\rho_s)_{s\in[0,1]}$ be the W_2-geodesic joining ρ_0 and ρ_1. Note that the measures $\rho_s = f_s\,\mathfrak{m}$, $s\in[0,1]$, are supported in $B_{2\delta}(z)$. The $\mathrm{CD}(K,\infty)$ condition furthermore ensures that the f_s are bounded uniformly in s, cf. [25, Theorem 1.3]. Applying Lemma 4.3 with $\kappa := K_z$ we get

$$\begin{aligned}
&\frac{1}{2t}\Big(W_2^2(\mathsf{H}_t\rho_1,\rho_0) - W_2^2(\rho_0,\rho_1)\Big)\\
&\quad= \frac{1}{2t}\Big(W_2^2(\mathsf{H}_t\rho_1,\rho_0) - W_2^2(\mathsf{H}_t^*\rho_1,\rho_0) + 2\int_X Q_1\phi_t\,\mathrm{d}\mathsf{H}_t^*\rho_1 - 2\int_X \phi_t\,\mathrm{d}\rho_0 - W_2^2(\rho_0,\rho_1)\Big)\\
&\quad\le \frac{1}{2t}\Big(W_2^2(\mathsf{H}_t\rho_1,\rho_0) - W_2^2(\mathsf{H}_t^*\rho_1,\rho_0) + 2\int_X Q_1\phi_t\,\mathrm{d}\mathsf{H}_t^*\rho_1 - 2\int_X Q_1\phi_t\,\mathrm{d}\mathsf{H}_t\rho_1\Big)\\
&\qquad\quad + \frac{1}{2t}\big(\mathrm{R}_{K_z}^2(t)-1\big)\,W_2^2(\rho_0,\rho_1) + \mathrm{Ent}_{\mathfrak{m}}(\rho_0) - \mathrm{Ent}_{\mathfrak{m}}(\mathsf{H}_t\rho_1)\\
&\qquad\quad - \frac{1}{t}\int_0^1 s\int_0^t\int_X \Gamma(Q_s\phi_t)\,\mathsf{P}_{s(t-r)}^{2(k-K_z)}\big((k-K_z)\,\mathsf{P}_{sr} f_{\vartheta_t(s)}\big)\,\mathrm{d}\mathfrak{m}\,\mathrm{d}r\,\mathrm{d}s,
\end{aligned}$$

where we have put $\vartheta_t := \vartheta_{K_z,t}$. Note that the $\limsup$ as $t\downarrow 0$ of the last term is nonnegative since $(k-K_z)\,f_s \ge 0$ $\mathfrak{m}$-a.e. on X for every $s\in[0,1]$ and

$$\lim_{t\downarrow 0}\frac{1}{t}\int_0^t \mathsf{P}_{s(t-r)}^{2(k-K_z)}\big((k-K_z)\,\mathsf{P}_{sr} f_{\vartheta_t(s)}\big)\,\mathrm{d}r = (k-K_z)\,f_s$$

w.r.t. convergence in $L^1(X,\mathfrak{m})$. Indeed, $\vartheta_t(s)\to s$ as $t\downarrow 0$ for every $s\in[0,1]$ and therefore $f_{\vartheta_t(s)}\to f_s$ pointwise $\mathfrak{m}$-a.e. As all considered functions are nonnegative and $\int_X f_{\vartheta_t(s)}\,\mathrm{d}\mathfrak{m} = \int_X f_s\,\mathrm{d}\mathfrak{m}$ for all $t>0$, we have $f_{\vartheta_t(s)}\to f_s$ in $L^1(X,\mathfrak{m})$ as $t\downarrow 0$. We conclude by strong continuity of the heat and the Schrödinger semigroup with potential $2(k-K_z)$ in $L^1(X,\mathfrak{m})$.

Lower semicontinuity of $\mathrm{Ent}_{\mathfrak{m}}$ yields $-\liminf_{t\downarrow 0}\mathrm{Ent}_{\mathfrak{m}}(\mathsf{H}_t\rho_1) \le -\mathrm{Ent}_{\mathfrak{m}}(\rho_1)$, and clearly $\mathrm{R}_{K_z}^2(t) = 1 - K_z t + \mathrm{o}(t)$ as $t\downarrow 0$. Lastly, observe that $\big(W_2^2(\mathsf{H}_t\rho_1,\rho_0) - W_2^2(\mathsf{H}_t^*\rho_1,\rho_0)\big)/2t \to 0$ according to Lemma 4.1 applied with $a := 2$. Thus, we finally deduce

$$\limsup_{t\downarrow 0}\frac{1}{2t}\big(W_2^2(\mathsf{H}_t\rho_1,\rho_0) - W_2^2(\rho_0,\rho_1)\big) + \frac{K_z}{2}W_2^2(\rho_0,\rho_1) \le \mathrm{Ent}_{\mathfrak{m}}(\rho_0) - \mathrm{Ent}_{\mathfrak{m}}(\rho_1). \qquad \square$$

Theorem 4.5. *The 2-gradient estimate* $\mathrm{GE}_2(k)$ *implies* $\mathrm{CD}(k,\infty)$.

Proof. Given $\varepsilon>0$, Proposition 4.4 translates into a "local" $\mathrm{EVI}(k-\varepsilon)$ property at time 0: for every $z\in X$, choosing $\delta>0$ and $K_z\in\mathbb{R}$ such that $K_z\le k\le K_z+\varepsilon$ in $B_{2\delta}(z)$, we obtain that for all $\mu,\nu\in\mathscr{P}_2(X)\cap\mathrm{Dom}(\mathrm{Ent}_{\mathfrak{m}})$ with support in $B_\delta(z)$ and bounded densities w.r.t. $\mathfrak{m}$, for $\boldsymbol{\pi}\in\mathscr{P}(\mathrm{Geo}(X))$ representing the W_2-geodesic from μ to ν, we have

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0}\frac{1}{2}W_2^2(\mathsf{H}_t\mu,\nu) + \int_0^1\int_{\mathrm{Geo}(X)}(1-s)\,\big(k(\gamma_s)-\varepsilon\big)\,|\dot\gamma|^2\,\mathrm{d}\boldsymbol{\pi}(\gamma)\,\mathrm{d}s \le \mathrm{Ent}_{\mathfrak{m}}(\nu) - \mathrm{Ent}_{\mathfrak{m}}(\mu).$$

With the same argument used in the proof of [31, Theorem 3.4] for the equivalence of $\mathrm{CD}(k,\infty)$ and $\mathrm{EVI}(k)$ (based on previous work [13] in the case of constant k), we conclude that this local $\mathrm{EVI}(k-\varepsilon)$ implies a "local" $\mathrm{CD}(k-\varepsilon,\infty)$ condition in the following sense: for all $z\in X$ there exists $\delta>0$ such that for all

$\mu_0, \mu_1 \in \mathscr{P}_2(X) \cap \mathrm{Dom}(\mathrm{Ent}_\mathfrak{m})$ with support in $B_\delta(z)$ and bounded densities w.r.t. $\mathfrak{m}$, if $\boldsymbol{\pi} \in \mathscr{P}(\mathrm{Geo}(X))$ represents the W_2-geodesic from μ_0 to μ_1, for every $t \in [0,1]$, we have

$$\mathrm{Ent}_\mathfrak{m}(\mu_t) \leq (1-t)\,\mathrm{Ent}_\mathfrak{m}(\mu_0) + t\,\mathrm{Ent}_\mathfrak{m}(\mu_1) - \int_0^1 \int_{\mathrm{Geo}(X)} \mathrm{g}(s,t)\,\big(k(\gamma_s) - \varepsilon\big)\,|\dot\gamma|^2 \,\mathrm{d}\boldsymbol{\pi}(\gamma)\,\mathrm{d}s.$$

Using the local-to-global property from [31, Theorem 3.7] and taking the limit $\varepsilon \downarrow 0$, noticing again that the choice of W_2-geodesics does not depend on ε, allows us to pass from this local $\mathrm{CD}(k-\varepsilon,\infty)$ property to $\mathrm{CD}(k-\varepsilon,\infty)$ and finally to $\mathrm{CD}(k,\infty)$. $\square$

4.3. From EVI *to a differential 2-transport estimate*

It has already been observed in [20] that $\mathrm{EVI}(k)$ yields contraction estimates for the 2-Kantorovich–Wasserstein distance along two heat flows starting at regular measures. For irregular initial data, we now aim in deducing a weak version of it, see also Remark 1.9.

Proposition 4.6. *The* $\mathrm{EVI}(k)$ *implies the following differential* 2*-transport estimates:*

(i) *for every* $\mu_1, \mu_2 \in \mathscr{P}_2(X) \cap \mathrm{Dom}(\mathrm{Ent}_\mathfrak{m})$, *one has*

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0} W_2^2(\mathsf{H}_t\mu_1, \mathsf{H}_t\mu_2) \leq -2\int_0^1 \int_{\mathrm{Geo}(X)} k(\gamma_s)\,|\dot\gamma|^2\,\mathrm{d}\boldsymbol{\pi}(\gamma)\,\mathrm{d}s, \tag{4.1}$$

where $\boldsymbol{\pi} \in \mathscr{P}(\mathrm{Geo}(X))$ *represents the* W_2*-geodesic from* μ_1 *to* μ_2, *and*

(ii) *for all* $x, y \in X$,

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0} W_2^2(\mathsf{H}_t\delta_x, \mathsf{H}_t\delta_y) \leq -2\underline{k}(x,y)\,\mathsf{d}^2(x,y).$$

Proof. Concerning (i), up to truncating k and using monotone convergence afterwards, we may assume that k is bounded. Naively, the claim follows by applying the $\mathrm{EVI}(k)$ to $(\mathsf{H}_t\mu_1)_{t\geq 0}$ and $(\mathsf{H}_t\mu_2)_{t\geq 0}$, respectively. Some care, however, is needed to deal with the double t-dependence of the nonsmooth function $t \mapsto W_2^2\big(\mathsf{H}_t\rho_0, \mathsf{H}_t\rho_1\big)$. To deal with this, one adds up the $\mathrm{EVI}(k)$, integrated from t to $t+h$, $h>0$, for the flow $(\mathsf{H}_t\mu_1)_{t\geq 0}$ with observation point $\mathsf{H}_{t+h}\mu_2$ and for the flow $(\mathsf{H}_t\mu_2)_{t\geq 0}$ with observation point $\mathsf{H}_t\mu_1$. The entropy terms cancel out, and we obtain the desired estimate after dividing by h and letting $h \downarrow 0$. See [20, Theorem 6.1] for details.

Next, we show (ii). Denote by $\underline{k}_n \in \mathrm{Lip}_\mathrm{b}(X \times X)$ a sequence converging pointwise from below in a monotone way to $\underline{k}$, see Lemma 2.1, and put $k_n(x) := \underline{k}_n(x,x)$ for $x \in X$. Given $x, y \in X$ and $t > 0$, select $\tau_* > 0$ small enough so that, for every $\tau \in (0, \tau_*)$,

$$W_2^2(\mathsf{H}_\tau\delta_x, \mathsf{H}_\tau\delta_y) \leq \mathsf{d}^2(x,y) + 2t^2.$$

The local absolute continuity of the curves $(\mathsf{H}_t\delta_x)_{t\geq 0}$ and $(\mathsf{H}_t\delta_y)_{t\geq 0}$ on $(0,\infty)$ w.r.t. W_2 and property (i) with k_n in place of k, since $k_n \leq k$ on X, yield

$$\frac{1}{2t}\big(W_2^2(\mathsf{H}_t\delta_x, \mathsf{H}_t\delta_y) - \mathsf{d}^2(x,y)\big) \leq t + \frac{1}{2t}\big(W_2^2(\mathsf{H}_t\delta_x, \mathsf{H}_t\delta_y) - W_2^2(\mathsf{H}_\tau\delta_x, \mathsf{H}_\tau\delta_y)\big)$$

$$\leq t - \frac{1}{t}\int_\tau^t \int_0^1 \int_{\mathrm{Geo}(X)} k_n(\gamma_s)\,|\dot\gamma|^2 \,\mathrm{d}\boldsymbol{\pi}_r(\gamma)\,\mathrm{d}s\,\mathrm{d}r,$$

where $\boldsymbol{\pi}_r \in \mathscr{P}(\mathrm{Geo}(X))$ represents the W_2-geodesic from $\mathsf{H}_r\delta_x$ to $\mathsf{H}_r\delta_y$. As $n \to \infty$, by monotone convergence, the above inequality still holds with k in place of k_n. Thus, the definition of $\underline{k}$ and the inequality $\underline{k}_n \leq \underline{k}$ on X for every $n \in \mathbb{N}$ give, setting $\pi_r := (\mathsf{e}_0, \mathsf{e}_1)_\sharp \boldsymbol{\pi}_r$,

$$\frac{1}{2t}\big(W_2^2(\mathsf{H}_t\delta_x, \mathsf{H}_t\delta_y) - \mathsf{d}^2(x,y)\big) \leq t - \frac{1}{t}\int_\tau^t \int_{X\times X} \underline{k}_n(x',y')\,\mathsf{d}^2(x',y')\,\mathrm{d}\pi_r(x',y')\,\mathrm{d}r.$$

Since $\mathsf{H}_r\delta_x \to \delta_x$ and $\mathsf{H}_r\delta_y \to \delta_y$ w.r.t. W_2 as $r \to 0$ and since $W_2(\mathsf{H}_r\delta_x, \mathsf{H}_r\delta_y)$ is bounded uniformly in for small r, stability of optimal couplings, see [1, Proposition 7.1.3], and uniqueness of the W_2-optimal coupling $\pi_0 := \delta_x \otimes \delta_y$ imply that $\pi_r \to \pi_0$ weakly as $r \to 0$. Thus, the map $r \mapsto \int_{X\times X} \underline{k}_n\,\mathsf{d}^2\,\mathrm{d}\pi_r$ is continuous at 0 by [33, Lemma 4.3]. The claim follows by taking successively $\tau \downarrow 0$, $t \downarrow 0$ and $n \to \infty$ in the above inequality. □

A posteriori, knowing from Theorem 1.1 that EVI(k) implies $\mathrm{GE}_1(k)$, we will be able to improve the bound (ii) from Proposition 4.6 even for exponents different from 2, see Remark 5.12 below.

5. Duality of p-transport estimates and q-gradient estimates

Throughout the rest of this article, given $t \geq 0$, we use the short-hand notation $\Pi_t := \mathrm{C}([0,t]; X \times X)$. Moreover, at several instances we consider a function $\underline{\ell}\colon X \times X \to \mathbb{R}$ which, unless stated otherwise, is assumed lower semicontinuous and lower bounded. However, it should practically rather be thought of as a bounded Lipschitz function "approximating" $\underline{k}$ from below without being of the particular form (1.2). This often allows us to assume that $\underline{\ell} \in \mathrm{Lip_b}(X \times X)$, while $\underline{k}$ is not continuous in general, even if k is Lipschitz.

5.1. *Perturbed costs and coupled Brownian motions*

Given any $p \in [1,\infty)$ and $\mu_1, \mu_2 \in \mathscr{P}_p(X)$, let us define the *perturbed p-transport cost with potential $-p\underline{\ell}$* at $t \geq 0$ by

$$W_p^{\underline{\ell}}(\mu_1,\mu_2,t) := \inf_{(\mathbb{P},\mathsf{b}^1,\mathsf{b}^2)} \mathbb{E}\Big[\mathrm{e}^{\int_0^{2t} p\underline{\ell}(\mathsf{b}_r^1,\mathsf{b}_r^2)/2\,\mathrm{d}r}\,\mathsf{d}^p\big(\mathsf{b}_{2t}^1,\mathsf{b}_{2t}^2\big)\Big]^{1/p}, \tag{5.1}$$

where the infimum is taken over all pairs of coupled Brownian motions $(\mathbb{P},\mathsf{b}^1)$ and $(\mathbb{P},\mathsf{b}^2)$ on X, restricted to $[0,2t]$ and modeled on a common probability space, with initial distributions μ_1 and μ_2, respectively. In more analytic words,

$$W_p^{\underline{\ell}}(\mu_1,\mu_2,t) = \inf_{\boldsymbol{\nu}}\Big(\int_{\Pi_{2t}} \mathrm{e}^{\int_0^{2t} p\underline{\ell}(\gamma_r^1,\gamma_r^2)/2\,\mathrm{d}r}\,\mathsf{d}^p\big(\gamma_{2t}^1,\gamma_{2t}^2\big)\,\mathrm{d}\boldsymbol{\nu}(\gamma)\Big)^{1/p}, \tag{5.2}$$

the infimum being taken over all $\boldsymbol{\nu} \in \mathscr{P}(\Pi_{2t})$ whose marginals $\boldsymbol{\nu}_1, \boldsymbol{\nu}_2 \in \mathscr{P}(\mathrm{C}([0,2t];X))$ are the laws of Brownian motions on X, restricted to $[0,2t]$, with initial distribution μ_1 and μ_2, respectively. If $\underline{\ell} = \underline{k}$, this is the usual perturbed p-transport cost from Section 1.4.

A natural, albeit nontrivial identity relates the perturbed p-transport cost in the case of constant k with the usual p-transport cost.

Lemma 5.1. *If $\underline{\ell}$ is constantly equal to $L \in \mathbb{R}$ then, for $t \geq 0$,*

$$W_p^{\underline{\ell}}(\mu_1, \mu_2, t) = \mathrm{e}^{Lt}\, W_p(\mathsf{H}_t\mu_1, \mathsf{H}_t\mu_2).$$

Proof. Since $W_p(\mathsf{H}_t\mu_1, \mathsf{H}_t\mu_2)^{1/p} = \inf_{(\mathsf{x},\mathsf{y})} \mathbb{E}\big[\mathsf{d}^p(\mathsf{x},\mathsf{y})\big]^{1/p}$, the infimum ranging over pairs of random variables $\mathsf{x} \sim \mathsf{H}_t\mu_1$ and $\mathsf{y} \sim \mathsf{H}_t\mu_2$ defined on a common probability space $(\Omega, \mathscr{F}, \mathbb{P})$, and as $\mathsf{b}_{2t} \sim \mathsf{H}_t\mu$ for every Brownian motion $(\mathbb{P}, \mathsf{b})$ with initial distribution $\mu \in \mathscr{P}(X)$, we get

$$W_p^{\underline{\ell}}(\mu_1, \mu_2, t) \geq \mathrm{e}^{Lt}\, W_p(\mathsf{H}_t\mu_1, \mathsf{H}_t\mu_2).$$

For the converse inequality, let $\pi_t \in \mathscr{P}(X \times X)$ be a W_p-optimal coupling of $\mathsf{H}_t\mu_1$ and $\mathsf{H}_t\mu_2$. Consider Brownian motions $\big(\mathbb{P}_1, \mathsf{b}^1\big)$ and $\big(\mathbb{P}_2, \mathsf{b}^2\big)$, restricted to $[0, 2t]$, starting at μ_1 and μ_2, defined on probability spaces $(\Omega_1, \mathscr{F}_1, \mathbb{P}_1)$ and $(\Omega_2, \mathscr{F}_2, \mathbb{P}_2)$, respectively. Define the "bridge measures" $\mathbb{P}_1^x$ for $x \in X$ by disintegrating $\mathbb{P}_1$ w.r.t. $\mathsf{H}_t\mu_1(\mathrm{d}x)$ or, in other words, by conditioning b^1 on the event $\{\mathsf{b}^1_{2t} = x\}$. Similarly, let $\mathbb{P}_2^y$ for $y \in X$ be the disintegration of $\mathbb{P}_2$ w.r.t. $\mathsf{H}_t\mu_2(\mathrm{d}y)$. Consider the "glued measure" $\widetilde{\mathbb{P}}$ defined by

$$\widetilde{\mathbb{P}} := \int_{X \times X} \mathbb{P}_1^x \otimes \mathbb{P}_2^y \,\mathrm{d}\pi_t(x, y)$$

on $\Omega := \Omega_1 \times \Omega_2$. Then $\big(\widetilde{\mathbb{P}}, \mathsf{b}^1\big)$ and $\big(\widetilde{\mathbb{P}}, \mathsf{b}^2\big)$ is a pair of coupled Brownian motions with joint distribution π_t at time $2t$. The desired inequality then follows directly, since

$$\widetilde{\mathbb{E}}\big[\mathsf{d}^p\big(\mathsf{b}^1_{2t}, \mathsf{b}^2_{2t}\big)\big] = \int_{X \times X} \mathsf{d}^p(x, y)\,\mathrm{d}\pi_t(x, y) = W_p^p(\mathsf{H}_t\mu_1, \mathsf{H}_t\mu_2). \quad \square$$

Lemma 5.2. *For every $p \in [1, \infty)$, $t \geq 0$ and $\mu_1, \mu_2 \in \mathscr{P}_p(X)$ as above, the infima in (5.1) and in (5.2) are attained.*

Moreover, for every sequence of lower semicontinuous functions $\underline{\ell}_n : X \times X \to \mathbb{R}$ converging pointwise to $\underline{\ell}$ from below in an increasing way, we have

$$\lim_{n\to\infty} W_p^{\underline{\ell}_n}(\mu_1, \mu_2, t) = W_p^{\underline{\ell}}(\mu_1, \mu_2, t).$$

Proof. The lower semicontinuity of $\underline{\ell}$ implies the one of

$$\gamma \quad \longmapsto \quad \mathrm{e}^{\int_0^{2t} p\underline{\ell}(\gamma_r^1, \gamma_r^2)/2\,\mathrm{d}r}\, \mathsf{d}^p\big(\gamma^1_{2t}, \gamma^2_{2t}\big)$$

w.r.t. the uniform topology on Π_{2t} which in turn implies weak lower semicontinuity of

$$\boldsymbol{\nu} \quad \longmapsto \quad \int_{\Pi_{2t}} \mathrm{e}^{\int_0^{2t} p\underline{\ell}(\gamma_r^1, \gamma_r^2)/2\,\mathrm{d}r}\, \mathsf{d}^p\big(\gamma^1_{2t}, \gamma^2_{2t}\big)\,\mathrm{d}\boldsymbol{\nu}(\gamma)$$

in $\mathscr{P}(\Pi_{2t})$. This gives the existence of a minimizer for (5.2) by a standard argument since, according to [33, Lemma 4.4], the family of $\boldsymbol{\nu} \in \mathscr{P}(\Pi_{2t})$ with given marginals is tight as the sets of marginals are both singletons.

The second assertion is a standard argument via Γ-convergence of the functionals whose infima give $W_p^{\underline{\ell}_n}(\mu_1, \mu_2, t)$ and $W_p^{\underline{\ell}}(\mu_1, \mu_2, t)$, respectively, in $\mathscr{P}(\Pi_{2t})$. $\square$

Let us denote by $\mathscr{B}^\nu(X \times X)$ the completion of the Borel σ-field on $X \times X$ w.r.t. a given $\nu \in \mathscr{P}(X \times X)$, and then

$$\mathscr{B}^{\mathrm{univ}}(X \times X) := \bigcap_{\nu \in \mathscr{P}(X\times X)} \mathscr{B}^{\nu}(X \times X)$$

is the σ-field of all *universally measurable* subsets of $X \times X$.

Lemma 5.3. *For every $t \geq 0$ and $p \in [1,\infty)$, there exists a universally measurable map*

$$\boldsymbol{\eta}^t\colon \quad X \times X \quad \longrightarrow \quad \mathscr{P}(\Pi_{2t})$$

such that for every $x, y \in X$, the marginals of $\boldsymbol{\eta}^t_{x,y} := \boldsymbol{\eta}^t(x,y)$ are laws of Brownian motions, restricted to $[0, 2t]$, starting in x and y, respectively, and $\boldsymbol{\eta}^t_{x,y}$ is a minimizer in the definition (5.2) of $W_p^{\underline{\ell}}(\delta_x, \delta_y, t)$.

Proof. According to Lemma 5.2, for each pair $(x, y) \in X \times X$ there exists an admissible measure on $\mathscr{P}(\Pi_{2t})$ which attains the infimum in (5.2). The class of all probability measures with this property is closed. Then a measurable selection argument, see [11, Theorem 6.9.2] and [31, Lemma 2.2], allows us to produce a family of measures $\boldsymbol{\eta}^t_{x,y}$ still satisfying the minimality property so that $(x, y) \mapsto \boldsymbol{\eta}^t_{x,y}$ is universally measurable in $(x, y) \in X \times X$. □

An important consequence of these observations is a type of Markov property which will be crucial in the proof of Proposition 5.6. For this and also for later use, fix $s, t \geq 0$, a measure $\boldsymbol{\nu} \in \mathscr{P}(\Pi_s)$ and a universally measurable map $\boldsymbol{\mu}\colon X \times X \to \mathscr{P}(\Pi_t)$ such that $(\mathsf{e}_0)_\sharp \boldsymbol{\mu}_{x,y} = \delta_x \otimes \delta_y$ for all $x, y \in X$. Define their composition $\boldsymbol{\mu} \circ \boldsymbol{\nu} \in \mathscr{P}(\Pi_{s+t})$ by

$$\int_{\Pi_{s+t}} f(\gamma)\,\mathrm{d}(\boldsymbol{\mu}\circ\boldsymbol{\nu})(\gamma) := \int_{\Pi_s}\int_{\Pi_t} f\big(\Phi_{s,t}(\alpha,\beta)\big)\,\mathrm{d}\boldsymbol{\mu}_{\alpha_s^1,\alpha_s^2}(\beta)\,\mathrm{d}\boldsymbol{\nu}(\alpha) \quad \text{for every } f \in \mathrm{C_b}(\Pi_{s+t}),$$

where

$$\Phi_{s,t}(\alpha,\beta)_r := \alpha_r \quad \text{if } r \in [0,s] \quad \text{and} \quad \Phi_{s,t}(\alpha,\beta)_r := \beta_{r-s} \quad \text{if } r \in (s, s+t]$$

denotes the *concatenation map* “gluing” together the curves $(\alpha_\sigma)_{\sigma\in[0,s]}$ and $(\beta_\tau)_{\tau\in[0,t]}$.

Proposition 5.4. *For every $p \in [1,\infty)$, every $s, t \geq 0$ and all $\mu_1, \mu_2 \in \mathscr{P}_p(X)$, there exists a pair $\big(\mathbb{P}, \mathsf{b}^1\big)$ and $\big(\mathbb{P}, \mathsf{b}^2\big)$ of coupled Brownian motions on X with initial distributions μ_1 and μ_2, respectively, which minimizes (5.1) for the given time t and such that*

$$W_p^{\underline{\ell}}(\mu_1,\mu_2,t+s)^p \leq \mathbb{E}\Big[\mathrm{e}^{\int_0^{2t} p\underline{\ell}(\mathsf{b}_r^1,\mathsf{b}_r^2)/2\,\mathrm{d}r}\, W_p^{\underline{\ell}}\big(\delta_{\mathsf{b}^1_{2t}}, \delta_{\mathsf{b}^2_{2t}}, s\big)^p\Big]. \tag{5.3}$$

Proof. Denote the map from Lemma 5.3 with s in place of t by $\boldsymbol{\eta}^s$, denote a minimizer of (5.2) for time t by $\boldsymbol{\nu}_t$, and define $\boldsymbol{\eta}^{t+s} := \boldsymbol{\eta}^s \circ \boldsymbol{\nu}_t \in \mathscr{P}(\Pi_{2(s+t)})$. This defines a coupling of the laws of two Brownian motions with initial distributions μ_1 and μ_2, respectively, restricted to $[0, 2(t+s)]$ such that

$$\begin{aligned}&\int_{\Pi_{2(s+t)}} \mathrm{e}^{\int_0^{2(t+s)} p\underline{\ell}(\gamma_r^1,\gamma_r^2)/2\,\mathrm{d}r}\,\mathsf{d}^p\big(\gamma^1_{2(t+s)},\gamma^2_{2(t+s)}\big)\,\mathrm{d}\boldsymbol{\nu}^{t+s}(\gamma)\\ &\qquad = \int_{\Pi_{2t}} \mathrm{e}^{\int_0^{2t} p\underline{\ell}(\alpha_r^1,\alpha_r^2)/2\,\mathrm{d}r}\, W_p^k\big(\delta_{\alpha^1_{2t}},\delta_{\alpha^2_{2t}},s\big)^p\,\mathrm{d}\boldsymbol{\nu}_t(\alpha).\end{aligned}$$

This proves the claim. □

Less formally, the previous construction can be described as follows. To estimate the perturbed p-transport cost at time $t+s$, we construct the required process by first choosing a pair process $(\mathsf{b}^1, \mathsf{b}^2)$ of Brownian motions with given initial distributions μ_1 and μ_2 which realizes the minimum for $W_p^{\underline{\ell}}(\mu_1, \mu_2, t)$. Then we switch to a pair of Brownian motions starting in b^1_{2t} and b^2_{2t}, respectively, which minimizes the cost at time s.

5.2. Differential p-transport inequalities and p-transport estimates

To deduce a p-transport estimate $\mathrm{TE}_p(k)$, we have to control the upper derivatives of the function $t \mapsto W_p^{\underline{k}}(\delta_x, \delta_y, t)^p$ or, more generally, of $t \mapsto W_p^{\underline{\ell}}(\delta_x, \delta_y, t)^p$ for $x, y \in X$.

Lemma 5.5. *Assume that $\underline{\ell} \in \mathrm{C_b}(X \times X)$. Then for all $x, y \in X$ and $p \in [1, \infty)$, we have*

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0} W_p^{\underline{\ell}}(\delta_x, \delta_y, t)^p \le p\, \underline{\ell}(x,y)\, \mathsf{d}^p(x,y) + \frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0} W_p^p(\mathsf{H}_t\delta_x, \mathsf{H}_t\delta_y).$$

Proof. Choose any exponent $p' \in (p, \infty)$ with dual exponent $q' \in (1, \infty)$. For all $t > 0$, denote by $(\mathbb{P}, \mathsf{b}^1)$ and $(\mathbb{P}, \mathsf{b}^2)$ a pair of coupled Brownian motions starting in (x, y) and such that the law of $(\mathsf{b}^1_{2t}, \mathsf{b}^2_{2t})$ constitutes a $W_{p'}$-optimal coupling of $\mathsf{H}_t\delta_x$ and $\mathsf{H}_t\delta_y$. Albeit this process still depends on t, we suppress this dependence in the sequel to simplify the notation. For a precise construction of such process, we refer to the proof of Lemma 5.1.

Observe that

$$\begin{aligned}
&\frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0} W_p^{\underline{\ell}}(\delta_x, \delta_y, t)^p \\
&\qquad \le \limsup_{t\downarrow 0} \frac{1}{t}\mathbb{E}\Big[\mathrm{e}^{\int_0^{2t} p\underline{\ell}(\mathsf{b}^1_r, \mathsf{b}^2_r)/2\,\mathrm{d}r}\, \mathsf{d}^p\big(\mathsf{b}^1_{2t}, \mathsf{b}^2_{2t}\big) - \mathsf{d}^p\big(\mathsf{b}^1_0, \mathsf{b}^2_0\big)\Big] \\
&\qquad \le \limsup_{t\downarrow 0} \frac{1}{t}\mathbb{E}\Big[\Big(\mathrm{e}^{\int_0^{2t} p\underline{\ell}(\mathsf{b}^1_r, \mathsf{b}^2_r)/2\,\mathrm{d}r} - 1\Big)\, \mathsf{d}^p\big(\mathsf{b}^1_{2t}, \mathsf{b}^2_{2t}\big)\Big] + \frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0} W_{p'}^p(\mathsf{H}_t\delta_x, \mathsf{H}_t\delta_y).
\end{aligned}$$

Each of the last two limits will be estimated separately. The last term will converge to the upper derivative of $W_p^p(\mathsf{H}_t\delta_x, \mathsf{H}_t\delta_y)$ at 0 as $p' \downarrow p$ by monotone convergence. Moreover, since $\underline{\ell}$ is bounded, the former term can be estimated through

$$\limsup_{t\downarrow 0} \frac{1}{t}\mathbb{E}\Big[\Big(\mathrm{e}^{\int_0^{2t} p\underline{\ell}(\mathsf{b}^1_r, \mathsf{b}^2_r)/2\,\mathrm{d}r} - 1\Big)\, \mathsf{d}^p\big(\mathsf{b}^1_{2t}, \mathsf{b}^2_{2t}\big)\Big] \le \limsup_{t\downarrow 0} \frac{p}{2t}\mathbb{E}\Big[\int_0^{2t} \underline{\ell}\big(\mathsf{b}^1_r, \mathsf{b}^2_r\big)\,\mathrm{d}r\, \mathsf{d}^p\big(\mathsf{b}^1_{2t}, \mathsf{b}^2_{2t}\big)\Big].$$

Now we split the expectation into a term where $(\mathsf{b}^1, \mathsf{b}^2)$ behaves well and a remainder term. Let $\varepsilon > 0$ and choose $\delta > 0$ such that

$$\max\Big\{\big|\underline{\ell}(x',y') - \underline{\ell}(x,y)\big|, \big|\,\mathsf{d}^p(x',y') - \mathsf{d}^p(x,y)\big|\Big\} \le \varepsilon \quad \text{for every } x' \in B_\delta(x),\ y' \in B_\delta(y),$$

and define the exceptional set $E_{r,2t}$ for $r \in (0, 2t)$ by

$$E_{r,2t} := \big\{\mathsf{b}^1_r \notin B_\delta(x)\big\} \cup \big\{\mathsf{b}^1_{2t} \notin B_\delta(x)\big\} \cup \big\{\mathsf{b}^2_r \notin B_\delta(y)\big\} \cup \big\{\mathsf{b}^2_{2t} \notin B_\delta(y)\big\}.$$

By these definitions and Fubini's theorem, since $\underline{\ell}$ is bounded,

$$\limsup_{t\downarrow 0} \frac{p}{2t}\mathbb{E}\Big[\int_0^{2t} \underline{\ell}\big(\mathsf{b}_r^1,\mathsf{b}_r^2\big)\,\mathbb{1}_{E_{r,2t}^{\mathrm{c}}}\,\mathrm{d}r\,\mathsf{d}^p\big(\mathsf{b}_{2t}^1,\mathsf{b}_{2t}^2\big)\Big]$$

$$\leq p\,\big(\underline{\ell}(x,y)+\varepsilon\big)\,\big(\mathsf{d}^p(x,y)+\varepsilon\big)\,\limsup_{t\downarrow 0}\frac{1}{2t}\int_0^{2t}\mathbb{P}\big[E_{r,2t}^{\mathrm{c}}\big]\,\mathrm{d}r.$$

According to Lemma 4.2, we have $\mathbb{P}[E_{r,2t}] \to 0$ as $r\downarrow 0$ and $t\downarrow 0$, therefore the latter lim sup is equal to 1. On the other hand, if $C>0$ denotes an upper bound for $\underline{\ell}$, using Hölder's inequality the second term can be bounded through

$$\limsup_{t\downarrow 0}\left|\frac{p}{2t}\mathbb{E}\Big[\int_0^{2t}\underline{\ell}\big(\mathsf{b}_r^1,\mathsf{b}_r^2\big)\,\mathbb{1}_{E_{r,2t}}\,\mathrm{d}r\;\mathsf{d}^p\big(\mathsf{b}_{2t}^1,\mathsf{b}_{2t}^2\big)\Big]\right|$$

$$\leq p\,C\,\limsup_{t\downarrow 0}\mathbb{E}\Big[\mathsf{d}^{p'}\big(\mathsf{b}_{2t}^1,\mathsf{b}_{2t}^2\big)\Big]^{p/p'}\;\limsup_{t\downarrow 0}\Big(\frac{1}{2t}\int_0^{2t}\mathbb{P}\big[E_{r,2t}\big]\,\mathrm{d}r\Big)^{1-p/p'}.$$

By the choice of the pair process $(\mathsf{b}^1,\mathsf{b}^2)$, the first lim sup is equal to $\mathsf{d}^p(x,y)$ while the second one is 0, as already observed above. Since ε was arbitrary, we obtain the claim. □

Proposition 5.6. *Fix $p\in[1,\infty)$ and assume the differential p-transport estimate*

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0} W_p^p(\mathsf{H}_t\delta_x,\mathsf{H}_t\delta_y) \leq -p\,\underline{k}(x,y)\,\mathsf{d}^p(x,y) \quad \textit{for every } x,y\in X. \tag{5.4}$$

Then the p-transport estimate $\mathrm{TE}_p(k)$ is satisfied.

Proof. We first show that for all $\mu_1,\mu_2\in\mathscr{P}_p(X)$, the function $t\mapsto W_p^{\underline{\ell}}(\mu_1,\mu_2,t)$ is nonincreasing on $[0,\infty)$ whenever $\underline{\ell}\in\mathrm{C_b}(X\times X)$ with $\underline{\ell}\leq\underline{k}$ on $X\times X$.

To get started, we demonstrate that its p-th power $t\mapsto W_p^{\underline{\ell}}(\mu_1,\mu_2,t)^p$ is upper Lipschitz continuous on $[0,\infty)$. To see this, fix $h\in(0,1]$ and $t>0$, and consider the pair process $(\mathsf{b}^1,\mathsf{b}^2)$ as provided by Proposition 5.4. By the estimate (5.3) of this proposition, Lemma 5.1 and contractivity of the Wasserstein heat flow, we have

$$\begin{aligned}
&\frac{1}{h}\big(W_p^{\underline{\ell}}(\mu_1,\mu_2,t+h)^p - W_p^{\underline{\ell}}(\mu_1,\mu_2,t)^p\big)\\
&\quad\leq \frac{1}{h}\mathbb{E}\Big[\mathrm{e}^{\int_0^{2t} p\underline{\ell}(\mathsf{b}_r^1,\mathsf{b}_r^2)/2\,\mathrm{d}r}\Big(W_p^{\underline{\ell}}\big(\delta_{\mathsf{b}_{2t}^1},\delta_{\mathsf{b}_{2t}^1},h\big)^p - \mathsf{d}^p\big(\mathsf{b}_{2t}^1,\mathsf{b}_{2t}^2\big)\Big)\Big]\\
&\quad\leq \frac{1}{h}\mathbb{E}\Big[\mathrm{e}^{\int_0^{2t} p\underline{\ell}(\mathsf{b}_r^1,\mathsf{b}_r^2)/2\,\mathrm{d}r}\,\mathsf{d}^p\big(\mathsf{b}_{2t}^1,\mathsf{b}_{2t}^2\big)\,\big(\mathrm{e}^{pCh}-1\big)\Big] \leq C'\,W_p^{\underline{\ell}}(\mu_1,\mu_2,t)^p
\end{aligned}\tag{5.5}$$

for suitable nonnegative constants C and C'. This proves upper Lipschitz continuity of the p-th power of the perturbed p-transport cost with potential $-p\underline{\ell}$, which in turn implies

$$W_p^{\underline{\ell}}(\mu_1,\mu_2,\tau)^p - W_p^{\underline{\ell}}(\mu_1,\mu_2,\sigma)^p \leq \int_\sigma^\tau \frac{\mathrm{d}^+}{\mathrm{d}t}W_p^{\underline{\ell}}\big(\mu_1,\mu_2,t\big)^p\,\mathrm{d}t \tag{5.6}$$

for every $\sigma,\tau\in[0,\infty)$ with $\sigma\leq\tau$. Letting $h\downarrow 0$, the estimate (5.5) and the observation

$$\mathbb{E}\Big[\mathrm{e}^{\int_0^{2t} p\underline{\ell}(\mathsf{b}_r^1,\mathsf{b}_r^2)/2\,\mathrm{d}r}\,\mathsf{d}^p\big(\mathsf{b}_{2t}^1,\mathsf{b}_{2t}^2\big)\Big] < \infty,$$

which justifies to apply Fatou's lemma, give

$$\frac{\mathrm{d}^+}{\mathrm{d}t}W_p^{\underline{\ell}}(\mu_1,\mu_2,t)^p \le \mathbb{E}\Big[\mathrm{e}^{\int_0^{2t} p\underline{\ell}(\mathsf{b}_r^1,\mathsf{b}_r^2)/2\,\mathrm{d}r}\,\frac{\mathrm{d}^+}{\mathrm{d}h}\Big|_{h=0} W_p^{\underline{\ell}}\big(\delta_{\mathsf{b}_{2t}^1},\delta_{\mathsf{b}_{2t}^2},h\big)^p\Big].$$

Finally, the inequality (5.6) for the upper derivative inside the expectation, Lemma 5.5 and then the assumed estimate (5.4), noting that $-\underline{k}\le-\underline{\ell}$ on $X\times X$, yield the initial claim.

The nonincreasingness of $t\mapsto W_p^{\underline{k}}(\mu_1,\mu_2,t)$ on $[0,\infty)$ is then immediate due to an easy approximation argument using Lemma 2.1 and Lemma 5.2. □

Theorem 5.7. *For every $p\in[1,\infty)$, $\mathrm{TE}_p(k)$ and the differential p-transport estimate (5.4) are equivalent.*

Proof. According to Proposition 5.6, it suffices to prove that $\mathrm{TE}_p(k)$ implies

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0} W_p^p(\mathsf{H}_t\delta_x,\mathsf{H}_t\delta_y) \le -p\,\underline{k}(x,y)\,\mathsf{d}^p(x,y)$$

for every $x,y\in X$. For every $t>0$ and $p'\in(p,\infty)$, we denote by $\big(\mathbb{P},\mathsf{b}^1\big)$ and $\big(\mathbb{P},\mathsf{b}^2\big)$ a pair of coupled Brownian motions which realizes the minimum in the definition of $W_{p'}^{\underline{\ell}}(\delta_x,\delta_y,t)$. This process does depend on t, but we leave out this dependency from the notation. Arguing as in the proof of Lemma 5.5, we get

$$\begin{aligned}\frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0} & W_p^p(\mathsf{H}_t\delta_x,\mathsf{H}_t\delta_y)\\ &\le \limsup_{t\downarrow 0}\frac{1}{t}\mathbb{E}\Big[\Big(1-\mathrm{e}^{\int_0^{2t} p\underline{\ell}(\mathsf{b}_r^1,\mathsf{b}_r^2)/2\,\mathrm{d}r}\Big)\,\mathsf{d}^p\big(\mathsf{b}_{2t}^1,\mathsf{b}_{2t}^2\big)\Big] + \frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0} W_{p'}^{\underline{k}}(\delta_x,\delta_y,t)^p\\ &\le -p\,\underline{\ell}(x,y)\,\mathsf{d}^p(x,y) + \frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0} W_{p'}^{\underline{k}}(\delta_x,\delta_y,t)^p\end{aligned}$$

for all $\underline{\ell}\in\mathrm{C_b}(X\times X)$ with $\underline{\ell}\le\underline{k}$ on $X\times X$. Letting $p'\downarrow p$, the last upper derivative becomes nonpositive due to $\mathrm{TE}_p(k)$, and approximating $\underline{k}$ from below using Lemma 2.1 gives the conclusion. □

Using this equivalence, Hölder's inequality and the chain rule, the subsequent nestedness of $\mathrm{TE}_p(k)$, which is the Lagrangian analogue of Lemma 3.3, is easily shown.

Corollary 5.8. *If $\mathrm{TE}_p(k)$ holds for some $p\in[1,\infty)$, then $\mathrm{TE}_{p'}(k)$ is satisfied for all $p'\in[1,p]$.*

5.3. Transport estimates via vertical Brownian perturbations

We prove the variable Kuwada duality from Theorem 1.8. We start by first showing the implication from $\mathrm{GE}_q(k)$ to $\mathrm{TE}_p(k)$, where $p,q\in(1,\infty)$ are dual to each other. Since the behavior of Brownian trajectories can only be controlled for small times, we show the equivalent infinitesimal first-order description of $\mathrm{TE}_p(k)$ in terms of a differential p-transport estimate. This is done by a localization argument.

Additionally, in the extremal case $q=1$, the argument mentioned above can actually be circumvented and we are able to derive the contraction estimate

$$\frac{\mathrm{d}^+}{\mathrm{d}t}W_p^p(\mathsf{H}_t\mu,\mathsf{H}_t\nu) \le -p\int_0^1\int_{\mathrm{Geo}(X)} k(\gamma_s)\,|\dot\gamma|^p\,\mathrm{d}\boldsymbol{\pi}_t(\gamma)\,\mathrm{d}s \quad \text{for every } t\ge 0$$

for all $\mu, \nu \in \mathscr{P}(X)$ of finite W_p-distance to each other, for *every* $p \in (1,\infty)$. The measure $\boldsymbol{\pi}_t \in \mathscr{P}(\mathrm{Geo}(X))$ induces an *arbitrary* W_p-optimal coupling of $\mathsf{H}_t\mu$ and $\mathsf{H}_t\nu$. This is discussed now, see Theorem 5.10 and Corollary 5.11, where, possibly replacing k by $\min\{k,n\}$ for $n \in \mathbb{N}$, we assume that k is bounded. This is not restrictive as, given these results for every $n \in \mathbb{N}$, they easily pass to the limit $n \to \infty$ by monotone convergence.

Recall that $\mathrm{G}_0(x,y)$ denotes the set of geodesics from $x \in X$ to $y \in X$. Given $p \in (1,\infty)$ and $t \geq 0$, we define the function $\mathsf{d}^0_{p,k,t} \colon X \times X \to \mathbb{R}$ by

$$\mathsf{d}^0_{p,k,t}(x,y) := \inf_{\gamma \in \mathrm{G}_0(x,y)} \Big(\int_0^1 \mathbb{E}_{\gamma_s}\Big[\mathrm{e}^{-\int_0^{2t} pk(\mathsf{b}_r)/2\,\mathrm{d}r}\Big]\,|\dot\gamma|^p\,\mathrm{d}s\Big)^{1/p}.$$

Here $(\mathbb{P}_{\gamma_s}, \mathsf{b})$ denotes Brownian motion starting in γ_s for every $s \in [0,1]$. We will not explicitly mention the dependence of the process b on s. The function $\mathsf{d}^0_{p,k,t}$ can be turned into a metric $\mathsf{d}_{p,k,t}$ on X by defining

$$\mathsf{d}_{p,k,t}(x,y) := \inf\Big\{\sum_{i=1}^n \mathsf{d}^0_{p,k,t}(x_{i-1},x_i) : n \in \mathbb{N},\ x =: x_0 < x_1 < \cdots < x_n := y\Big\}.$$

It is equivalent to d by boundedness of k since d is a length metric. Let us denote by $W^0_{p,k,t}$ and $W_{p,k,t}$ the transport "distances" w.r.t. $\mathsf{d}^0_{p,k,t}$ and $\mathsf{d}_{p,k,t}$, respectively. Then $W_{p,k,t}$ is a metric on $\mathscr{P}_p(X)$, which is equivalent to the usual p-Kantorovich–Wasserstein metric W_p. Compared to the perturbed p-transport cost W_p^k which measures Brownian evolutions "horizontally" by following their trajectories with *fixed* starting points, the distance $W_{p,k,t}$ varies the initial points along a geodesic and may thus be seen as a "vertical" counterpart of W_p^k.

Let Q_s be the p-Hopf–Lax semigroup and $q \in (1,\infty)$ such that $1/p + 1/q = 1$. Similarly to [22, Proposition 3.7], the key point will be the following Lipschitz regularity along geodesics.

Lemma 5.9. *Let $f \in \mathrm{Lip_b}(X)$. Then for every $x,y \in X$, each $t > 0$ and all $\gamma \in \mathrm{G}_0(y,x)$, the map $s \mapsto \mathsf{P}_tQ_sf(\gamma_s)$ belongs to* $\mathrm{Lip}([0,1])$*, and*

$$\mathsf{P}_tQ_1f(x) - \mathsf{P}_tf(y) \leq \int_0^1 \Big(\limsup_{h\downarrow 0}\frac{1}{h}\big(\mathsf{P}_tQ_sf(\gamma_{s+h}) - \mathsf{P}_tQ_sf(\gamma_s)\big) - \frac{1}{q}\mathsf{P}_t\big(\mathrm{lip}(Q_sf)^q\big)(\gamma_s)\Big)\,\mathrm{d}s.$$

Proof. Let $h > 0$ and $s \in [0,1-h]$. Notice that

$$\begin{aligned}
&\frac{1}{h}\big|\mathsf{P}_tQ_{s+h}f(\gamma_{s+h}) - \mathsf{P}_tQ_sf(\gamma_s)\big| \\
&\quad\leq \frac{1}{h}\big|\mathsf{P}_tQ_{s+h}f(\gamma_{s+h}) - \mathsf{P}_tQ_{s+h}f(\gamma_s)\big| + \frac{1}{h}\Big|\int_X \big(Q_{s+h}f - Q_sf\big)\,\mathrm{d}\mathsf{H}_t\delta_{\gamma_s}\Big| \\
&\quad\leq \frac{\mathsf{d}(x,y)}{h}\int_s^{s+h} |\mathrm{D}\mathsf{P}_tQ_{s+h}f|(\gamma_v)\,\mathrm{d}v + \int_X \frac{1}{h}|Q_{s+h}f - Q_sf|\,\mathrm{d}\mathsf{H}_t\delta_{\gamma_s}.
\end{aligned}$$

The latter is bounded uniformly in s and h since the first integral can be controlled using the Lipschitz regularization estimate (2.1) of the heat flow while the second one exploits the fact that the map $s \mapsto Q_sf$ is Lipschitz from $[0,\infty)$ to $\mathrm{C}(X)$.

It follows that $\mathsf{P}_t Q_1 f(x) - \mathsf{P}_t f(y)$ is bounded from above by

$$\int_0^1 \Big(\limsup_{h\downarrow 0} \frac{1}{h}\big(\mathsf{P}_t Q_s f(\gamma_{s+h}) - \mathsf{P}_t Q_s f(\gamma_s)\big) + \limsup_{h\downarrow 0} \frac{1}{h} \int_X \big(Q_{s+h} f - Q_s f\big)\, \mathrm{d}\mathsf{H}_t \delta_{\gamma_{s+h}} \Big)\, \mathrm{d}s. \tag{5.7}$$

The Kantorovich–Rubinstein formula (2.3) for W_1, the W_1-contractivity of the heat flow and the duality of P_t and H_t give us the following upper bound for the second lim sup in (5.7)

$$\begin{aligned}
&\limsup_{h\downarrow 0} \frac{1}{h} \int_X \big(Q_{s+h} f - Q_s f\big)\, \mathrm{d}\big(\mathsf{H}_t\delta_{\gamma_{s+h}} - \mathsf{H}_t\delta_{\gamma_s}\big) + \limsup_{h\to 0} \frac{1}{h} \int_X \big(Q_{s+h} f - Q_s f\big)\, \mathrm{d}\mathsf{H}_t\delta_{\gamma_s} \\
&\qquad \le \mathrm{Lip}(Q_\bullet f)\, \limsup_{h\downarrow 0} W_1\big(\mathsf{H}_t\delta_{\gamma_{s+h}}, \mathsf{H}_t\delta_{\gamma_s}\big) + \int_X \frac{\mathrm{d}}{\mathrm{d}s} Q_s f\, \mathrm{d}\mathsf{H}_t\delta_{\gamma_s} \\
&\qquad = -\frac{1}{q} \int_X \mathrm{lip}(Q_s f)^q\, \mathrm{d}\mathsf{H}_t\delta_{\gamma_s} = -\frac{1}{q} \mathsf{P}_t\big(\mathrm{lip}(Q_s f)^q\big)(\gamma_s).
\end{aligned}$$

Here we used $\mathrm{Lip}(Q_\bullet f)$ as a shorthand for the Lipschitz constant of the map $s \mapsto Q_s f$ from $[0,\infty)$ to $\mathrm{C}(X)$. These estimates conclude the proof. □

Theorem 5.10. *Assume the* 1*-gradient estimate* $\mathrm{GE}_1(k)$*. Then for every* $p \in (1,\infty)$*,* $t \ge 0$ *and* $\mu,\nu \in \mathscr{P}(X)$*,*

$$W_p(\mathsf{H}_t\mu, \mathsf{H}_t\nu) \le W_{p,k,t}(\mu,\nu) \le W^0_{p,k,t}(\mu,\nu).$$

Proof. The second inequality is trivial since by definition $\mathsf{d}_{p,k,t} \le \mathsf{d}^0_{p,k,t}$, thus we concentrate on the first one.

Let us initially consider the case $\mu := \delta_x$ and $\nu := \delta_y$ for $x, y \in X$, and $t > 0$. By the duality (2.3), we have to estimate $\mathsf{P}_t Q_1 f(x) - \mathsf{P}_t f(y)$ from above for every $f \in \mathrm{Lip_b}(X)$. Pick a geodesic $\gamma \in \mathrm{G}_0(y,x)$. By the upper gradient property of $|\mathrm{D}\mathsf{P}_t Q_s f|$ and the $\mathrm{GE}_1(k)$ inequality, we deduce for $\mathscr{L}^1$-a.e. $s \in [0,1]$ that

$$\begin{aligned}
\limsup_{h\downarrow 0} \frac{1}{h}\big(\mathsf{P}_t Q_s f(\gamma_{s+h}) - \mathsf{P}_t Q_s f(\gamma_s)\big) &\le \limsup_{h\downarrow 0} \frac{\mathsf{d}(x,y)}{h} \int_s^{s+h} \mathsf{P}^k_t |\mathrm{D}Q_s f|(\gamma_v)\, \mathrm{d}v \\
&\le \mathsf{d}(x,y)\, \mathbb{E}_{\gamma_s}\Big[\mathrm{e}^{-\int_0^{2t} pk(\mathsf{b}_r)/2\, \mathrm{d}r}\Big]^{1/p} \mathsf{P}_t\big(\mathrm{lip}(Q_s f)^q\big)^{1/q}(\gamma_s),
\end{aligned}$$

denoting by $\big(\mathbb{P}_{\gamma_s}, \mathsf{b}\big)$ Brownian motion on X starting in γ_s. Invoking Lemma 5.9 and Young's inequality, we infer that

$$\mathsf{P}_t Q_1 f(x) - \mathsf{P}_t f(y) \le \frac{\mathsf{d}^p(x,y)}{p} \int_0^1 \mathbb{E}_{\gamma_s}\Big[\mathrm{e}^{-\int_0^{2t} pk(\mathsf{b}_r)/2\, \mathrm{d}r}\Big]\, \mathrm{d}s. \tag{5.8}$$

Taking the supremum over $f \in \mathrm{Lip_b}(X)$ and then infimizing over all geodesics γ connecting y to x, we deduce the inequality $W_p(\mathsf{H}_t\mu, \mathsf{H}_t\nu) \le W^0_{p,k,t}(\mu,\nu) = \mathsf{d}^0_{p,k,t,}(x,y)$. By the triangle inequality and the definition of $\mathsf{d}_{p,k,t}$, this already implies that $W_p(\mathsf{H}_t\mu, \mathsf{H}_t\nu) \le W_{p,k,t}(\mu,\nu) = \mathsf{d}_{p,k,t,}(x,y)$ under the above assumptions.

The case $t = 0$ follows by letting $t \downarrow 0$ in (5.8) and then concluding as above.

Lastly, the inequality $W_p(\mathsf{H}_t\mu, \mathsf{H}_t\nu) \le W_{p,k,t}(\mu,\nu)$ for general $\mu,\nu \in \mathscr{P}(X)$ follows by a standard coupling argument. Given any $W_{p,k,t}$-optimal coupling π of μ and ν, fix W_p-optimal couplings $\pi_{t,x,y}$ of $\mathsf{H}_t\delta_x$ and $\mathsf{H}_t\delta_y$,

$x, y \in X$, in a π-measurable way according to a measurable selection theorem [11, Theorem 6.9.2], and define a coupling of $\mathsf{H}_t\mu$ and $\mathsf{H}_t\nu$ by

$$\mathrm{d}\varpi_t(u,v) := \int_{X\times X} \mathrm{d}\pi_{t,x,y}(u,v)\,\mathrm{d}\pi(x,y).$$

Thus we obtain

$$\begin{aligned} W_p^p(\mathsf{H}_t\mu, \mathsf{H}_t\nu) &\leq \int_{X\times X} \mathsf{d}^p(u,v)\,\mathrm{d}\varpi_t(u,v) = \int_{X\times X} W_p^p(\mathsf{H}_t\delta_x, \mathsf{H}_t\delta_y)\,\mathrm{d}\pi(x,y) \\ &\leq \int_{X\times X} \mathsf{d}_{p,k,t}^p(x,y)\,\mathrm{d}\pi(x,y) = W_{p,k,t}^p(\mu,\nu). \quad \square \end{aligned}$$

With this in hand, we can proceed to what we have indicated in Remark 1.9.

Corollary 5.11. *Assume that* $\mathrm{GE}_1(k)$ *is satisfied. Let* $\mu, \nu \in \mathscr{P}(X)$ *so that* $W_p(\mu,\nu) < \infty$, *let* $t \geq 0$, *and let* $\boldsymbol{\pi}_t \in \mathscr{P}(\mathrm{Geo}(X))$ *represent an arbitrary* W_p*-optimal coupling between* $\mathsf{H}_t\mu$ *and* $\mathsf{H}_t\nu$, *i.e.* $(\mathsf{e}_0, \mathsf{e}_1)_\sharp\boldsymbol{\pi}_t$ *is a* W_p*-optimal coupling of* $\mathsf{H}_t\mu$ *and* $\mathsf{H}_t\nu$. *Then*

$$\frac{\mathrm{d}^+}{\mathrm{d}t} W_p^p(\mathsf{H}_t\mu, \mathsf{H}_t\nu) \leq -\int_0^1 \int_{\mathrm{Geo(X)}} k(\gamma_s)\,|\dot\gamma|^p\,\mathrm{d}\boldsymbol{\pi}_t(\gamma)\,\mathrm{d}s.$$

Proof. Given any optimal geodesic plan $\boldsymbol{\pi}_t$ as above, using Theorem 5.10 gives

$$\begin{aligned} &\limsup_{h\to 0} \frac{1}{ph}\big(W_p^p(\mathsf{H}_{t+h}\mu, \mathsf{H}_{t+h}\nu) - W_p^p(\mathsf{H}_t\mu, \mathsf{H}_t\nu)\big) \\ &\quad\leq \limsup_{h\downarrow 0} \frac{1}{ph}\big(W_{p,k,h}^0(\mathsf{H}_t\mu, \mathsf{H}_t\nu)^p - W_p^p(\mathsf{H}_t\mu, \mathsf{H}_t\nu)\big) \\ &\quad\leq \limsup_{h\downarrow 0} \frac{1}{ph} \int_{\mathrm{Geo}(X)} \left(\int_0^1 \mathbb{E}_{\gamma_s}\Big[\mathrm{e}^{-\int_0^{2h} pk(\mathsf{b}_r)/2\,\mathrm{d}r}\Big]\,\mathrm{d}s - 1\right) \mathsf{d}^p(\gamma_0,\gamma_1)\,\mathrm{d}\boldsymbol{\pi}_t(\gamma) \\ &\quad= -\int_0^1 \int_{\mathrm{Geo}(X)} k(\gamma_s)\,|\dot\gamma|^p\,\mathrm{d}\boldsymbol{\pi}_t(\gamma)\,\mathrm{d}s, \end{aligned}$$

where $(\mathbb{P}_{\gamma_s}, \mathsf{b})$ denotes Brownian motion on X starting in γ_s. In the very last step, we used the assumed boundedness of k together with the dominated convergence theorem. $\square$

Remark 5.12. The previous corollary applied to $\mu := \delta_x$ and $\nu := \delta_y$ for $x, y \in X$ at $t = 0$, choosing $\boldsymbol{\pi}_0$ as the Dirac mass on an arbitrary geodesic $\gamma \in \mathrm{G}_0(x,y)$, yields the estimate

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0} W_p^p(\mathsf{H}_t\delta_x, \mathsf{H}_t\delta_y) \leq -p \sup_{\gamma\in\mathrm{G}_0(x,y)} \int_0^1 k(\gamma_s)\,\mathrm{d}s\,\mathsf{d}^p(x,y) \leq -p\,\overline{k}(x,y)\,\mathsf{d}^p(x,y),$$

where, as in (1.3), the function $\overline{k}\colon X\times X \to \mathbb{R}$ is defined by

$$\overline{k}(x,y) := \liminf_{(x_n,y_n)\to(x,y)} \sup_{\gamma\in \mathrm{G}_0(x_n,y_n)} \int_0^1 k(\gamma_s)\,\mathrm{d}s.$$

Note that $\overline{k}$ is lower semicontinuous and bounded from below.

This improves the differential p-transport estimate (5.4), since $\underline{k} \leq \overline{k}$ on $X\times X$, see also Proposition 4.6. In Chapter 6, we shall construct a coupling of Brownian motions obeying pathwise bounds involving the larger function $\overline{k}$ in place of $\underline{k}$. In particular, using Theorem 5.17, all equivalences from Theorem 1.1 and Theorem 1.8 are still valid when replacing the function $\underline{k}$ by $\overline{k}$ in all relevant quantities. ■

The proof of the $\mathrm{TE}_p(k)$ property starting from $\mathrm{GE}_q(k)$ with dual $p,q\in(1,\infty)$ is slightly more involved as a control of the error terms is only possible "locally" for small times. A crucial ingredient is the subsequent result.

Lemma 5.13. *Let u and v be bounded Borel functions on X such that $u\leq v$ on a ball $B_\delta(z)$, $z\in X$ and $\delta>0$. Then for every $p\in(1,\infty)$ and $\varepsilon>0$, there exists $t_*>0$ such that for every $t\in[0,t_*]$, every nonnegative Borel function g on X, and every Brownian motion $(\mathbb{P}_x,\mathsf{b})$ on X starting in $x\in B_{\delta/2}(z)$, we have*

$$\mathbb{E}_x\Big[\mathrm{e}^{\int_0^t u(\mathsf{b}_r)\,\mathrm{d}r} g(\mathsf{b}_t)\Big] \leq \mathbb{E}_x\Big[\mathrm{e}^{p\int_0^t (v(\mathsf{b}_r)+\varepsilon)\,\mathrm{d}r} g^p(\mathsf{b}_t)\Big]^{1/p}.$$

Proof. The condition on u and v guarantees that for fixed $T>0$ and every $t\in[0,T]$,

$$\mathrm{e}^{\int_0^t u(\mathsf{b}_r)\,\mathrm{d}r} - \mathrm{e}^{\int_0^t v(\mathsf{b}_r)\,\mathrm{d}r} = \int_0^t \mathrm{e}^{\int_0^s u(\mathsf{b}_r)\,\mathrm{d}r + \int_s^t v(\mathsf{b}_r)\,\mathrm{d}r}(u-v)(\mathsf{b}_s)\,\mathrm{d}s \leq M\int_0^t \mathbb{1}_{\{\mathsf{b}_s\notin B_\delta(z)\}}\,\mathrm{d}s.$$

Here, $M>0$ is a constant depending only on u, v and T. Therefore,

$$\begin{aligned}\mathbb{E}_x\Big[\mathrm{e}^{\int_0^t u(\mathsf{b}_r)\,\mathrm{d}r} g(\mathsf{b}_t)\Big] &\leq \mathbb{E}_x\Big[\mathrm{e}^{\int_0^t v(\mathsf{b}_r)\,\mathrm{d}r} g(\mathsf{b}_t)\Big] + M\int_0^t \mathbb{E}_x\Big[\mathrm{e}^{\int_0^t v(\mathsf{b}_r)\,\mathrm{d}r} g(\mathsf{b}_t)\mathbb{1}_{\{\mathsf{b}_s\notin B_\delta(z)\}}\Big]\,\mathrm{d}s\\ &\leq \mathbb{E}_x\Big[\mathrm{e}^{\int_0^t p v(\mathsf{b}_r)\,\mathrm{d}r} g^p(\mathsf{b}_t)\Big]^{1/p}\Big(1+M\int_0^t \mathbb{P}_x\big[\mathsf{b}_s\notin B_\delta(z)\big]^{1/q}\,\mathrm{d}s\Big),\end{aligned}$$

where $q\in(1,\infty)$ denotes the dual exponent to p. By Lemma 4.2, we know that $\mathbb{P}_x[\mathsf{b}_s\notin B_\delta(z)]\leq s^q$ for every $s\in[0,t]$ and small enough t. Thus, $1+M\int_0^t \mathbb{P}_x[\mathsf{b}_s\notin B_\delta(z)]^{1/q}\,\mathrm{d}s \leq \mathrm{e}^{\varepsilon t}$, which directly proves the claim. □

Remark 5.14. With the very same strategy, also estimates for Feynman–Kac-type expressions in terms of pairs of Brownian motions can be derived, each component being required to start within $B_{\delta/2}(z)$. Moreover, the integrands u and v are then supposed to be functions on $X\times X$ with $u\leq v$ on $B_\delta(z)\times B_\delta(z)$. ■

Proposition 5.15. *Let $p,q\in(1,\infty)$ such that $1/p+1/q=1$ and assume the q-gradient estimate $\mathrm{GE}_q(k)$. Assume that $\underline{\ell}\in \mathrm{C_b}(X\times X)$ with $\underline{\ell}\leq\underline{k}$ on $X\times X$, and put $\ell(x):=\underline{\ell}(x,x)$ for $x\in X$. Then for every $\varepsilon>0$, $p'\in(1,p)$ and $z\in X$, there exist $\delta>0$ and $t_*>0$ such that for every $x,y\in B_\delta(z)$, every $\gamma\in \mathrm{G}_0(y,x)$ and every $t\in[0,t^*]$, we have*

$$W_{p'}^{p'}(\mathsf{H}_t\delta_x,\mathsf{H}_t\delta_y) \leq \mathsf{d}(x,y)\,\mathrm{e}^{-\left(\int_0^1 \ell(\gamma_r)\,\mathrm{d}r-\varepsilon\right)t},$$

and thus in particular,

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0} W_{p'}(\mathsf{H}_t\delta_x, \mathsf{H}_t\delta_y) \leq -\,\mathsf{d}(x,y)\,\Big(\int_0^1 \ell(\gamma_r)\,\mathrm{d}r - \varepsilon\Big).$$

Proof. We adapt the proof of Theorem 5.10 by adding a localization argument. Given $z \in X$ and $\varepsilon > 0$, choose $\delta > 0$ and $L_z \in \mathbb{R}$ such that $L_z \leq \ell \leq L_z + \varepsilon/2$ on $B_{3\delta}(z)$. Let $x, y \in B_\delta(z)$ and $\gamma \in \mathrm{G}_0(y,x)$, and note that $L_z \leq \int_0^1 \ell(\gamma_r)\,\mathrm{d}r \leq L_z + \varepsilon/2$.

Denote by Q_s the p'-Hopf–Lax semigroup with dual exponent $q' \in (q, \infty)$. Since $|\mathrm{D}\mathsf{P}_t Q_s f|$ is a weak upper gradient and using $\mathrm{GE}_q(k)$, which clearly implies $\mathrm{GE}_q(\ell)$, we directly obtain, for $\mathscr{L}^1$-a.e. $s \in [0,1]$,

$$\limsup_{h\downarrow 0} \frac{1}{h}\big(\mathsf{P}_t Q_s f(\gamma_{s+h}) - \mathsf{P}_t Q_s f(\gamma_s)\big) \leq \mathsf{d}(x,y)\,\big(\mathsf{P}_t^{q\ell}|\mathrm{D}Q_s f|^q\big)^{1/q}(\gamma_s).$$

Applying Lemma 5.13 with $\varepsilon/2$ and $t/2$ in place of ε and t, respectively, we get, for small enough t,

$$\big(\mathsf{P}_t^{q\ell}|\mathrm{D}Q_s f|^q\big)^{1/q}(\gamma_s) \leq \mathrm{e}^{-(L_z-\varepsilon/2)t}\,\mathsf{P}_t\big(\mathrm{lip}(Q_s f)^{q'}\big)^{1/q'}(\gamma_s),$$

and thus

$$\mathsf{d}(x,y)\,\big(\mathsf{P}_t^{q\ell}|\mathrm{D}Q_s f|\big)^{1/q}(\gamma_s) \leq \frac{\mathsf{d}^{p'}(x,y)}{p'}\,\mathrm{e}^{-p'(L_z-\varepsilon/2)t} + \frac{1}{q'}\mathsf{P}_t\big(\mathrm{lip}(Q_s f)^{q'}\big)(\gamma_s)$$

for $\mathscr{L}^1$-a.e. $s \in [0,1]$ by Young's inequality. Therefore, Lemma 5.9 with q' in place of q yields

$$\mathsf{P}_t Q_1 f(x) - \mathsf{P}_t f(y) \leq \frac{\mathsf{d}^{p'}(x,y)}{p'}\,\mathrm{e}^{-p'(L_z-\varepsilon/2)t} \leq \frac{\mathsf{d}^{p'}(x,y)}{p'}\,\mathrm{e}^{-p'(\int_0^1 \ell(\gamma_r)\,\mathrm{d}r - \varepsilon)t}.$$

Taking the supremum over $f \in \mathrm{Lip}_\mathrm{b}(X)$, we conclude by (2.3). □

Theorem 5.16. *Given $p, q \in (1,\infty)$ with $1/p+1/q = 1$, the q-gradient estimate $\mathrm{GE}_q(k)$ implies the p-transport estimate $\mathrm{TE}_p(k)$.*

Proof. Fix $x, y \in X$, an arbitrary geodesic $\gamma \in \mathrm{G}_0(y,x)$ and ℓ as in Proposition 5.15. Given $\varepsilon > 0$, choose a finite covering of $\gamma([0,1])$ by metric balls $B_{\delta_i/2}(\gamma_{s_i})$, $i \in \{1,\dots,n\}$ and $n \in \mathbb{N}$, such that each of the enlarged balls $B_{\delta_i}(\gamma_{s_i})$ satisfies the assumption of the previous Proposition 5.15. Without restriction, we may assume $s_1 = 0$ and $s_n = 1$. Applying this proposition to pairs of intermediate points $\gamma_{s_{i-1}}$ and γ_{s_i} and the reparameterized geodesics $\gamma^i \in \mathrm{G}_0(\gamma_{s_{i-1}}, \gamma_{s_i})$ defined by $\gamma^i_r := \gamma_{s_{i-1}+r(s_i-s_{i-1})}$, $r \in [0,1]$, yields

$$\begin{aligned}
\frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0} W_{p'}(\mathsf{H}_t\delta_x, \mathsf{H}_t\delta_y) &\leq \sum_{i=1}^n \frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0} W_{p'}\big(\mathsf{H}_t\delta_{\gamma_{s_{i-1}}}, \mathsf{H}_t\delta_{\gamma_{s_i}}\big)\\
&\leq -\sum_{i=1}^n \mathsf{d}(\gamma_{s_{i-1}}, \gamma_{s_i})\,\Big(\int_0^1 \ell\big(\gamma^i_r\big)\,\mathrm{d}r - \varepsilon\Big)\\
&= -\,\mathsf{d}(x,y)\,\Big(\int_0^1 \ell(\gamma_r)\,\mathrm{d}r - \varepsilon\Big).
\end{aligned}$$

Since ℓ is arbitrary, this bound holds with k in place of ℓ by Lemma 2.1. Furthermore, by definition of $\underline{k}$ and the arbitrariness of $\varepsilon > 0$, we deduce the differential transport estimate (5.4) with p replaced by p'. Since this true for every $p' \in (1,p)$, this finally yields $\mathrm{TE}_p(k)$ by Proposition 5.6 and monotone convergence. □

5.4. Gradient estimates out of pathwise and transport estimates

A modification of the arguments given in [22, Proposition 3.1] allows us to prove the converse direction of Theorem 1.8, i.e. that the p-transport estimate $\mathrm{TE}_p(k)$ implies the q-gradient estimate $\mathrm{GE}_q(k)$, where $1/p + 1/q = 1$. As in the previous section, a control of the error terms can only be achieved for small times. Therefore, instead of deriving $\mathrm{GE}_q(k)$ directly, it is more convenient to establish a local version of the q-Bochner inequality $\mathrm{BE}_q(k,\infty)$.

As in the preceding Section 5.3, the extremal version $q = 1$ is much easier to treat: in this case, the condition "$\mathrm{TE}_\infty(k)$" is to be interpreted as "$\mathrm{TE}_p(k)$ holds for any $p \in [1,\infty)$", which translates into the requirement of $\mathrm{PCP}(k)$ as discussed in Chapter 6.

Theorem 5.17. *The property* $\mathrm{PCP}(k)$ *implies the 1-gradient estimate* $\mathrm{GE}_1(k)$*, that is, for every* $f \in W^{1,2}(X)$ *and* $t \geq 0$*, we have*

$$\Gamma(\mathsf{P}_t f)^{1/2} \leq \mathsf{P}_t^k\big(\Gamma(f)^{1/2}\big) \quad \mathfrak{m}\text{-}a.e.$$

Proof. Fix $f \in \mathrm{Lip}_{\mathrm{bs}}(X)$ and $x \in X$. Recall that $\mathsf{P}_{t/2} f(x) = \mathbb{E}_x[f(\mathsf{b}_t)]$, where $(\mathbb{P}_x, \mathsf{b})$ denotes Brownian motion on X starting in x. Pick a function $\underline{\ell} \in \mathrm{Lip}_{\mathrm{b}}(X \times X)$ with $\underline{\ell} \leq \underline{k}$ on $X \times X$, and set $\ell(x) := \underline{\ell}(x,x)$ for $x \in X$. By $\mathrm{PCP}(k)$, given any $\varrho > 0$ and $y \in B_\varrho(x)$, we may choose a pair $\big(\mathbb{P}_{x,y}, \mathsf{b}^1\big)$ and $\big(\mathbb{P}_{x,y}, \mathsf{b}^2\big)$ of coupled Brownian motions in such a way that $\mathbb{P}_{x,y}$-a.s., we have

$$\mathsf{d}\big(\mathsf{b}_t^1, \mathsf{b}_t^2\big) \leq \mathrm{e}^{-\int_0^t \underline{k}(\mathsf{b}_r^1,\mathsf{b}_r^2)/2\,\mathrm{d}r}\,\mathsf{d}(x,y) \leq \mathrm{e}^{-\int_0^t \underline{\ell}(\mathsf{b}_r^1,\mathsf{b}_r^2)/2\,\mathrm{d}r}\,\mathsf{d}(x,y) \tag{5.9}$$

for every $t \geq 0$. With this in hand, we can estimate

$$\begin{aligned}
|\mathrm{D}\mathsf{P}_{t/2} f|(x) &\leq \lim_{\varrho\downarrow 0} \sup_{y \in B_\varrho(x)} \frac{|\mathsf{P}_{t/2} f(x) - \mathsf{P}_{t/2} f(y)|}{\mathsf{d}(x,y)} \\
&\leq \lim_{\varrho\downarrow 0} \sup_{y \in B_\varrho(x)} \mathbb{E}_{x,y}\Big[\frac{|f(\mathsf{b}_t^1) - f(\mathsf{b}_t^2)|}{\mathsf{d}(\mathsf{b}_t^1,\mathsf{b}_t^2)}\,\frac{\mathsf{d}(\mathsf{b}_t^1,\mathsf{b}_t^2)}{\mathsf{d}(x,y)}\,\big(\mathbb{1}_{U_{\varrho,t}} + \mathbb{1}_{V_{\varrho,t}} + \mathbb{1}_{W_{\varrho,t}}\big)\Big],
\end{aligned}$$

where $V_{\varrho,t} := \big\{\mathsf{d}\big(\mathsf{b}_t^1, \mathsf{b}_t^2\big) \geq \varrho^{1/2}\big\}$, $W_{\varrho,t} := \big\{\int_0^t \mathsf{d}\big(\mathsf{b}_r^1, \mathsf{b}_r^2\big)\,\mathrm{d}r/t \geq \varrho^{1/2}\big\}$ and $U_{\varrho,t} := V_{\varrho,t}^{\mathrm{c}} \cap W_{\varrho,t}^{\mathrm{c}}$.

Let us consider this upper bound for the weak upper gradient $|\mathrm{D}\mathsf{P}_{t/2} f|(x)$ term by term, starting with the contribution coming from $U_{\varrho,t}$. We have the inequality $\int_0^t \underline{\ell}\big(\mathsf{b}_r^1, \mathsf{b}_r^2\big)\,\mathrm{d}r \geq \int_0^t \ell\big(\mathsf{b}_r^1\big)\,\mathrm{d}r - \mathrm{Lip}\big(\underline{\ell}\big) t\varrho^{1/2}$ on $W_{\varrho,t}^{\mathrm{c}}$, which gives

$$\begin{aligned}
&\lim_{\varrho\downarrow 0} \sup_{y \in B_\varrho(x)} \mathbb{E}_{x,y}\Big[\frac{|f(\mathsf{b}_t^1) - f(\mathsf{b}_t^2)|}{\mathsf{d}(\mathsf{b}_t^1,\mathsf{b}_t^2)}\,\frac{\mathsf{d}(\mathsf{b}_t^1,\mathsf{b}_t^2)}{\mathsf{d}(x,y)}\,\mathbb{1}_{U_{\varrho,t}}\Big] \\
&\qquad \leq \lim_{\varrho\downarrow 0} \sup_{y \in B_\varrho(x)} \mathbb{E}_{x,y}\Big[\mathrm{e}^{-\int_0^t \ell(\mathsf{b}_r^1)/2\,\mathrm{d}r + \mathrm{Lip}(\underline{\ell}) t\varrho^{1/2}/2} \sup_{z \in B_{\varrho^{1/2}}(\mathsf{b}_t^1)} \Big|\frac{f(\mathsf{b}_t^1) - f(z)}{\mathsf{d}(\mathsf{b}_t^1, z)}\Big|\Big] \\
&\qquad = \lim_{\varrho\downarrow 0} \mathbb{E}_x\Big[\mathrm{e}^{-\int_0^t \ell(\mathsf{b}_r^x)/2\,\mathrm{d}r + \mathrm{Lip}(\underline{\ell}) t\varrho^{1/2}/2} \sup_{z \in B_{\varrho^{1/2}}(\mathsf{b}_t^x)} \Big|\frac{f(\mathsf{b}_t^x) - f(z)}{\mathsf{d}(\mathsf{b}_t^x, z)}\Big|\Big]
\end{aligned}$$

$$= \mathbb{E}_x\Big[\mathrm{e}^{-\int_0^t \ell(\mathsf{b}_r^x)/2\,\mathrm{d}r}\,|\mathrm{D}f|(\mathsf{b}_t^x)\Big] = \mathsf{P}_{t/2}^{\ell}\big(\Gamma(f)^{1/2}\big)(x).$$

We point out the intermediate change from the process b^1, which in general also depends on y, to a Brownian motion $\big(\mathbb{P}_x, \mathsf{b}^x\big)$ on X starting in x, chosen independently of y.

Next we consider the term involving $\mathbb{1}_{V_{\varrho,t}}$. Denoting by $C > 0$ a suitable upper bound on $\underline{\ell}$, we obtain by (5.9) that

$$\lim_{\varrho\downarrow 0} \sup_{y\in B_\varrho(x)} \mathbb{E}_{x,y}\Big[\frac{|f(\mathsf{b}_t^1) - f(\mathsf{b}_t^2)|}{\mathsf{d}(\mathsf{b}_t^1,\mathsf{b}_t^2)}\,\frac{\mathsf{d}(\mathsf{b}_t^1,\mathsf{b}_t^2)}{\mathsf{d}(x,y)}\,\mathbb{1}_{V_{\varrho,t}}\Big]$$
$$\leq \mathrm{Lip}(f)\,\lim_{\varrho\downarrow 0}\frac{1}{\varrho^{1/2}}\sup_{y\in B_\varrho(x)}\mathbb{E}_{x,y}\Big[\frac{\mathsf{d}^2(\mathsf{b}_t^1,\mathsf{b}_t^2)}{\mathsf{d}(x,y)}\Big] \leq \mathrm{Lip}(f)\,\mathrm{e}^{Ct}\lim_{\varrho\downarrow 0}\frac{1}{\varrho^{1/2}}\sup_{y\in B_\varrho(x)}\mathsf{d}(x,y) = 0.$$

Similarly, the last expression which involves $W_{\varrho,t}$ can be bounded through

$$\lim_{\varrho\downarrow 0} \sup_{y\in B_\varrho(x)} \mathbb{E}_{x,y}\Big[\frac{|f(\mathsf{b}_t^1) - f(\mathsf{b}_t^2)|}{\mathsf{d}(\mathsf{b}_t^1,\mathsf{b}_t^2)}\,\frac{\mathsf{d}(\mathsf{b}_t^1,\mathsf{b}_t^2)}{\mathsf{d}(x,y)}\,\mathbb{1}_{W_{\varrho,t}}\Big]$$
$$\leq \mathrm{Lip}(f)\,\lim_{\varrho\downarrow 0}\frac{1}{t\varrho^{1/2}}\sup_{y\in B_\varrho(x)}\int_0^t \mathbb{E}_{x,y}\Big[\frac{\mathsf{d}(\mathsf{b}_t^1,\mathsf{b}_t^2)\,\mathsf{d}(\mathsf{b}_r^1,\mathsf{b}_r^2)}{\mathsf{d}(x,y)}\Big]\,\mathrm{d}r$$
$$\leq \mathrm{Lip}(f)\,\mathrm{e}^{Ct}\lim_{\varrho\downarrow 0}\frac{1}{\varrho^{1/2}}\sup_{y\in B_\varrho(x)}\mathsf{d}(x,y) = 0.$$

Finally, we have to extend the class of admissible functions f and pass to $\mathrm{GE}_1(k)$. By uniform convexity of $\mathscr{E}$, every $f \in W^{1,2}(X)$ can be approximated strongly in $W^{1,2}(X)$ by a sequence of Lipschitz functions f_n with bounded support [3]. Thus, possibly passing to a subsequence, we get, for some suitable $c \in \mathbb{R}$, that

$$\lim_{n\to\infty}\mathsf{P}_t^{\ell}\big(\Gamma(f-f_n)^{1/2}\big) \leq \mathrm{e}^{ct}\lim_{n\to\infty}\mathsf{P}_t\big(\Gamma(f-f_n)^{1/2}\big) = 0 \quad \mathfrak{m}\text{-a.e.}$$

Moreover, $\Gamma(\mathsf{P}_t f_n) \to \Gamma(\mathsf{P}_t f)$ in $L^1(X,\mathfrak{m})$ as $n\to\infty$ and thus, up to a subsequence, this convergence holds $\mathfrak{m}$-a.e., which then proves $\mathrm{GE}_1(\ell)$ for arbitrary $f \in W^{1,2}(X)$. By the arbitrariness of $\underline{\ell}$, Lemma 2.1 and the identity $k(x) = \underline{k}(x,x)$ for every $x \in X$, we deduce $\mathrm{GE}_1(k)$ by the monotone convergence theorem. □

Proposition 5.18. *Let $\varepsilon > 0$, $z \in X$ and $q \in (1,\infty)$. Assume the transport estimate $\mathrm{TE}_p(k)$, where $1/p+1/q = 1$. Suppose that $\underline{\ell} \in \mathrm{C_b}(X\times X)$ with $\underline{\ell} \leq \underline{k}$ on $X \times X$. Then for every $q' \in (q,\infty)$, there exist $t_* > 0$ and $\delta > 0$ such that*

$$\Gamma(\mathsf{P}_t f)^{q'/2} \leq \mathsf{P}_t^{q'(\ell-\varepsilon)}\big(\Gamma(f)^{q'/2}\big) \quad \mathfrak{m}\text{-}a.e.\ on\ B_\delta(z)$$

for every $t \in [0,t_*]$ and all bounded Lipschitz functions f on X.

Proof. Fix $T > 0$. Given $\varepsilon > 0$, choose $\delta > 0$ and $L_z \in \mathbb{R}$ such that $L_z \leq \underline{\ell}(x,y) \leq L_z + \varepsilon/3$ for every $x,y \in B_{3\delta}(z)$. Given $t \in [0,T]$, $x \in B_\delta(z)$ and $y \in B_\varrho(z)$ with $\varrho \leq \delta$, select a pair $\big(\mathbb{P}_{x,y}, \mathsf{b}^1\big)$ and $\big(\mathbb{P}_{x,y}, \mathsf{b}^2\big)$ of coupled Brownian motions starting in (x,y) which attains the minimum in the definition of $W_p^{\underline{k}}(\delta_x,\delta_y,t/2) \leq \mathsf{d}(x,y)$. The choice of this pair does depend on x, y and t, but these dependencies are suppressed in the notation. Similarly to the proof of Theorem 5.17, for every $f \in \mathrm{Lip_b}(X)$, we have

$$|\mathrm{D}\mathsf{P}_{t/2}f|(x) \leq \lim_{\varrho\downarrow 0}\sup_{y\in B_\varrho(x)}\mathbb{E}_{x,y}\Big[\frac{|f(\mathsf{b}_t^1) - f(\mathsf{b}_t^2)|}{\mathsf{d}(\mathsf{b}_t^1,\mathsf{b}_t^2)}\,\frac{\mathsf{d}(\mathsf{b}_t^1,\mathsf{b}_t^2)}{\mathsf{d}(x,y)}\,\big(\mathbb{1}_{V_{\varrho,t}} + \mathbb{1}_{V_{\varrho,t}^c}\big)\Big]$$

where $V_{\varrho,t} := \{\mathsf{d}(\mathsf{b}_t^1, \mathsf{b}_t^2) \geq \varrho^{1/2q}\}$. The contribution of $V_{\varrho,t}$ vanishes as $\varrho \downarrow 0$ due to

$$\begin{aligned}
&\lim_{\varrho\downarrow 0} \sup_{y\in B_\varrho(x)} \mathbb{E}_{x,y}\Big[\frac{|f(\mathsf{b}_t^1) - f(\mathsf{b}_t^2)|}{\mathsf{d}(x,y)}\,\mathbb{1}_{V_{\varrho,t}}\Big] \\
&\qquad \leq \mathrm{Lip}(f)\,\mathrm{e}^{Ct} \lim_{\varrho\downarrow 0} \varrho^{(1-p)/2q} \sup_{y\in B_\varrho(x)} \frac{1}{\mathsf{d}(x,y)}\,\mathbb{E}_{x,y}\Big[\mathrm{e}^{\int_0^t p\underline{k}(\mathsf{b}_r^1,\mathsf{b}_r^2)/2\,\mathrm{d}r}\,\mathsf{d}^p\big(\mathsf{b}_t^1,\mathsf{b}_t^2\big)\Big] \\
&\qquad \leq \mathrm{Lip}(f)\,\mathrm{e}^{Ct} \lim_{\varrho\downarrow 0} \varrho^{(1-p)/2q} \sup_{y\in B_\varrho(x)} \mathsf{d}^{p-1}(x,y) = 0
\end{aligned}$$

for a suitable $C > 0$, where we used the assumption that $\underline{\ell} \leq \underline{k}$ in the first inequality and the $\mathrm{TE}_p(k)$ condition in the last inequality.

Next we study the influence coming from $V_{\varrho,t}^{\mathrm{c}}$. Choosing some exponents $q'' \in (q, q')$ and $p'' \in (1, p')$ dual to each other, using Hölder's inequality, Lemma 5.13 with $\varepsilon/3$ and $t/2$ in place of ε and t, respectively, and eventually assumption $\mathrm{TE}_p(k)$, we obtain for sufficiently small t that

$$\begin{aligned}
&\lim_{\varrho\downarrow 0} \sup_{y\in B_\varrho(x)} \mathbb{E}_{x,y}\Big[\frac{|f(\mathsf{b}_t^1) - f(\mathsf{b}_t^2)|}{\mathsf{d}(\mathsf{b}_t^1,\mathsf{b}_t^2)}\,\frac{\mathsf{d}(\mathsf{b}_t^1,\mathsf{b}_t^2)}{\mathsf{d}(x,y)}\,\mathbb{1}_{V_{\varrho,t}^{\mathrm{c}}}\Big] \\
&\qquad \leq \mathrm{e}^{-(L_z-\varepsilon/3)t/2} \lim_{\varrho\downarrow 0} \sup_{y\in B_\varrho(x)} \mathbb{E}_{x,y}\Big[\Big|\frac{f(\mathsf{b}_t^1) - f(\mathsf{b}_t^2)}{\mathsf{d}(\mathsf{b}_t^1,\mathsf{b}_t^2)}\Big|^{q''}\,\mathbb{1}_{V_{\varrho,t}^{\mathrm{c}}}\Big]^{1/q''} \\
&\qquad\qquad \cdot \lim_{\varrho\downarrow 0} \sup_{y\in B_\varrho(x)} \mathbb{E}_{x,y}\Big[\mathrm{e}^{p''(L_z-\varepsilon/3)t/2}\,\Big|\frac{\mathsf{d}(\mathsf{b}_t^1,\mathsf{b}_t^2)}{\mathsf{d}(x,y)}\Big|^{p''}\Big]^{1/p''} \\
&\qquad \leq \mathrm{e}^{-(L_z-\varepsilon/3)t/2} \lim_{\varrho\downarrow 0} \mathbb{E}_x\Big[\sup_{z\in B_{\varrho^{1/2q}}(\mathsf{b}_t^x)} \Big|\frac{f(\mathsf{b}_t^x) - f(z)}{\mathsf{d}(\mathsf{b}_t^x, z)}\Big|^{q''}\Big]^{1/q''}\,\frac{1}{\mathsf{d}(x,y)}\,W_p^{\underline{k}}(\delta_x,\delta_y,t) \\
&\qquad \leq \mathrm{e}^{-(L_z-\varepsilon/3)t/2}\,\mathbb{E}_x\big[|\mathrm{D}f|^{q''}(\mathsf{b}_t^x)\big]^{1/q''}.
\end{aligned}$$

Here $(\mathbb{P}_x, \mathsf{b}^x)$ is a Brownian motion on X starting in x which is chosen independently of y. Once again using Lemma 5.13 as above to estimate the last expression, we finally obtain

$$\lim_{\varrho\downarrow 0} \sup_{y\in B_\varrho(x)} \mathbb{E}_{x,y}\Big[\frac{|f(\mathsf{b}_t^1) - f(\mathsf{b}_t^2)|}{\mathsf{d}(\mathsf{b}_t^1,\mathsf{b}_t^2)}\,\frac{\mathsf{d}(\mathsf{b}_t^1,\mathsf{b}_t^2)}{\mathsf{d}(x,y)}\,\mathbb{1}_{V_{\varrho,t}^{\mathrm{c}}}\Big] \leq \mathsf{P}_t^{q'(\ell-\varepsilon)}\big(|\mathrm{D}f|^{q'}\big)^{1/q'}(x). \quad \square$$

Theorem 5.19. *Given $p, q \in (1,\infty)$ with $1/p + 1/q = 1$, the p-transport estimate $\mathrm{TE}_p(k)$ implies the q-gradient estimate $\mathrm{GE}_q(k)$.*

Proof. Let $\underline{\ell}$ be as in Proposition 5.18 and put $\ell(x) := \underline{\ell}(x,x)$ for $x \in X$. First, we assume that $q \in [2,\infty)$. Given $\varepsilon > 0$, $z \in X$, $t_* > 0$, $q' \in (q,\infty)$ and the associated time $t_* > 0$ from in Proposition 5.18, arguing as in the proof of Theorem 3.4, the function $F\colon [0,t_*] \to \mathbb{R}$ defined by

$$F(t) := \int_X \Big(\mathsf{P}_t^{q'(\ell-\varepsilon)}\big(\Gamma(f)^{q'/2}\big) - \Gamma(\mathsf{P}_t f)^{q'/2}\Big)\,\phi\,\mathrm{d}\mathfrak{m}$$

belongs to $\mathrm{C}^1([0,t_*])$ for every $f \in \mathrm{TestF}(X)$ and all nonnegative functions $\phi \in W^{1,2}(X) \cap L^\infty(X,\mathfrak{m})$ supported in $B_\delta(z)$. The function F itself and its derivative at 0 are nonnegative by Proposition 5.18. The latter translates into

$$-\int_X \Big(\frac{1}{q'}\Gamma\big(\Gamma(f)^{q'/2},\phi\big) + \Gamma(f)^{q'/2}\,\Gamma(f,\Delta f)\,\phi\Big)\,\mathrm{d}\mathfrak{m} \geq \int_X (\ell-\varepsilon)\,\Gamma(f)^{q'/2}\,\phi\,\mathrm{d}\mathfrak{m}.$$

Approximating k from below by the sequence $k_n \in \mathrm{Lip_b}(X)$ of functions $k_n(x) := \underline{k}_n(x,x)$ for $x \in X$, or in other words, replacing $\underline{\ell}$ by $\underline{k}_n$ for every $n \in \mathbb{N}$, where $\underline{k}_n$ tends to $\underline{k}$ from below as provided by Lemma 2.1, and letting $q' \downarrow q$ and $\varepsilon \downarrow 0$, we obtain precisely the local q-Bakry–Émery inequality $\mathrm{BE}_{q,\mathrm{loc}}(k,\infty)$ according to Definition 3.9. Since the latter implies $\mathrm{BE}_q(k,\infty)$ by Theorem 3.10, the equivalence with $\mathrm{GE}_q(k)$ finishes the proof in the case $q \in [2,\infty)$.

If $q \in [1,2)$, choosing $q' := 2$ in Proposition 5.18 and arguing as above, we obtain $\mathrm{BE}_2(k,\infty)$, which in turn implies $\mathrm{BE}_q(k,\infty)$. □

6. A pathwise coupling estimate

It remains to treat the pathwise coupling property w.r.t. k to finish the proof of Theorem 1.1. By Theorem 5.17, we know that $\mathrm{PCP}(k)$ implies $\mathrm{GE}_1(k)$. Conversely, letting $\overline{k}$ be the function from Remark 5.12 which is even larger than $\underline{k}$, $\mathrm{GE}_1(k)$ implies that, for any $p \in (1,\infty)$,

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=0} W_p^p(\mathsf{H}_t\delta_x, \mathsf{H}_t\delta_y) \leq -p\overline{k}(x,y)\,\mathsf{d}^p(x,y) \quad \text{for every } x,y \in X.$$

The same argument as for Proposition 5.6 then shows that $t \mapsto W_p^{\overline{k}}(\delta_x,\delta_y,t)$ is nonincreasing for every $p \in (1,\infty)$ and every $x,y \in X$. Therefore, $\mathrm{PCP}(k)$ follows once having proven the subsequent even stronger statement which has a weaker assumption.

Theorem 6.1. *Suppose that, for all large enough $p \in (1,\infty)$, the map $t \mapsto W_p^{\overline{k}}(\delta_x,\delta_y,t)$ is nonincreasing on $[0,\infty)$ for every $x,y \in X$. Then for every $\mu_1,\mu_2 \in \mathscr{P}(X)$ there exists a pair $\big(\mathbb{P},\mathsf{b}^1\big)$ and $\big(\mathbb{P},\mathsf{b}^2\big)$ of coupled Brownian motions on X with initial distributions μ_1 and μ_2, respectively, such that $\mathbb{P}$-a.s., we have*

$$\mathsf{d}\big(\mathsf{b}_t^1,\mathsf{b}_t^2\big) \leq \mathrm{e}^{-\int_s^t \overline{k}(\mathsf{b}_r^1,\mathsf{b}_r^2)/2\,\mathrm{d}r}\,\mathsf{d}\big(\mathsf{b}_s^1,\mathsf{b}_s^2\big) \quad \text{for every } s,t \in [0,\infty) \text{ with } s \leq t.$$

In particular, the pathwise coupling property $\mathrm{PCP}(k)$ *holds.*

For the proof of Theorem 6.1, it is necessary to adapt the arguments from [31, Section 2] in a nontrivial way, since our pathwise estimate requires control of the entire path of $\big(\mathsf{b}^1,\mathsf{b}^2\big)$ on the interval $[s,t]$ and not just at the endpoints.

The proof of Theorem 6.1 will be subdivided into multiple steps. Firstly, we construct a coupled process starting in $\delta_x \otimes \delta_y$, $x,y \in X$, satisfying the desired pathwise contraction estimate on the interval $[0,1]$. Secondly, a gluing procedure will let us extend the process to $[0,\infty)$. Finally, we use a coupling technique to allow for arbitrary initial distributions.

Proposition 6.2. *Under the same assumptions as in Theorem 6.1, for every $t \geq 0$, there exists a universally measurable map*

$$\boldsymbol{\mu}^t\colon \quad X \times X \quad \longrightarrow \quad \mathscr{P}(\Pi_t)$$

such that for every $x,y \in X$, the marginals of $\boldsymbol{\mu}^t_{x,y} := \boldsymbol{\mu}^t(x,y)$ are laws of Brownian motions, restricted to $[0,t]$, starting in x and y, respectively, and

$$\mathsf{d}\big(\gamma_t^1,\gamma_t^2\big) \leq \mathrm{e}^{-\int_0^t \overline{k}(\gamma_r^1,\gamma_r^2)/2\,\mathrm{d}r}\,\mathsf{d}(x,y) \quad \text{for } \boldsymbol{\mu}^t_{x,y}\text{-a.e. } \gamma \in \Pi_t.$$

Proof. Given $x,y \in X$ and an increasing sequence $(p_n)_{n\in\mathbb{N}}$ tending to ∞, denote by $\boldsymbol{\eta}^{t,n}_{x,y} \in \mathscr{P}(\Pi_t)$ the measure obtained by Lemma 5.3 for the exponent p_n, $\underline{\ell}$ replaced by $\overline{k}$, and time $t/2$ in place of t. As for

Lemma 5.2, we see that the sequence $(\boldsymbol{\eta}^{t,n}_{x,y})_{n\in\mathbb{N}}$ is tight. Hence it converges weakly to some $\boldsymbol{\eta}^t_{x,y} \in \mathscr{P}(\Pi_t)$ along a subsequence which we do not relabel.

Let $p \in (1,\infty)$ arbitrary, and fix $\overline{\ell} \in \mathrm{C_b}(X\times X)$ with $\overline{\ell} \leq \overline{k}$ on $X\times X$. Then by Hölder's inequality and the nonincreasingness of $t \mapsto W^{\overline{k}}_{p_n}(\delta_x,\delta_y,t)$ for large enough n, we obtain

$$\begin{aligned}
&\Big(\int_{\Pi_t} \mathrm{e}^{\int_0^t p\overline{\ell}(\gamma_r^1,\gamma_r^2)/2\,\mathrm{d}r}\,\mathsf{d}^p\big(\gamma_t^1,\gamma_t^2\big)\,\mathrm{d}\boldsymbol{\eta}^t_{x,y}(\gamma)\Big)^{1/p}\\
&\qquad\leq \liminf_{n\to\infty}\Big(\int_{\Pi_t} \mathrm{e}^{\int_0^t p\overline{\ell}(\gamma_r^1,\gamma_r^2)/2\,\mathrm{d}r}\,\mathsf{d}^p\big(\gamma_t^1,\gamma_t^2\big)\,\mathrm{d}\boldsymbol{\eta}^{t,n}_{x,y}(\gamma)\Big)^{1/p}\\
&\qquad\leq \limsup_{n\to\infty}\Big(\int_{\Pi_t} \mathrm{e}^{\int_0^t p_n\overline{k}(\gamma_r^1,\gamma_r^2)/2\,\mathrm{d}r}\,\mathsf{d}^{p_n}\big(\gamma_t^1,\gamma_t^2\big)\,\mathrm{d}\boldsymbol{\eta}^{t,n}_{x,y}(\gamma)\Big)^{1/p_n} \leq \mathsf{d}(x,y).
\end{aligned}$$

Sending $p\to\infty$ and then approximating $\overline{k}$ from below by means of Lemma 2.1 gives

$$\mathsf{d}\big(\gamma_t^1,\gamma_t^2\big) \leq \mathrm{e}^{-\int_0^t \overline{k}(\gamma_r^1\gamma_r^2)/2\,\mathrm{d}r}\,\mathsf{d}(x,y) \quad \text{for } \boldsymbol{\eta}^t_{x,y}\text{-a.e. } \gamma\in\Pi_t.$$

A measurable selection argument as in the proof of Lemma 5.3 establishes the claim. □

The next goal is to obtain a measure which obeys such pathwise bound at *every* initial and terminal time instance in, say, $[0,1]$. Indeed, this is the point where the main work has to be done.

Theorem 6.3. *Under the same assumptions as in Theorem 6.1, there exists a universally measurable map*

$$\boldsymbol{\mu}\colon \quad X\times X \quad \longrightarrow \quad \mathscr{P}(\Pi_1)$$

such that for every $x,y\in X$, we have that the marginals of $\boldsymbol{\mu}_{x,y} := \boldsymbol{\mu}(x,y)$ are laws of Brownian motions, restricted to $[0,1]$, starting in x and y, respectively, and that there exists a $\boldsymbol{\mu}_{x,y}$-negligible Borel set $E\subset\Pi_1$ such that

$$\mathsf{d}\big(\gamma_t^1,\gamma_t^2\big) \leq \mathrm{e}^{-\int_s^t \overline{k}(\gamma_r^1,\gamma_r^2)/2\,\mathrm{d}r}\,\mathsf{d}\big(\gamma_s^1,\gamma_s^2\big) \quad \textit{for every } s,t\in[0,1] \textit{ with } s\leq t$$

for all $\gamma\in\Pi_1\setminus E$.

Proof. The strategy relies on patching the laws obtained in the previous proposition together on small dyadic partitions of $[0,1]$. Denote by $\boldsymbol{\mu}^{2^{-n}}$ the map from Proposition 6.2 and define

$$\boldsymbol{\mu}_{n,x,y} := \underbrace{\boldsymbol{\mu}^{2^{-n}}\circ\cdots\circ\boldsymbol{\mu}^{2^{-n}}}_{2^{n-1}\text{ kernels}}\circ\,\boldsymbol{\mu}^{2^{-n}}_{x,y} \in \mathscr{P}(\Pi_1),$$

that is, at every dyadic partition point of $[0,1]$ at scale 2^{-n}, we attach a new random curve evolving according to the law obtained in Proposition 6.2 to the random endpoint of the previous curve. The marginals of $\boldsymbol{\mu}_{n,x,y}$ are the laws of Brownian motions on X, restricted to $[0,1]$, starting in x and y, respectively. As in the proof of Lemma 5.2, we may exhibit a subsequence, not relabeled in the sequel, weakly converging to some $\boldsymbol{\mu}_{x,y}\in\mathscr{P}(\Pi_1)$.

The key point lies in proving that for every $s,t\in\mathbb{Q}\cap[0,1]$ with $s\leq t$, there exists a $\boldsymbol{\mu}_{x,y}$-negligible Borel set $E_{s,t}\subset\Pi_1$ such that, for every $\gamma\in\Pi_1\setminus E_{s,t}$,

$$\mathsf{d}\big(\gamma_t^1,\gamma_t^2\big) \leq \mathrm{e}^{-\int_s^t \overline{k}(\gamma_r^1,\gamma_r^2)/2\,\mathrm{d}r}\,\mathsf{d}\big(\gamma_s^1,\gamma_s^2\big). \tag{6.1}$$

By continuity of curves, the desired requirements are then satisfied by the $\boldsymbol{\mu}_{x,y}$-null set

$$E := \bigcup_{\substack{s,t\in\mathbb{Q}\cap[0,1],\\ s\leq t}} E_{s,t}.$$

Let $\overline{\ell} \in \mathrm{C_b}(X\times X)$ as above, i.e. $\overline{\ell} \leq \overline{k}$ on $X\times X$. Pick s and t as above and notice that the sequences $s_m := 2^{-m}\lfloor 2^m s\rfloor$ and $t_m := 2^{-m}\lfloor 2^m t\rfloor$ tend to s and t, respectively, as $m\to\infty$. Fix $m\in\mathbb{N}$ and an arbitrary $n\geq m$. Given any $i\in\{1,\dots,2^n-1\}$, for every path $\widetilde{\gamma}\in\Pi_{2^{-n}}$ one gets

$$\mathsf{d}\big(\gamma^1_{2^{-n}},\gamma^2_{2^{-n}}\big) \leq \mathrm{e}^{-\int_0^{2^{-n}}\overline{\ell}(\gamma^1_r,\gamma^2_r)/2\,\mathrm{d}r}\,\mathsf{d}\big(\widetilde{\gamma}^1_{2^{-n}},\widetilde{\gamma}^2_{2^{-n}}\big)\quad \text{for } \boldsymbol{\mu}^{2^{-n}}_{\widetilde{\gamma}^1_{2^{-n}},\widetilde{\gamma}^2_{2^{-n}}}\text{-a.e. } \gamma\in\Pi_{2^{-n}}.$$

Observing that the dyadic partition of $[0,1]$ of step size 2^{-n} contains the one at scale 2^{-m} and then integrating the resulting $\boldsymbol{\mu}_{n,x,y}$-a.e. valid estimate, truncated at large enough $C>0$, against an arbitrary nonnegative function $\phi\in\mathrm{C_b}(\Pi_1)$, we obtain

$$\int_{\Pi_1}\phi(\gamma)\,\mathsf{d}_C\big(\gamma^1_{t_m},\gamma^2_{t_m}\big)\,\mathrm{d}\boldsymbol{\mu}_{n,x,y}(\gamma) \leq \int_{\Pi_1}\phi(\gamma)\,\mathrm{e}^{-\int_{2^{-n}\lfloor 2^n s_m\rfloor}^{2^{-n}\lfloor 2^n t_m\rfloor}\overline{\ell}(\gamma^1_r,\gamma^2_r)/2\,\mathrm{d}r}\,\mathsf{d}_C\big(\gamma^1_{s_m},\gamma^2_{s_m}\big)\,\mathrm{d}\boldsymbol{\mu}_{n,x,y}(\gamma),$$

where $\mathsf{d}_C := \min\{\mathsf{d},C\}$. Since $\overline{\ell}$ is bounded, for all $m\in\mathbb{N}$ and every $\varepsilon>0$, this yields

$$\begin{aligned}\int_{\Pi_1}\phi(\gamma)\,\mathsf{d}_C\big(\gamma^1_{t_m},\gamma^2_{t_m}\big)\,\mathrm{d}\boldsymbol{\mu}_{n,x,y}(\gamma) &\leq \int_{\Pi_1}\phi(\gamma)\,\mathrm{e}^{-\int_{s_m}^{t_m}\overline{\ell}(\gamma^1_r,\gamma^2_r)/2\,\mathrm{d}r}\,\mathsf{d}_C\big(\gamma^1_{s_m},\gamma^2_{s_m}\big)\,\mathrm{d}\boldsymbol{\mu}_{n,x,y}(\gamma)\\ &\quad + \varepsilon\int_{\Pi_1}\phi(\gamma)\,\mathsf{d}_C\big(\gamma^1_{s_m},\gamma^2_{s_m}\big)\,\mathrm{d}\boldsymbol{\mu}_{n,x,y}(\gamma)\end{aligned}$$

for all large enough n. Letting $n\to\infty$, $\varepsilon\downarrow 0$ and then $C\to\infty$ in the previous estimate as well as extending the class of ϕ to nonnegative, bounded Borel functions by a routine approximation argument, we get

$$\mathsf{d}\big(\gamma^1_{t_m},\gamma^2_{t_m}\big) \leq \mathrm{e}^{-\int_{s_m}^{t_m}\overline{\ell}(\gamma^1_r,\gamma^2_r)/2\,\mathrm{d}r}\,\mathsf{d}\big(\gamma^1_{s_m},\gamma^2_{s_m}\big)\quad\text{for } \boldsymbol{\mu}_{x,y}\text{-a.e. } \gamma\in\Pi_1. \tag{6.2}$$

Let us now put

$$\widetilde{E}_{s,t} := \bigcup_{m\in\mathbb{N}}\{\gamma\in\Pi_1 : \gamma \text{ does not satisfy (6.2)}\}$$

which clearly satisfies $\boldsymbol{\mu}_{x,y}\big[\widetilde{E}_{s,t}\big]=0$, and (6.1) holds on $\Pi_1\setminus\widetilde{E}_{s,t}$ with $\overline{\ell}$ in place of $\overline{k}$ by the convergences $s_m\to s$ and $t_m\to t$ as $m\to\infty$. Finally, denoting by $\overline{k}_n\in\mathrm{Lip_b}(X)$ a sequence approximating $\overline{k}$ from below as provided by Lemma 2.1, the above reasoning gives Borel subsets $\widetilde{E}^n_{s,t}$ of Π_1 such that $\boldsymbol{\mu}_{x,y}\big[\widetilde{E}^n_{s,t}\big]=0$ and

$$\mathsf{d}\big(\gamma^1_t,\gamma^2_t\big) \leq \mathrm{e}^{-\int_s^t\overline{k}_n(\gamma^1_r,\gamma^2_r)/2\,\mathrm{d}r}\,\mathsf{d}\big(\gamma^1_s,\gamma^2_s\big)$$

for every $\gamma\in\Pi_1\setminus\widetilde{E}^n_{s,t}$. Putting

$$E_{s,t} := \bigcup_{n=1}^{\infty}\widetilde{E}^n_{s,t},$$

we see that $\boldsymbol{\mu}_{x,y}\big[E_{s,t}\big]=0$ and that (6.1) holds for all $\gamma\in\Pi_1\setminus E_{s,t}$ by monotone convergence.

A similar argument and arguing as for Lemma 5.3 shows that we can then select the obtained measures in a universally measurable way. □

The cases of arbitrary initial distributions $\mu \in \mathscr{P}(X \times X)$ and an infinite time horizon are immediate given the construction in the proof of Theorem 6.3.

Proof of Theorem 6.1. By iteratively composing copies of $\boldsymbol{\mu}$ with $\boldsymbol{\mu} \circ \mu$, we obtain a measure $\boldsymbol{\rho}_\mu \in \mathscr{P}(\mathrm{C}([0,\infty); X \times X))$ such that $(\mathsf{e}_0)_\sharp \boldsymbol{\rho}_\mu = \mu$. The pathwise coupling properties on each interval $[n-1, n]$, $n \in \mathbb{N}$, which are inherited by $\boldsymbol{\mu}$ carry over to the entire space.

By considering the canonical process $(\mathsf{b}^1, \mathsf{b}^2)$ defined by $\mathsf{b}^1_t(\gamma) := \gamma^1_t$ and $\mathsf{b}^2_t(\gamma) := \gamma^2_t$ under the measure $\boldsymbol{\rho}_\mu$, we immediately obtain the assertion of Theorem 6.1, which is just a stochastic rephrasing of the previous considerations. □

References

[1] L. Ambrosio, N. Gigli, G. Savaré, Gradient Flows in Metric Spaces and in the Space of Probability Measures, second edition, Lectures in Mathematics ETH Zürich, Birkhäuser, Basel, 2008.

[2] L. Ambrosio, N. Gigli, G. Savaré, Density of Lipschitz functions and equivalence of weak gradients in metric measure spaces, Rev. Mat. Iberoam. 29 (3) (2013) 969–996.

[3] L. Ambrosio, N. Gigli, G. Savaré, Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below, Invent. Math. 195 (2) (2014) 289–391.

[4] L. Ambrosio, N. Gigli, G. Savaré, Metric measure spaces with Riemannian Ricci curvature bounded from below, Duke Math. J. 163 (7) (2014) 1405–1490.

[5] L. Ambrosio, N. Gigli, G. Savaré, Bakry–Émery curvature-dimension condition and Riemannian Ricci curvature bounds, Ann. Probab. 43 (1) (2015) 339–404.

[6] L. Ambrosio, A. Mondino, G. Savaré, On the Bakry-Émery condition, the gradient estimates and the local-to-global property of $\mathrm{RCD}^*(K,N)$ metric measure spaces, J. Geom. Anal. 26 (1) (2016) 24–56.

[7] M. Arnaudon, K.A. Coulibaly, A. Thalmaier, Horizontal diffusion in C^1 path space, in: Séminaire de Probabilités XLIII, in: Lecture Notes in Math., vol. 2006, Springer, Berlin, 2011, pp. 73–94.

[8] K. Bacher, K.-T. Sturm, Localization and tensorization properties of the curvature-dimension condition for metric measure spaces, J. Funct. Anal. 259 (1) (2010) 28–56.

[9] D. Bakry, Transformations de Riesz pour les semi-groupes symétriques. II. Étude sous la condition $\Gamma_2 \geq 0$, in: Séminaire de Probabilités, XIX, 1983/84, in: Lecture Notes in Math., vol. 1123, Springer, Berlin, 1985, pp. 145–174.

[10] D. Bakry, M. Émery, Diffusions hypercontractives, in: Séminaire de Probabilités, XIX, 1983/84, in: Lecture Notes in Math., vol. 1123, Springer, Berlin, 1985, pp. 177–206.

[11] V.I. Bogachev, Measure Theory. II, Springer-Verlag, Berlin, 2007.

[12] Z.-Q. Chen, M. Fukushima, Symmetric Markov Processes, Time Change, and Boundary Theory, London Mathematical Society Monographs Series, vol. 35, Princeton University Press, Princeton, NJ, 2012, xvi+479 pp.

[13] S. Daneri, G. Savaré, Eulerian calculus for the displacement convexity in the Wasserstein distance, SIAM J. Math. Anal. 40 (3) (2008) 1104–1122.

[14] M. Erbar, K. Kuwada, K.-T. Sturm, On the equivalence of the entropic curvature-dimension condition and Bochner's inequality on metric measure spaces, Invent. Math. 201 (3) (2015) 993–1071.

[15] N. Gigli, K. Kuwada, S.-I. Ohta, Heat flow on Alexandrov spaces, Commun. Pure Appl. Math. 66 (3) (2013) 307–331.

[16] N. Gigli, On the differential structure of metric measure spaces and applications, Mem. Am. Math. Soc. 236 (1113) (2015), vi+91 pp.

[17] N. Gigli, Nonsmooth differential geometry – an approach tailored for spaces with Ricci curvature bounded from below, Mem. Am. Math. Soc. 251 (1196) (2018), v+161 pp.

[18] N. Gozlan, C. Roberto, P.-M. Samson, Hamilton Jacobi equations on metric spaces and transport entropy inequalities, Rev. Mat. Iberoam. 30 (1) (2014) 133–163.

[19] B.-X. Han, New characterizations of Ricci curvature on RCD metric measure spaces, Discrete Contin. Dyn. Syst. 38 (10) (2018) 4915–4927.

[20] C. Ketterer, Evolution variational inequality and Wasserstein control in variable curvature context, Preprint, arXiv: 1509.02178, 2015.

[21] C. Ketterer, On the geometry of metric measure spaces with variable curvature bounds, J. Geom. Anal. 27 (3) (2017) 1951–1994.

[22] K. Kuwada, Duality on gradient estimates and Wasserstein controls, J. Funct. Anal. 258 (11) (2010) 3758–3774.

[23] S. Lisini, Characterization of absolutely continuous curves in Wasserstein spaces, Calc. Var. Partial Differ. Equ. 28 (1) (2007) 85–120.

[24] J. Lott, C. Villani, Ricci curvature for metric-measure spaces via optimal transport, Ann. Math. (2) 169 (3) (2009) 903–991.

[25] T. Rajala, Interpolated measures with bounded density in metric spaces satisfying the curvature-dimension conditions of Sturm, J. Funct. Anal. 263 (4) (2012) 896–924.

[26] T. Rajala, K.-T. Sturm, Non-branching geodesics and optimal maps in strong $\mathrm{CD}(K,\infty)$-spaces, Calc. Var. Partial Differ. Equ. 50 (3–4) (2014) 831–846.

[27] G. Savaré, Self-improvement of the Bakry–Émery condition and Wasserstein contraction of the heat flow in $\mathrm{RCD}(K,\infty)$ metric measure spaces, Discrete Contin. Dyn. Syst. 34 (4) (2014) 1641–1661.

[28] K.-T. Sturm, Analysis on local Dirichlet spaces. II. Upper Gaussian estimates for the fundamental solutions of parabolic equations, Osaka J. Math. 32 (2) (1995) 275–312.
[29] K.-T. Sturm, On the geometry of metric measure spaces. I, Acta Math. 196 (1) (2006) 65–131.
[30] K.-T. Sturm, On the geometry of metric measure spaces. II, Acta Math. 196 (1) (2006) 133–177.
[31] K.-T. Sturm, Metric measure spaces with variable Ricci bounds and couplings of Brownian motions, in: Festschrift Masatoshi Fukushima, in: Interdiscip. Math. Sci., vol. 17, World Sci. Publ., Hackensack, NJ, 2015, pp. 553–575.
[32] L. Veysseire, Improved spectral gap bounds on positively curved manifolds, Preprint, arXiv:1105.6080, 2011.
[33] C. Villani, Optimal Transport. Old and New, Grundlehren der Mathematischen Wissenschaften., vol. 338, Springer, Berlin, 2009.