Heat Flows on Time-dependent Metric Measure Spaces

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Eva Kopfer
aus
Siegen

Bonn
2017

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Karl-Theodor Sturm

2. Gutachter: Prof. Dr. Jan Maas

Tag der Promotion: 19.12.2017

Erscheinungsjahr: 2018

Acknowledgement

At first I would like to thank my supervisor Prof. Dr. Karl-Theodor Sturm for supporting me during my PhD studies. I am very grateful for the close collaboration. I benefited from his immense knowledge and experience in many helpful discussions.

Secondly, I thank Prof. Dr. Jan Maas for refereeing my thesis as well as the other members of the committee.

Thirdly, I thank my colleagues for the warm atmosphere in our research group. In particular I thank Dr. Matthias Erbar for his constant helpfulness and my office mate Angelo Profeta. Further I thank Dr. Martin Huesmann and Lorenzo Dello Schiavo.

I wish to say thank you to my boyfriend Peter Gladbach, to who I owe so many things. His love and advice helped me greatly.

I thank my family, especially my parents Annette and Herbert Kopfer, for standing behind me during my studies, and in life.

Lastly, I do not want to conclude without appreciating my four-legged friends, Ivo and Bonny.

Summary

In this thesis we discuss the heat flow on time-dependent metric measure spaces. It will be useful to distinguish between the heat flow (on functions forwards in time), the adjoint heat flow (on functions backwards in time) and the dual heat flow (on measures backwards in time). We obtain existence of the heat flow and its adjoint in two different ways, in particular, first, by solving it in a suitable weak sense, and second, by applying a minimizing movement scheme (also referred to as JKO-scheme) using Cheeger's energy on time-dependent L^2-space and the relative entropy on time-dependent L^2-Kantorovich space, respectively. Let us remark that the latter way requires less regularity assumptions on the space. Of particular interest are properties which characterize the underlying space as a super-Ricci flow as introduced by Sturm in [59]. Similar to weak lower Ricci curvature bounds in the sense of Lott, Sturm and Villani, the definition of super-Ricci flows employs a convexity property of the time-dependent relative entropy called dynamic convexity. This thesis is subdivided into three parts.

In the first part we establish the equivalence of dynamic convexity of the relative entropy on the (time-dependent) L^2-Kantorovich space, monotonicity of L^2-Kantorovich distances under the dual heat flow, gradient estimates for the heat flow and a dynamic version of Bochner's inequality involving the time-derivative of the metric. We also give a characterization for the dynamic N-convexity of the relative entropy, where N can be thought of as an upper bound on the dimension. These results represent a dynamic analogue to the characterization of weak curvature-dimension bounds obtained in [6] and [24] and can be seen as a contribution to the research topic of weak Ricci flows, cf. e.g. [46], [29]. Moreover, we characterize the heat flow on functions as the unique forward EVI-flow for Cheeger's energy on the Hilbert space of square integrable functions and the dual heat flow on probability measures as the unique backward EVI-flow for the relative entropy on the L^2-Kantorovich space.

In the second part we strengthen our assumptions on the metrics and obtain refined gradient and transport estimates. As an application we construct Brownian motions such that the distance of their paths is controlled.

In the last part we introduce notions of dynamic gradient flows on time-dependent metric spaces as well as on time-dependent Hilbert spaces. We prove existence of solutions for a given class of time-dependent energy functionals in both settings via a JKO-scheme adapted to our time-dependent setting. In particular we apply our results to the relative entropy on the space of probability measures endowed with the time-dependent L^2-Kantorovich distance and to Cheeger's energy on the time-dependent Hilbert space of L^2-integrable functions. As in the static setting, it is crucial for the existence concerning the relative entropy gradient flow that each underlying metric measure space satisfies a lower Ricci curvature bound. We identify the gradient flow for the time-dependent Cheeger's energy and the gradient flow of the time-dependent relative entropy with the heat flow and the forward dual heat flow, respectively introduced in the first part. This is possible since we obtain uniqueness for the gradient flows of Cheeger's energy and the relative entropy.

Contents

1 Introduction

Over the last several years the theory of optimal transport has proven to be an effective instrument for studying geometric structures for non-smooth spaces on the one hand and for studying diffusion equations on the other. The relative entropy and the Kantorovich distance, objects which are defined on the space of probability measures, play a crucial role in both applications. Concerning the geometry of the underlying space, its curvature is captured in the behavior of the relative entropy along Kantorovich geodesics, while diffusion equations can be characterized as a Kantorovich gradient flow of the relative entropy. Interestingly there is a strong interplay between diffusion equations on the one hand and the geometry of the underlying space which is again reflected in terms of optimal transport.

In this thesis we study diffusion equations on spaces with geometries which evolve in time. The diffusion equations we consider are given by the heat equation and its adjoint. In the first part we prove existence and uniqueness of solutions for both types of equations and give a characterization of super-Ricci flows introduced by Sturm in [59]. The defining property of super-Ricci flows is given by the so-called dynamic convexity of the relatice entropy. We show the one-to-one correspondence of gradient estimates for the heat flow in the sense of Bakry-Émery on the one hand and dynamic convexity of the relativ entropy on the other. Equivalently we obtain that the adjoint heat flow satisfies contraction estimates with respect to the Kantorovich distance. We show in the second part that the gradient estimate and the transport estimate possess the property of self-improving. This leads to pathwise contraction estimates for the trajectories of Brownian motions. In the last part we introduce notions of dynamic gradient flows and prove existence via a time-dependent JKO-scheme. We identify the heat flow with the dynamic gradient flow of Cheeger's energy and the (forward) adjoint heat flow with the dynamic gradient flow of the relative entropy.

In the following we give a brief survey of optimal transport, weak notions of Ricci curvature and gradient flows. Then we give an informal overview of the results obtained in this thesis. The precise statements can be found in the Chapters 2, 3 and 4.

1.1 Optimal Transport

The problem of optimal transport goes back to Monge's work "Mémoire sur la théorie des déblais et des remblais" ([47]) in the late 18th century. He questioned how to transport a certain amount of soil from one place to the other such that the total cost is as low as possible. The modern way to describe the *Monge's optimal transport problem* goes as follows. Let X, Y be two Polish spaces and $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$ two probability measures. Fix a *cost function* $c\colon X \times Y \to \mathbb{R}$. We minimize

$$T \mapsto \int c(x, T(x))\, d\mu(x)$$

among all *transport maps* T from μ to ν, i.e. all measurable maps $T\colon X \to Y$ such that $T_\#\mu = \nu$. The measure $T_\#\mu \in \mathcal{P}(Y)$ is called *push forward of* μ *through* T and is characterized by $\int f\, dT_\#\mu = \int f \circ T\, d\mu$ for all functions

$f\colon Y \to \mathbb{R}$. Unfortunately this formulation carries some disadvantages. For instance if μ is a Dirac measure and ν not, there exists no admissible T. It took almost two hundred years for Kantorovich to propose a relaxation to overcome these difficulties, (see [31] for an English translation of the Russian article from 1942).

Given a cost function $c\colon X \times Y \to \mathbb{R}$ and two probability measures $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$ *Kantorovich's optimal transport problem* consists of minimizing

$$\gamma \mapsto \int_{X\times Y} c(x,y)\, d\gamma(x,y)$$

among all *transport plans* γ from μ to ν, i.e. all probability measures $\gamma \in \mathcal{P}(X \times Y)$ such that $\gamma(A \times Y) = \mu(A)$ and $\gamma(X \times B) = \nu(B)$ for all measurable sets $A \subset X$, $B \subset Y$. Transport plans can be thought of as multivalued transport maps. Every transport map T admits a transport plan $\gamma = (Id \times T)_{\#}\mu$. Moreover there always exists a plan, e.g. $\mu \times \nu$, and under mild assumptions on the cost c there exists even a minimizer. It is well-known that these kind of problems, where a linear functional has to be minimized under a affine constraint, admits a dual problem, where a linear functional has to be maximized.

Kantorovich himself introduced the associated *dual problem* ([31]). It consists of maximizing

$$(\varphi, \psi) \mapsto \int \varphi(x)\, d\mu(x) + \int \psi(y)\, d\nu(y),$$

among all functions $\varphi \in L^1(\mu)$, $\psi \in L^1(\nu)$ such that $\varphi(x) + \psi(y) \leq c(x,y)$. For a cost function which is continuous and bounded from below the minimum of the Kantorovich problem is equal to the supremum of the dual problem,

$$\min_{\gamma} \int_{X\times Y} c(x,y)\, d\gamma(x,y) = \sup_{\varphi,\psi} \int \varphi(x)\, d\mu(x) + \int \psi(y)\, d\nu(y).$$

The supremum is actually a maximum and is of the form (φ, φ^{c+}), where φ^{c+} is the *c_+-transform*

$$\varphi^{c+}(y) := \inf_{x\in X} c(x,y) - \varphi(x).$$

The study of the dual problem reveals significant information for the transport problem and has been performed by several authors, e.g. Knott and Smith [33] and Rachev and Rüschendorf [51].

If $X = Y$ and the cost function is given by the squared distance d^p of X, where p is a natural number, we recover the *L^p-Kantorovich distance* on measures defined by

$$W_p(\mu,\nu) := \Big(\inf_{\gamma} \int d^p(x,y)\, d\gamma(x,y) \Big)^{1/p},$$

where the infimum is taken among all transport plans $\gamma \in \mathcal{P}(X^2)$ from μ to ν. Strictly speaking, this does not define a distance since it is possible that $W_p(\mu,\nu) = \infty$, but if W_p is restricted to the space $\mathcal{P}_p(X)$ of Borel probability measures with finite moments of order p

$$\mathcal{P}_p(X) := \left\{ \mu \in \mathcal{P}(X) : \int d^p(x,x_0)\, d\mu(x) < \infty \right\},$$

where $x_0 \in X$ can be chosen arbitrarily, we recover all the axioms of a distance. A basic fact [62, Theorem 6.9] is that given a Polish space X, i.e. a complete separable metric space, the space $(\mathcal{P}_p(X), W_p)$ is a Polish space as well.

A simple consequence of Hölder's inequality is the fact that

$$p \leq q \Rightarrow W_p \leq W_q.$$

Hence, the metric W_1, also known as the Kantorovich-Rubinstein distance, is the weakest of all the W_p's. The other extreme case is given by $W_\infty := \lim_{p\to\infty} W_p$, which is the most restrictive of all the L^p-Kantorovich distances.

Another famous representative is given by the L^2-Kantorovich distance W_2. One of the interesting features of $(\mathcal{P}_2(X), W_2)$ is that it inherits certain geometric properties of the space X. If (X, d) is a geodesic space, i.e. for each $x, y \in X$ there exists a curve $\gamma\colon [0,1] \to X$ with $\gamma(0) = x$ and $\gamma(1) = y$ such that $d(\gamma(s), \gamma(t)) = |s-t| d(x,y)$, then $(\mathcal{P}_2(X), W_2)$ is geodesic as well. Furthermore, each geodesic $t \mapsto \mu_t$ in $\mathcal{P}_2(X)$ can be lifted to a measure on the geodesics in X such that the joint law of the start and end point produces an optimal transport plan between μ_0 and μ_1, cf. [2, Theorem 2.10]. As an example we consider the Dirac measures δ_x and δ_y for $x, y \in X$. It is important to note that the classical linear interpolation between δ_x and δ_y

$$t \mapsto \mu_t := (1-t)\delta_x + t\delta_y,$$

is not the right object since it has infinite length as soon as $x \neq y$. The right way to interpolate between these measures is given by *displacement interpolation*, i.e. $t \mapsto \delta_{\gamma_t}$, where $t \mapsto \gamma_t$ is a geodesic on X connecting x to y. The terminology for probability measures on $\mathbb{R}^N$ goes back to McCann [45].

1.2 Weak Notions of Lower Ricci Curvature Bounds

The concept of curvature is closely related to the behavior of geodesics. Imagine a point x on a smooth N-dimensional Riemannian manifold (M, g) and a tangential vector v attached to x. Let U be an arbitrary neighbourhood of the point x. Now we transport every point in the neighbourhood along the geodesic with initial velocity v. If we assume that the manifold has positive curvature the geodesics will tend to diverge (at least for short times), whereas negative curvature will mean that geodesics will tend to converge (at least for short times). Both scenarios result in a distortion of the initial neighbourhood U. A simple formula where the Ricci curvature comes into play is given in the following. Let U_t denote the image set of the geodesics at time t, then the volume $\mathrm{vol}(U_t)$ is given by the following formula

$$\mathrm{vol}(U_t) = \mathrm{vol}(U)\Big(1 - \frac{t^2}{2}\mathrm{Ric}(v) + \text{lower order terms}\Big),$$

where $\mathrm{Ric}(v) = \mathrm{Ric}(v,v)$ denotes the Ricci tensor.

The Ricci curvature is said to be bounded from below by some $K \in \mathbb{R}$ if for every v in the tangent space

$$\mathrm{Ric}(v) \geq K g(v).$$

By now lower Ricci curvature bounds are well-understood in terms of optimal transport. Crucial in this context is the *relative entropy functional* or *Boltzmann entropy* on the space of probability measures $\mathcal{P}(M)$, which is given by

$$\mathrm{Ent}(\mu|\mathrm{vol}) = \int \rho \log \rho \, d\,\mathrm{vol} \qquad \text{with } \mu = \rho\,\mathrm{vol}.$$

It turns out that convexity properties of $\mu \mapsto \mathrm{Ent}(\mu|\mathrm{vol})$ are directly related to the curvature of the underlying space. Sturm and von Renesse proved that M has Ricci curvature bounded from below by some $K \in \mathbb{R}$ if and only if the relative entropy is *K-geodesically convex*, i.e. for any pair of measures $\mu, \nu \in Dom(\mathrm{Ent}(\cdot|\mathrm{vol})) \cap \mathcal{P}_2(M)$ there exists a geodesic $(\mu_t) \subset \mathcal{P}_2(M)$ such that $\mu_0 = \mu$ and $\mu_1 = \nu$ and

$$\mathrm{Ent}(\mu_t|\mathrm{vol}) \leq (1-t)\mathrm{Ent}(\mu_0|\mathrm{vol}) + t\mathrm{Ent}(\mu_1|\mathrm{vol}) - \frac{K}{2}t(1-t)W_2^2(\mu_0,\mu_1)$$

for every $t \in [0,1]$. A first hint in that direction has been achieved by Otto and Villani in [50] and later by Cordero-Erasquin, McCann and Schmuckenschläger in [22].

Surprisingly, the notion of geodesic convexity does not employ any differentiable structure of the manifold M, and is suitable to generalize the notion of Ricci curvature bounded from below to the class of metric measure spaces, i.e. metric spaces equipped with a Borel reference measure. We say that a metric measure space (X,d,m) has *Ricci curvature bounded from below by* $K \in \mathbb{R}$ (in short $\mathrm{CD}(K,\infty)$) if the relative entropy $\mathrm{Ent}(\cdot|m)$ is K-geodesically convex on $(\mathcal{P}_2(X), W_2)$. This definition, introduced independently by Sturm in [57] and Lott and Villani in [43], is consistent with the smooth Riemannian case and stable under measured Gromov-Hausdorff convergence.

This notion of lower Ricci bounds is dimension independent, but many geometric applications are not provided until the additional presence of an upper dimension bound. The *curvature-dimension condition* (in short $\mathrm{CD}(K,N)$ where N is an upper bound for the dimension) was introduced by Sturm in [57] and constitutes a tightening up of the much simpler $\mathrm{CD}(K,\infty)$ condition. It provides geometric inequalities such as Brunn-Minkowski, which further leads to volume growth estimates (Bishop-Gromov inequality) and diameter estimates (Bonnet-Myers theorem).

A different approach to describe curvature-dimension bounds has been initiated by Bakry and Émery in [13] by means of the functional Γ-calculus in Dirichlet spaces. Given is a strongly local, symmetric Dirichlet form $\mathcal{E}\colon L^2(X,m) \to [0,\infty]$ on a measure space $(X,\mathcal{B},m)$ generating the Markov semigroup $(P_t)_{t\geq 0}$ in $L^2(X,m)$ with operator $\Delta_\mathcal{E}$. The Dirichlet form admits the representation formula

$$\mathcal{E}(u,v) = \int \Gamma(u,v)\,dm = -\int u\Delta_\mathcal{E} v\,dm,$$

where Γ denotes the so-called *Carré du champ* $\Gamma(u,v) := \frac{1}{2}(\Delta_\mathcal{E}(uv) - u\Delta_\mathcal{E} v - v\Delta_\mathcal{E} u)$ on a suitable algebra $\mathcal{A}$ of functions which are dense in the domain of $\Delta_\mathcal{E}$. The intrinsic distance $d_\mathcal{E}$ induced by the Dirichlet form is given by

$$d_\mathcal{E}(x,y) := \sup\{\psi(y) - \psi(x) | \psi \in \mathcal{A}, \Gamma(\psi) \leq 1\}.$$

As a basic example consider the Dirichlet energy

$$\mathcal{E}(u,v) = \int \nabla u \cdot \nabla v \, d\mathrm{vol}$$

on a n-dimensional smooth Riemannian manifold (M, vol) endowed with its natural volume. Consequently, $\Gamma(u) = |\nabla u|^2$ and $\Delta_{\mathcal{E}} = \Delta$, where Δ denotes the usual Laplace-Beltrami operator on M, and one recovers the geodesic distance d in M by the intrinsic distance $d(x,y) = d_{\mathcal{E}}(x,y)$. The crucial observation is the fact that the manifold has Ricci curvature bounded from below by K if and only if *Bochner's inequality* holds

$$\frac{1}{2}\Delta_{\mathcal{E}}\Gamma(u) - \Gamma(u, \Delta_{\mathcal{E}} u) \geq \frac{1}{N}(\Delta_{\mathcal{E}} u)^2 + K\Gamma(u), \tag{1}$$

where $N \geq n$. Using the notion of the *Carré du champ itéré*

$$2\Gamma_2(u,v) := \Delta_{\mathcal{E}}\Gamma(u,v) - \Gamma(u, \Delta_{\mathcal{E}} v) - \Gamma(v, \Delta_{\mathcal{E}} u),$$

Bochner's inequality can be expressed as $\Gamma_2(u) := \Gamma_2(u,u) \geq \frac{1}{N}(\Delta_{\mathcal{E}} u)^2 + K\Gamma(u)$.

The resulting weak notion of curvature-dimension bounds called *Bakry-Émery condition* and in short $\mathrm{BE}(K,N)$, is obtained by using (1) as definition. Essentially, considering the case $N = \infty$ for plainness, the property $\mathrm{BE}(K,\infty)$ is equivalent to the pointwise gradient estimate for the Markov semigroup

$$\Gamma(P_t u) \leq e^{-2Kt} P_t(\Gamma(u)),$$

see e.g. [64]. This curvature-dimension condition implies many functional and geometric inequalities like Poincaré, Log-Sobolev and Talagrand inequality (see e.g. [14], [24], [62]).

Another feature of the gradient estimate is that it is self-improving, since it leads to the stronger contraction estimate

$$\big(\Gamma(P_t u)\big)^\alpha \leq e^{-2\alpha K t} P_t\big(\Gamma(u)^\alpha\big) \text{ for every } \alpha \in [1/2, 2].$$

This has been shown by Bakry in [12], and later by Savaré in the setting of metric measure spaces [55]. Both authors prove the stronger gradient estimate by showing

$$\Gamma(\Gamma(u)) \leq 4\big(\Gamma_2(u) - K\Gamma(u)\big)\Gamma(u),$$

which represents an already stronger version of Bochner's inequality.

1.3 Gradient Flows

In [30] Jordan, Kinderlehrer and Otto showed that the solution to the heat equation

$$\partial_t \rho_t = \Delta \rho_t \quad \text{on } \mathbb{R}^n \times (0, \infty)$$

is the relative entropy gradient flow on the space of probability measures with respect to the L^2-Kantorovich distance. They constructed the solution via a discrete approximation procedure. This procedure is called by many names in the literature: minimizing movement-, implicit Euler-, or JKO-scheme.

By now this result has been extended to more general settings, like Riemannian manifolds [23], Hilbert spaces [9], Finsler spaces [49], Alexandrov spaces [28] and metric measure spaces satisfying Ricci curvature bounds [5].

There are several ways to define gradient flows in metric spaces, which are not necessarily equivalent. For a comprising study we refer to the monograph [4]. Let us start with a very strong formulation called *EVI-gradient flow.* To motivate this let $E\colon \mathbb{R}^n \to (-\infty, +\infty]$ be a convex and lower semicontinuous functional. A smooth curve $x\colon [0,\infty) \to \mathbb{R}^n$ solves the gradient flow equation $\dot{x}_t = -\nabla E(x_t)$ if and only if it satisfies

$$\frac{d}{dt}\frac{1}{2}|x_t - y|^2 \leq E(y) - E(x_t) \qquad \forall y \in \mathbb{R}^n.$$

The latter formulation requires only the metric structure of the space and is therefore suitable to be taken as the definition of a gradient flow in metric spaces. Applied to the metric space $(\mathcal{P}_2(\mathbb{R}^n), W_2)$ the solution ρ_t to the heat equation is the gradient flow in the following sense

$$\frac{d}{dt}\frac{1}{2}W_2(\mu_t, \sigma)^2 \leq E(\sigma) - E(\mu_t) \qquad \forall \sigma \in \mathcal{P}_2(\mathbb{R}^n), \tag{2}$$

where $\mu_t = \rho_t\, dx$ and $E(\mu_t) = \int \rho_t \log \rho_t\, dx$. One consequence of estimate (2) is the contraction of flows, i.e. for two flows μ_t, ν_t solving (2) we have

$$W_2(\mu_t, \nu_t) \leq W_2(\mu_0, \nu_0).$$

In particular we immediately obtain that EVI-gradient flows are unique, i.e. given a probability measure $\bar{\mu}$, there exists at most one EVI-gradient flow μ_t starting in μ_0.

Let us now come back to the case where the underlying space is a general metric measure space (X, d, m). The *heat flow* on metric measure spaces is defined as the EVI-gradient flow of Cheeger's energy

$$\mathrm{Ch}(u) = \frac{1}{2} \inf \left\{ \liminf_{n\to\infty} \int_X (\mathrm{lip}\, u_n)^2 dm \mid u_n \in \mathrm{Lip}(X), \int_X |u_n - u|^2 dm \to 0 \right\},$$

on the Hilbert space of L^2-integrable functions $L^2(X, m)$. Here $\mathrm{lip}\, u\colon X \to [0, \infty]$ denotes the local Lipschitz constant. Since $u \mapsto \mathrm{Ch}(u)$ defines a lower semicontinuous and convex functional on a Hilbert space, existence and uniqueness is guaranteed by the general theory of monotone operators, cf. [19]. This flow is characterized by the fact that it solves the *heat equation* in the following sense

$$\frac{d^+}{dt} u_t = \Delta_{\mathrm{Ch}} u_t,$$

where $\frac{d^+}{dt}$ denotes the right derivative and $-\Delta_{\mathrm{Ch}} u$ is the element of minimal $L^2(X, m)$ norm in the subdifferential $D^-\mathrm{Ch}(u)$. The subdifferential $D^- F$ is a generalization of the gradient ∇F for convex functionals F which are not necessarily differentiable. $D^-\mathrm{Ch}(u)$ consists of all $v \in L^2(X, m)$ such that

$$\int v(g - u)\, dm \leq \mathrm{Ch}(g) - \mathrm{Ch}(u) \text{ for every } g \in L^2(X, m).$$

We will see soon that it cannot be taken for granted that the solution to the heat equation is an EVI-gradient flow for the relative entropy on $(\mathcal{P}_2(X), W_2)$. This is closely related to the potential lack of linearity of the operator Δ_{Ch}. Nonetheless under the assumption that (X, d, m) satisfies $\mathrm{CD}(K, \infty)$ the solution to the heat equation (ρ_t) solves

$$\mathrm{Ent}(\mu_0|m) = \mathrm{Ent}(\mu_t|m) + \frac{1}{2}\int_0^t |\dot{\mu}_s|^2\, ds + \int_0^t |\nabla \mathrm{Ent}|^2(\mu_s|m)\, ds, \tag{3}$$

where $\mu_t = \rho_t m$, $|\dot{\mu}_t|$ denotes the metric speed and $|\nabla \mathrm{Ent}|$ denotes the slope. A curve μ_t which solves (3) is called *EDE-gradient flow*. Moreover the heat flow can be unambiguously defined as the EVI-gradient flow of Cheeger's energy on $L^2(X)$ or as the EDE-gradient flow of the relative entropy on $\mathcal{P}_2(X)$. The identification is feasible thanks to Gigli, who showed uniqueness for solutions of (3) in [26]. This result is surprising since no contraction properties can be expected at this general level.

On the other hand if there exists a curve (μ_t) satisfying

$$\frac{d}{dt}\frac{1}{2}W_2(\mu_t, \sigma)^2 + \frac{K}{2}W_2(\mu_t, \sigma)^2 \leq \mathrm{Ent}(\sigma|m) - \mathrm{Ent}(\mu_t|m) \qquad \forall \sigma \in \mathcal{P}_2(X), \tag{4}$$

then (X, d, m) has a *Riemannian Ricci curvature* bounded from below by $K \in \mathbb{R}$ (in short $\mathrm{RCD}(K, \infty)$), i.e. (X, d, m) satisfies $\mathrm{CD}(K, \infty)$ and the heat flow is linear. The latter is also equivalent to saying Cheeger's energy constitutes a bilinear form in $L^2(X, m)$. Remarkably, this is also true for the converse implication; if (X, d, m) is a $\mathrm{RCD}(K, \infty)$ space then estimate (4) holds.

This notion of Riemannian curvature bounds has been introduced by Ambrosio, Gigli and Savaré in [6] and provides a bridge between the gap of $\mathrm{CD}(K, \infty)$ and $\mathrm{BE}(K, \infty)$ spaces in the sense that $\mathrm{BE}(K, \infty)$ is equivalent to $\mathrm{RCD}(K, \infty)$: If the Polish space X endowed with probability measure m and Dirichlet form $\mathcal{E}$ satisfies $\mathrm{BE}(K, \infty)$ then under minimal technical assumptions the measure space (X, m) endowed with the induced metric $d_{\mathcal{E}}$ is a $\mathrm{RCD}(K, \infty)$ space. Conversely, if (X, d, m) is a $\mathrm{RCD}(K, \infty)$ space then (X, d, m) equipped by $\mathcal{E} := 2\mathrm{Ch}$ is a $\mathrm{BE}(K, \infty)$ space and $d_{\mathcal{E}} = d$.

By Kuwada's duality approach in [36], the self-improvement of the gradient estimates shown by Savare in [55] leads to stronger contraction estimates for the heat flow on measures

$$W_p(\mu_t, \nu_t) \leq e^{-Kt} W_p(\mu, \nu) \text{ for every } \mu, \nu \in \mathcal{P}(X), p \in [1, \infty],$$

which is consistent with the Riemannian case, see [63].

Finally, as already mentioned above, many geometric and functional inequalities are present only under combined curvature-dimension condition. The results in [6, 7] has been generalized to $\mathrm{RCD}(K, N)$ and $\mathrm{BE}(K, N)$ spaces respectively by Erbar, Kuwada and Sturm in [24].

1.4 The Results of Chapter 2

In the first part of this thesis we describe the evolution of geometries in terms of optimal transport and present a time-dependent version of the characterization for lower Ricci curvature bounds obtained by Erbar, Kuwada and Sturm in [24].

In the smooth Riemannian setting a family of metric tensors (g_t) is a Ricci flow if $-\frac{1}{2}\partial_t g = \mathrm{Ric}$, where Ric is the Ricci tensor of g. Similar to the static case where one studies lower curvature bounds we relax the notion of Ricci flows by requiring g to be "only" a super-Ricci flow $-\frac{1}{2}\partial_t g \leq \mathrm{Ric}$.

The aim of Chapter 2 is to characterize weak super-Ricci flows by means of optimal transport and Bakry-Émery calculus. We will rely on the notion of weak super-Ricci flows on time-dependent metric measure spaces introduced by Sturm via optimal transport in [59]. The defining property is obtained by the notion of *dynamic convexity* of the relative entropy, which has been initiated in [59], too.

Let $(X, d_t, m_t)_{t\in(0,T)}$ be a family of metric measure spaces such that the map $t \mapsto \log d_t(x, y)$ is Lipschitz continuous uniformly in x and y and $m_t = e^{-f_t} m$ for some reference measure m. We assume that the logarithmic densities $f\colon (0, T) \times X \to \mathbb{R}$ are Lipschitz continuous in time and space and each (X, d_t, m_t) is a $\mathrm{RCD}(K, N)$ space for some finite numbers K and N. Hence $\mathcal{E}_t := 2\mathrm{Ch}_t$ defines a symmetric strongly local Dirichlet form.

We introduce the heat equation

$$\partial_t u_t = \Delta_t u_t \quad \text{ on } (s, \tau) \times X$$

as well as the adjoint heat equation

$$\partial_s v_s = -\Delta_s v_s + (\partial_s f_s) v_s \quad \text{ on } (\sigma, t) \times X.$$

Both equations are interpreted in a distributional sense, i.e. u solves the heat equation if

$$\int_s^\tau \int_X w_t \partial_t u_t \, dm_t dt = -\int_s^\tau \mathcal{E}_t(u_t, w_t)\, dt \quad \forall w$$

and v solves the adjoint heat equation if

$$\int_\sigma^t \int_X w_s \partial_s v_s \, dm_s \, ds = \int_\sigma^t \mathcal{E}_s(v_s, w_s) + \int_X (\partial_s f_s) v_s w_s \, ds \quad \forall w,$$

where w is chosen from a suitable class of test functions. Here, the adjoint heat equation has to be understood backwards in time, i.e. in order to obtain existence of solutions one has to prescribe terminal data.

We show that a number of regularity properties hold, e.g. existence, uniqueness and kernel representations. We refer to Chapter 2 for the detailed statements. Here, we will only stick to the core message.

We denote the heat flow $t \mapsto P_{t,s} u$ as the solution to the heat equation such that $\lim_{t\searrow s} P_{t,s} u = u$ in $L^2(X, m)$ and the adjoint heat flow $s \mapsto P^*_{t,s} v$ as the solution to the adjoint heat equation such that $\lim_{s\nearrow t} P^*_{t,s} v = v$ in $L^2(X, m)$. They are adjoint in the following way

$$\int u P^*_{t,s} v \, dm_s = \int P_{t,s} u v \, dm_t.$$

It turns out that the heat flow preserves constants, whereas the adjoint heat flow is mass preserving in the sense that $\int P^*_{t,s} v \, dm_s = \int v \, dm_t$. Consequently, we define the dual heat flow $s \mapsto \hat{P}_{t,s}\mu$ on measures by duality $\int u \, d\hat{P}_{t,s}\mu = \int P_{t,s} u \, d\mu$.

In [59] Sturm introduced the notion of dynamic convexity. We say that the relative entropy $S_t(\mu) = \mathrm{Ent}(\mu|m_t)$ is *dynamically convex* if for a.e. $t \in (0,T)$ and every W_t-geodesic $(\mu_a)_{a\in[0,1]}$ it holds

$$\partial_a S_t(\mu_a)|_{a=1} - \partial_a S_t(\mu_a)|_{a=0} \geq -\frac{1}{2} W_t^2(\mu^0, \mu^1),$$

where $W_t = W_{2,t}$ denotes the L^2-Kantorovich distance with respect to d_t.

We prove that dynamic convexity of the relative entropy can be equivalently obtained in terms of contraction estimates for the dual heat flow $\hat{P}_{t,s}$ and the heat flow $P_{t,s}$;

$$\Gamma_t(P_{t,s}u) \leq P_{t,s}(\Gamma_s u), \tag{5}$$

$$W_s(\hat{P}_{t,s}\mu, \hat{P}_{t,s}\nu) \leq W_t(\mu,\nu). \tag{6}$$

The contraction estimates for the dual heat flow result from the fact that this flow solves a dynamic version of EVI. The contraction estimates for the heat flow on the other hand constitute a gradient estimate in the spirit of Bakry-Émery, which can be equivalently stated as a time-dependent version of Bochner's inequality

$$\frac{1}{2}\Delta_r\Gamma_r(u_r) - \Gamma_r(u_r, \Delta_r u_r) \geq \frac{1}{2} \overset{\bullet}{\Gamma}_r (u_r) \text{ for a.e. } r,$$

where $u_r = P_{r,s}u$ and $\overset{\bullet}{\Gamma}_r$ is a weak version of the time-derivative $\partial_r \Gamma_r$. This has to be interpreted in a distributional sense, namely tested against adjoint heat flows $\hat{P}_{t,r}g$, where $g \geq 0$.

In this sense our main result can be seen as the dynamic counterpart of $\mathrm{RCD}(K,\infty) \Leftrightarrow \mathrm{BE}(K,\infty)$ obtained in [7, 6].

Recalling the static setting, the curvature-dimension condition $\mathrm{CD}(K,N)$ has been introduced since it provides a broader range of geometric applications. In the same spirit the notion of super-Ricci flows has been tightened up to N-super-Ricci flows [59], which we will use to obtain the dynamic counterpart $\mathrm{RCD}(K,N) \Leftrightarrow \mathrm{BE}(K,N)$ from [24]. In particular we characterize N-super-Ricci flows by means of N-dimensional contraction estimates of the heat flows and a dynamic version of Bochner's inequality.

But let us emphasize that many properties which are available for the heat semigroup on static metric measure spaces hold no longer true for the heat propagator on time-dependent metric measure spaces. For example it is not clear whether the operator and semigroup commute, or that the semigroup maps L^2 into the domain of the operator. In particular the domain of the Laplace operator $Dom(\Delta_t)$ will depend on time.

A similar result in the framework of smooth families of compact Riemannian manifolds which characterizes super solutions of Ricci flows has been established by McCann and Topping in [46]. They show that super solutions of Ricci flows can be equivalently characterized by the contraction estimate (6). Arnaudon and Coulibaly and Thalmeier [10] define a Brownian motion with time-dependent metric on a compact manifold and obtain a Bismut type formula if the metrics evolve as a Ricci flow. Philipowski and Kuwada [38, 39] obtain McCann's and Topping's result as a corollary for non compact Riemannian manifolds with uniform lower Ricci curvature bound by constructing couplings of Brownian motions. Lakzian and Munn [40] adopted the characterization of super-Ricci flow by McCann and Topping to a family of distance metrics defined on the disjoint union of closed manifolds M_1, M_2. They show that this is a super-Ricci flow provided that the distance function itself is a super solution to the heat equation on $M_1 \times M_2$. Haslhofer and Naber [29] characterize Ricci flows of Riemannian manifolds in terms of infinite-dimensional gradient estimates and suggest a weak notion of Ricci flows based on this characterization. Kleiner and Lott [32] introduce singular Ricci flows, which arise from Ricci flows with surgery starting from a compact three-dimensional Riemannian manifold as the surgery parameter tends to zero.

The results presented in Chapter 2 are obtained together with Karl-Theodor Sturm in the preprint [35] and can be seen as a contribution to the research of weak Ricci flows.

1.5 The Results of Chapter 3

Starting from the fact shown in Chapter 2 that being a super-Ricci flow implies the gradient and transport estimates

$$\big(\Gamma_t(P_{t,s}u)\big) \le P_{t,s}\big(\Gamma_s(u)\big), \qquad W_s(\hat{P}_{t,s}\mu, \hat{P}_{t,s}\nu) \le W_t(\mu,\nu)$$

for the heat flows $P_{t,s}, \hat{P}_{t,s}$, it is quite natural to ask whether we can obtain stronger estimates as in the case of static $\mathrm{RCD}(K,\infty)$ spaces. Crucial for this is the self-improvement property of the gradient estimate which itself arises from the self-improving of Bochner's inequality.

Let $(X, d_t, m_t)_{t\in(0,T)}$ be a family of metric measure spaces. As before we assume that each (X, d_t, m_t) is a $\mathrm{RCD}(K,N)$ space for some finite numbers K and N and $m_t = e^{-f_t}m$ for some reference measure m such that the logarithmic densities $f\colon (0,T)\times X \to \mathbb{R}$ are Lipschitz continuous in time and space. We will strengthen our assumptions on the metrics in the sense that the map $t \mapsto \log d_t(x,y)$ is continuously differentiable with logarithmic derivative which is uniformly bounded and in a certain sense "well-behaved" on the diagonal.

Firstly, we show time-differentiability of the Γ-operator. Along with the L^2-Kantorovich transport estimate we obtain a "real" dynamic Bochner inequality

$$\Delta_t\Gamma_t(u) - \Gamma_t(u, \Delta_t u) \ge \partial_t\Gamma_t(u),$$

real in the sense that it involves the derivative of the Γ-operator, and it holds in a class of test function which do not only arise as a heat flow.

With this we can proceed and adapt the strategy of Bakry and Savaré in [12] and [55] respectively, and derive the crucial estimate

$$\Gamma_t(\Gamma_t(u)) \le 4\Big(\Gamma_{2,t}(u) - \frac{1}{2}\partial_t\Gamma_t(u)\Big)\Gamma_t(u),$$

by applying Bochner's inequality to polynomials. Note that this approach has not been applicable in Chapter 2, since arbitrary polynomials are not necessarily descended from a heat flow. We then follow the ideas from Chapter 2 and derive

$$\big(\Gamma_t(P_{t,s}u)\big)^\alpha \le P_{t,s}\big(\Gamma_s(u)^\alpha\big) \quad \text{ for every } \alpha \in [1/2, 1].$$

Finally, applying the duality approach by Kuwada in [36], we obtain for the heat flow on measures the stronger transport estimate

$$W_{p,s}(\hat{P}_{t,s}\mu, \hat{P}_{t,s}\nu) \leq W_{p,t}(\mu, \nu) \quad \text{for every } p \in [1, \infty],$$

where $W_{p,s}$ denotes the L^p-Kantorovich distance with respect to the metric d_s. In particular, for $\mu = \delta_x, \nu = \delta_y$ and $p = \infty$ we find the following estimate for the heat kernel

$$W_{\infty,s}(p_{t,s}(y, \cdot), p_{t,s}(x, \cdot)) \leq d_t(x, y),$$

where $p_{t,s}(x, dz) = p_{t,s}(x, z)\, dm_s(z)$. As an application we will introduce Brownian motions on time-dependent metric measure spaces and construct a stochastic process on $X \times X$ which is a coupling of Brownian motions $(X_s^1)_{s\leq t}, (X_s^2)_{s\leq t}$ on X such that almost surely

$$d_s(X_s^1, X_s^2) \leq d_t(x, y).$$

These types of transport estimates are reminiscent of the static $\mathrm{RCD}(K, \infty)$ case, where the lower Ricci curvature bound K controls how fast the distance between two distributions may expand, or has to diminish, in time. Here, the super-Ricci flow, which is given by the lower Ricci curvature bound $-\partial_t g_t/2$ if we think of a smooth setting, controls the expansion of mass in time in a similar way. Note that the map $s \mapsto W_{p,s}(\mu_s, \nu_s)$ is non-decreasing. Speaking of a backward super-Ricci flow, as in [46] on Riemannian manifolds, this leads to a contraction of mass in time.

A similar result as in Theorem 3.1 and Theorem 3.2 has been derived by Haslhofer and Naber in [29] in the case of smooth Riemannian manifolds evolving as a super-Ricci flow. They give a characterization of super-Ricci flows in terms of a gradient estimate as in Theorem 3.2 with $\alpha = 1$ and $\alpha = 1/2$ and in terms of a Bochner's formula.

Arnaudon, Coulibaly and Thalmeier [10] showed existence of Brownian motions on a smooth time-dependent setting and apply their results to Ricci flows. Kuwada and Philipowski [38] construct couplings of Brownain motions such that the normalized Perelman's $\mathcal{L}$-distance of the coupling is a supermartingale, see also [60]. This construction is obtained on smooth Riemannian manifolds evolving as a super-Ricci flow.

1.6 The Results of Chapter 4

We have seen that the solution to the heat equation in $\mathbb{R}^n$ can be obtained as a gradient flow of the relative entropy with respect to the L^2-Kantorovich distance via the JKO-scheme. We will show a similar result in Chapter 4.

Let $(X, d_t, m_t)_{t\in[0,T]}$ be a family of metric measure spaces, such that the map $t \mapsto \log d_t(x, y)$ is Lipschitz continuous uniformly in x and y and there exist a Borel probability measure m and a measurable function $f\colon [0, T] \times X \to \mathbb{R}$ such that $e^{-f_t}m = m_t$. We assume that the logarithmic densities f are Lipschitz continuous with respect to the time variable.

We show that the solution to the forward adjoint heat equation introduced in Chapter 2

$$\partial_t \rho_t = \Delta_t \rho_t + (\partial_t f_t)\rho_t \quad \text{on } (0, T) \times X \tag{7}$$

can be obtained as a gradient flow of the relative entropy with respect to the L^2-Kantorovich distance on the one hand and on the other we obtain the solution to the heat equation from Chapter 2

$$\partial_t u_t = \Delta_t u_t \quad \text{ on } (0,T) \times X \tag{8}$$

as a gradient flow of Cheeger's energy with respect to the L^2-norm.

In the case of the relative entropy $S_t = \mathrm{Ent}(\cdot|m_t)$ on the time-dependent metric space $(\mathcal{P}_2(X), W_t)$ we obtain existence of a gradient flow μ_t in the EDE sense, i.e.

$$S_t(\mu_t) + \frac{1}{2}\int_0^t |\dot{\mu}_r|_r^2 dr + \frac{1}{2}\int_0^t |\nabla_r S_r|^2(\mu_r) dr = S_0(\mu_0) + \int_0^t (\partial_r S_r)(\mu_r) dr, \tag{9}$$

if we additionally assume that each underlying metric measure space (X, d_t, m_t) satisfies a lower Ricci curvature bound in the sense of Lott, Sturm and Villani. This equality can be seen as a time-dependent extension of equality (3). The extension appears in the time-dependence of the metric speed $|\dot{\mu}_r|_r^2$, the slope $|\nabla_r S_r|^2$ and the time derivative $(\partial_r S_r)$ of the functional. Further we prove that solutions of (9) are unique. This result enables us to identify the entropy gradient flow with the solution ρ_t to the forward adjoint heat equation (7) by showing that $\rho_t m_t$ solves (9).

We will prove existence for such types of gradient flows on more general time-dependent metric spaces (X, d_t) and for a broader class of energy functionals $E\colon [0,T] \times X \to (-\infty, \infty]$. For this we use a JKO-scheme adapted to our time-dependent setting in the following way. We fix a step size $h > 0$ and an initial value $\bar{x}$, and define recursively for $nh \leq T$

$$x_0^h := \bar{x}, \qquad x_n^h := \arg\min_x \Big\{ E_{nh}(x) + \frac{1}{2h} d_{nh}^2(x, x_{n-1}^h) \Big\}. \tag{10}$$

Under sufficient regularity assumption (see Section 4.3) we are able to show existence of a subsequence $h \to 0$ and a limit curve (x_t) such that the constant interpolations $\bar{x}_t^h$ converge to x_t as h goes to 0. This limit curve constitutes a dynamic gradient flow in the EDE sense. In the special case of the entropy on probability space we even obtain uniqueness of the flow using the convexity properties of the squared metric speed and squared slope, noting that the time derivative $\partial_t S_t$ is a linear perturbation.

Concerning the gradient flow for Cheeger's energy Ch_t on the time-dependent Hilbert space $L^2(X, m_t)$, we show that there exists a gradient flow u_t in the sense that

$$\partial_t u_t \in -D_t^- \mathrm{Ch}_t(u_t) \qquad \text{for a.e. } t \in (0,T), \tag{11}$$

where D_t^- denotes the subdifferential with respect to the scalar product $\langle g, h \rangle_t = \int_X gh\, dm_t$. We identify Cheeger's energy gradient flow with the heat flow on functions via the dynamic EVI introduced in Chapter 2.

We obtain existence via the JKO-scheme (10) applied to a given class of convex energy functionals on time-dependent Hilbert spaces. Let us emphasize that many properties we have in the static setting for this kind of gradient flow are no longer true in the time-dependent setting, e.g. a minimal selection principle, i.e. that the minimal element with respect to the norm is attained.

Let us conclude that the existence of the entropy gradient flow as well as the existence of Cheeger's energy gradient flow are obtained in a more general framework than the one in Chapter 2. For the entropy gradient flow we require that each static space (X, d_t, m_t) satisfies $\mathrm{CD}(K, \infty)$ instead of $\mathrm{RCD}(K, N)$ and that the logarithmic density $f_t \colon X \to \mathbb{R}$ of the measure $m_t = e^{-f_t} m$ are Lipschitz continuous only in time and not in space. For Cheeger's energy gradient flow we only require the Lipschitz continuity of the logarithmic densities f_t.

Gradient flow formulations for time-dependent functionals similar to (9) and (11) have been considered recently. Rossi, Mielke and Savaré in [53] investigate the doubly nonlinear evolution equation on a reflexive Banach space V

$$D^- \Psi(\partial_t u_t) + F_t(u_t) \ni 0 \text{ in } V^* \text{ a.e.},$$

where Ψ is a convex potential and F is a time-dependent family of multivalued maps. They prove existence of gradient flows using a time-dependent JKO scheme. Ferreira and Valencia-Guevara [25] introduce gradient flows for time-dependent functionals on metric spaces and apply their results to a class of PDEs on $\mathbb{R}^n$ such as the Fokker-Planck equation

$$\partial_t \rho = \kappa \Delta \rho + \nabla \cdot (\nabla V(t, x) \rho).$$

The results presented in Chapter 4 are obtained in the preprint [34]. The techniques we use to obtain existence of the entropy gradient flow are inspired by [53] and [25]. Concerning Cheeger's energy gradient flow we adopt the methods in [48] to our time-dependent setting.

2 Heat Flows on Time-dependent Metric Measure Spaces and Super-Ricci Flows

In this chapter we study the heat flow on time-dependent metric measure spaces. With the help of the heat flow we obtain equivalent characterizations to the notion of weak super-Ricci flows introduced by Sturm in [59] in terms of optimal transport. These notions consist of Bakry-Émery-like gradient estimates, a dynamic version of Bochner's inequality and L^2-Kantorovich contraction estimates. We prove that the heat flow emerges as a EVI-like gradient flow with respect to Cheeger's energy on the space of L^2-integrable functions on the one hand and with respect to the relative entropy on the space of probability measures on the other. These results represent a time-dependent version of the characterization of curvature-dimension bounds obtained by Ambrosio, Gigli and Savaré in [7] and Erbar, Kuwada and Sturm in [24] respectively.

2.1 Main Results

We consider a *time-dependent metric measure space* $\big(X, d_t, m_t\big)_{t\in I}$ where $I = (0,T)$ and X is a compact space equipped with one-parameter families of geodesic metrics d_t and Borel measures m_t. We always assume the measures m_t are mutually absolutely continuous with bounded, Lipschitz continuous logarithmic densities and that the metrics d_t are uniformly bounded and equivalent to each other with

$$\left|\log \frac{d_t(x,y)}{d_s(x,y)}\right| \leq L \cdot |t-s| \tag{12}$$

('log Lipschitz continuity'). Moreover, we assume that for each t the static space (X, d_t, m_t) satisfies a Riemannian curvature-dimension condition in the sense of [3], [24]. (In various respects, the latter is not really a restriction, see Lemma 2.8.)

Thus for each t, the detailed analysis in [7] guarantees a well-defined *Laplacian* Δ_t on $L^2(X, m_t)$ characterized by $-\int_X \Delta_t u\, v\, dm_t = \mathcal{E}_t(u,v)$ where the Dirichlet energy

$$\mathcal{E}_t(u,u) = \int_X |\nabla_t u|^2 dm_t = \liminf_{\substack{v\to u \text{ in } L^2(X,m_t)\\ v\in \mathrm{Lip}(X,d_t)}} \int_X (\mathrm{lip}_t v)^2\, dm_t$$

is defined either in terms of the minimal weak upper gradient $|\nabla_t u|$ of $u \in L^2(X, m_t)$ or alternatively in terms of the pointwise Lipschitz constant $\mathrm{lip}_t v(.)$.

Heat equation

Our first important result concerns existence and uniqueness for solutions of two types of diffusion equations on the time-dependent metric measure space $(X, d_t, m_t)_{t\in I}$. The heat equation acting on functions forward in time as well as for the adjoint heat equation acting on functions backward in time. Moreover, it yields regularity of solutions and representation as integrals with respect to a heat kernel.

Theorem 2.1. *There exists a* heat kernel p *on* $\{(t,s,x,y) \in I^2 \times X^2 : t > s\}$*, Hölder continuous in all variables and satisfying the propagator property* $p_{t,r}(x,z) = \int p_{t,s}(x,y) p_{s,r}(y,z)\, dm_s(y)$*, such that* *(i) for each $s \in I$ and $h \in L^2(X, m_s)$*

$$(t,x) \mapsto P_{t,s}h(x) := \int p_{t,s}(x,y)h(y)\,dm_s(y)$$

is the unique solution to the heat equation

$$\partial_t u_t = \Delta_t u_t \qquad \text{on } (s,T)\times X$$

with $u_s = h$;

(ii) for each $t \in I$ and $g \in L^2(X, m_t)$

$$(s,y) \mapsto P^*_{t,s}g(y) := \int p_{t,s}(x,y)g(x)\,dm_t(x)$$

is the unique solution to the adjoint heat equation

$$\partial_s v_s = -\Delta_s v_s + \dot{f}_s \cdot v_s \qquad \text{on } (0,t)\times X$$

with $v_t = g$. Here $\dot{f}_s = -\partial_t\big(\frac{dm_t}{dm_s}\big)\big|_{t=s}$.

Let us emphasize that many properties of the heat kernel available in the static setting drop away in the time-dependent setting. For example we cannot hope that the propagator $P_{t,s}$ is symmetric, neither to m_s nor to m_t, or that it commutes with the operator Δ_t, or Δ_s. Moreover the operators Δ_t depend strongly on time and the the propagator $P_{t,s}$ does not map $L^2(X)$ into the domain $Dom(\Delta_t)$ for each t. Nonetheless we derive various important L^2-properties and estimates – partly in the more general setting of heat flows for time-dependent Dirichlet forms – the most prominent of them being the EVI-characterization, the energy estimate and the commutator lemma.

Theorem 2.2. *(i) The heat flow is uniquely characterized as the* dynamic forward EVI$(-L/2,\infty)$-flow *for $\frac{1}{2}\mathcal{E}$ on $L^2(X,m_t)_{t\in I}$ in the following sense: for all solutions $(u_t)_{t\in(s,\tau)}$ to the heat equation, for all $\tau \le T$ and all $w \in Dom(\mathcal{E})$*

$$-\frac{1}{2}\partial_s^+ \big\|u_s - w\big\|^2_{s,t}\Big|_{s=t} + \frac{L}{4}\cdot\big\|u_s - w\big\|^2_{s,t} \;\ge \frac{1}{2}\mathcal{E}_t(u_t) - \frac{1}{2}\mathcal{E}_t(w).$$

(ii) For all $s \in (0,T)$ and $u \in Dom(\mathcal{E}_s)$

$$P_{t,s}u \in Dom(\Delta_t) \qquad \text{for a.e. } t > s$$

and $\int_s^\tau e^{-3L(t-s)}\int|\Delta_t P_{t,s}u|^2 dm_t\,dt \le \frac{1}{2}\mathcal{E}_s(u)$ for all $\tau > s$..

(iii) For all $\sigma < \tau$, all $u, v \in L^2$ and a.e. $s,t \in (\sigma,\tau)$ with $s < t$

$$\int \Big[\Delta_t P_{t,s}u_s - P_{t,s}\Delta_s u_s\Big] v_t\,dm_t \le C\cdot\sqrt{t-s}$$

where $u_s = P_{s,\sigma}u$, $v_t = P^*_{\tau,t}v$.

We define the *dual heat flow* $\hat{P}_{t,s} : \mathcal{P}(X) \to \mathcal{P}(X)$ by

$$(\hat{P}_{t,s}\mu)(dy) = \Big[\int p_{t,s}(x,y)\,d\mu(x)\Big]\, m_s(dy).$$

In particular, $(\hat{P}_{t,s}\delta_x)(dy) = p_{t,s}(x,dy)$ and $\hat{P}_{t,s}\big(g\cdot m_t\big) = \big(P^*_{t,s}g\big)\cdot m_s$.

Characterization of super-Ricci flows

In [59], Sturm introduced and analyzed the notion of *super-Ricci flows* for time-dependent metric measure $(X, d_t, m_t)_{t\in I}$. The defining property of the latter is the so-called dynamic convexity of the Boltzmann entropy $S : I \times \mathcal{P} \to (-\infty, \infty]$ with

$$S_t(\mu) = \int u \log u \, dm_t \qquad \text{if } \mu = u\, m_t$$

and $S_t(\mu) = \infty$ if $\mu \not\ll m_t$. Here $\mathcal{P} = \mathcal{P}(X)$ will denote the space of probability measures on X, equipped with time-dependent Kantorovich-Wasserstein distances W_t induced by d_t, $t \in I$.

The main goal of this chapter is to characterize super-Ricci flows in terms of the heat flow (acting on functions, forwards in time) and of the dual heat flow (acting on probability measures, backwards in time). Our first result in this direction is a complete analogue to the characterization of synthetic lower Ricci bounds in the sense of Lott-Sturm-Villani for 'static' metric measure spaces derived by Ambrosio, Gigli, Savaré [7].

Theorem 2.3. *The following assertions are equivalent:*

(I) *For a.e.* $t \in (0, T)$ *and every* W_t*-geodesic* $(\mu^a)_{a\in[0,1]}$ *in* $\mathcal{P}$ *with* $\mu^0, \mu^1 \in$ *Dom*(S)

$$\partial_a^+ S_t(\mu^a)\big|_{a=1-} - \partial_a^- S_t(\mu^a)\big|_{a=0+} \geq -\frac{1}{2}\partial_t^- W_{t-}^2(\mu^0, \mu^1) \tag{13}$$

('dynamic convexity').

(II) *For all* $0 \leq s < t \leq T$ *and* $\mu, \nu \in \mathcal{P}$

$$W_s(\hat{P}_{t,s}\mu, \hat{P}_{t,s}\nu) \leq W_t(\mu, \nu) \tag{14}$$

('transport estimate').

(III) *For all* $u \in$ *Dom*$(\mathcal{E})$ *and all* $0 < s < t < T$

$$\big|\nabla_t (P_{t,s}u)\big|^2 \leq P_{t,s}\big(|\nabla_s u|^2\big) \tag{15}$$

('gradient estimate').

(IV) *For all* $0 < s < t < T$ *and for all* $u_s, g_t \in \mathcal{F}$ *with* $g_t \geq 0$, $g_t \in L^\infty$, $u_s \in \mathrm{Lip}(X)$ *and for a.e.* $r \in (s, t)$

$$\mathbf{\Gamma}_{2,r}(u_r)(g_r) \geq \frac{1}{2}\int \overset{\bullet}{\Gamma}_r (u_r) g_r dm_r \tag{16}$$

('dynamic Bochner's inequality' or 'dynamic Bakry-Émery condition') where $u_r = P_{r,s}u_s$ *and* $g_r = P^*_{t,r}g_t$. *Moreover, the following regularity assumption is satisfied:*

$$u_r \in \mathrm{Lip}(X) \text{ for all } r \in (s,t) \text{ with } \sup_{r,x} \mathrm{lip}_r u_r(x) < \infty. \tag{17}$$

Here and in the sequel

$$\mathbf{\Gamma}_{2,r}(u_r)(g_r) := \int \Big[\frac{1}{2}\Gamma_r(u_r)\Delta_r g_r + (\Delta_r u_r)^2 g_r + \Gamma_r(u_r, g_r)\Delta_r u_r\Big] dm_r$$

denotes the distribution valued Γ_2-operator (at time r) applied to u_r and tested against g_r and

$$\overset{\bullet}{\Gamma}_r (u_r) := \text{w-}\lim_{\delta\to 0} \frac{1}{\delta}\Big(\Gamma_{r+\delta}(u_r) - \Gamma_r(u_r)\Big)$$

denotes any subsequential weak limit of $\frac{1}{2\delta}\big(\Gamma_{r+\delta} - \Gamma_{r-\delta}\big)(u_r)$ in $L^2((s,t)\times X)$.

EVI characterization of the dual heat flow

It turns out that the dual heat flow (acting on probability measures, backward in time) is the backward gradient flow for the Boltzmann entropy – in a very precise, strong sense – and it is the only one with this property.

Theorem 2.4. *Each of the assertions of the previous theorem implies that the dual heat flow $t \mapsto \mu_t = \hat{P}_{\tau,t}\mu$ is the unique* dynamical backward EVI-gradient flow *for the Boltzmann entropy S in the following sense:*
For every $\mu \in Dom(S)$ and every $\tau < T$ the absolutely continuous curve $t \mapsto \mu_t$ satisfies

$$\frac{1}{2}\partial_s^- W^2_{s,t}(\mu_s, \sigma)\big|_{s=t-} \geq S_t(\mu_t) - S_t(\sigma)$$

for all $\sigma \in Dom(S)$ and all $t \leq \tau$.

Characterization of super-N-Ricci flows

For static metric measure spaces, it turned out that many powerful applications of synthetic lower bounds on the Ricci curvature are available only in combination with some synthetic upper bound on the dimension. This lead to the so-called curvature-dimension condition $\mathrm{CD}(K,N)$. In a similar spirit, in [59] the notion of super Ricci flows for time-dependent metric measure spaces was tightened up towards N-super Ricci flows.

We aim to characterize super-N-Ricci flows in terms of the heat flow, the dual heat flow, and the time-dependent Bochner inequality. Our main result provides a complete characterization, analogous to the proof of the equivalence of the curvature-dimension condition of Lott-Stum-Villani and the Bochner inequality of Bakry-Émery for 'static' metric measure spaces derived by Erbar, Kuwada, and Sturm in [24].

Theorem 2.5. *For each N the following are equivalent:*

($\mathbf{I}_N$) *For a.e. $t \in (0,T)$ and every W_t-geodesic $(\mu^a)_{a\in[0,1]}$ in $\mathcal{P}$ with $\mu^0, \mu^1 \in Dom(S)$*

$$\partial_a^+ S_t(\mu^a)\big|_{a=1-} - \partial_a^- S_t(\mu^a)\big|_{a=0+} \geq -\frac{1}{2}\partial_t^- W^2_{t-}(\mu^0,\mu^1) + \frac{1}{N}\big|S_t(\mu^0) - S_t(\mu^1)\big|^2. \tag{18}$$

(II$_N$) *For all* $0 \le s < t \le T$ *and* $\mu, \nu \in \mathcal{P}$

$$W_s^2(\hat{P}_{t,s}\mu, \hat{P}_{t,s}\nu) \le W_t^2(\mu,\nu) - \frac{2}{N}\int_s^t \Big[S_r(\hat{P}_{t,r}\mu) - S_r(\hat{P}_{t,r}\nu)\Big]^2 dr. \quad (19)$$

(III$_N$) *For all* $u \in Dom(\mathcal{E})$ *and all* $0 < s < t < T$

$$\big|\nabla_t(P_{t,s}u)\big|^2 \le P_{t,s}\big(|\nabla_s(u)|^2\big) - \frac{2}{N}\int_s^t \Big(P_{t,r}\Delta_r P_{r,s}u\Big)^2 dr. \quad (20)$$

(IV$_N$) *For all* $0 < s < t < T$ *and for all* $u_s, g_t \in \mathcal{F}$ *with* $g_t \ge 0$, $g_t \in L^\infty$, $u_s \in \mathrm{Lip}(X)$ *the regularity assumption* (17) *is satisfied and for a.e.* $r \in (s,t)$

$$\mathbf{\Gamma}_{2,r}(u_r)(g_r) \ge \frac{1}{2}\int \overset{\bullet}{\Gamma}_r(u_r)g_r dm_r + \frac{1}{N}\Big(\int \Delta_r u_r g_r dm_r\Big)^2 \quad (21)$$

('dynamic Bochner inequality' or 'dynamic Bakry-Émery condition') where $u_r = P_{r,s}u_s$ *and* $g_r = P^*_{t,r}g_t$.

Remark 2.6. *a. In* **(I$_N$)**, *the requested property for a.e. t will imply that it holds true for all* $t \in (0,T)$.

b. The transport estimate **(II$_N$)** *implies the 'stronger' property*

$$W_s^2(\hat{P}_{t,s}\mu, \hat{P}_{t,s}\nu) \le W_t^2(\mu,\nu) - \frac{2}{N}\int_s^t \int_0^1 \Big(\partial_a S_r(\rho_r^a)\Big)^2 da\, dr$$

where $(\rho_r^a)_a$ *denotes the* W_r*-geodesic connecting* $\hat{P}_{r,t}\mu$ *and* $\hat{P}_{r,t}\nu$.

The strategy for the proof is as follows. In Chapter 2.7, we present the implications **(I$_N$)** $\Longrightarrow$ **(II$_N$)** and **(III$_N$)** $\Longrightarrow$ **(II$_N$)** as well as the converse of the latter in the case $N = \infty$. Chapter 2.8 is devoted to the proof of the implication **(III$_N$)** $\Longleftrightarrow$ **(IV$_N$)** as well as to the proof of the equivalence **(II$_N$)**$\Longrightarrow$ **(IV$_N$)**.

In Chapter 2.9 we prove that **(III)** implies the dynamic **EVI** ('evolution variation inequality'). More precisely, we derive two versions, the dynamic EVI$^-$ and a relaxed form of the dynamic EVI$^+$. The combination of these two versions implies that the dual heat flow is the **unique EVI** flow for the Boltzmann entropy.

The latter will be proven in a more abstract context in Chapter 2.10 which is devoted to the study of dynamical EVI-flows in a general framework. Here in particular, it will also be shown that **(III$_N$) & EVI$^-$** $\Longrightarrow$ **(I$_N$)**. □

Remark 2.7. *Note that the regularity assumption* (17) *in our formulation of the dynamic Bochner inequality is not really a restriction. Indeed, such an estimate with* $C = 2(K+L)$ *will always follow from the log-Lipschitz bound* (12) *and the* $RCD(-K,\infty)$*-condition for the static mm-spaces* (X, d_t, m_t).

Let us give two motivating examples of super-Ricci flows as defined in [59, Definition 2.4]. In the first example we construct a super-Ricci flow on the spherical cone by means of a Ricci flow on the punctured spherical cone, while in the second example we only draw a rough sketch.

Example. *Consider the product* $M \times [0,\pi]$*, where* $M = S^2(1/\sqrt{3}) \times S^2(1/\sqrt{3})$ *and* $S^2(r)$ *denotes the 2-dimensional sphere with radius* r*. We contract each of the fibers* $\mathcal{S} := M \times \{0\}$ *and* $\mathcal{N} := M \times \{\pi\}$ *to a point, the* south *and the* north pole*, respectively. The resulting space is called* spherical cone *and is denoted by* $\Sigma(M)$*. We endow* $\Sigma(M)$ *with*

- *metric* $d_{\Sigma(M)}$ *defined by*

$$\cos(d_{\Sigma(M)}((x,s),(x',s'))) := \cos s \cos s' + \sin s \sin s' \cos(d(x,x') \wedge \pi),$$

  where $(x,s),(x',s') \in M \times [0,\pi]$ *and* d *is the metric of* M*,*

- *measure* $d\hat{m}(x,s) := dm(x) \otimes (\sin^4 s\, ds)$*, where* m *is the volume of* M*.*

Since M *is a* $RCD^*(3,4)$ *space, the cone of it is a* $RCD^*(4,5)$ *space. The punctured cone* $\Sigma_0 := \Sigma(M) \setminus \{\mathcal{S},\mathcal{N}\}$ *is an incomplete 5-dimensional Riemannian manifold. Let* g_0 *denote the metric tensor of* Σ_0*. The curvature of the punctured cone can be calculated explicitly and is given by* $\mathrm{Ric}(g_0) = 4g_0$*. Then*

$$g(t) := (1-8t)g_0.$$

defines a solution to the Ricci flow $\mathrm{Ric}(g_t) = -\frac{1}{2}\partial_t g_t$ *with* $g(0) = g_0$*, which collapses to a point at time* $T = \frac{1}{8}$*.*

Let $I = (0,T')$ *with* $T' < T$*. We claim that the associated metric measure space* $(\Sigma(M), d_{\Sigma(M)}(t), \hat{m}_t)_{t\in I}$ *is a super-Ricci flow. Fix* $t \in I$ *and let* $\mu_0,\mu_1 \in$ *Dom*(S_t) *on* $\Sigma(M)$ *be given. Let* $(\mu_a)_{a\in[0,1]}$ *be a* W_t*-geodesic connecting* μ_0,μ_1*. Then,* $\mu_a = (e_a)_*\nu$*, where* ν *is an optimal path measure, i.e. a probability measure on the* d_t*-geodesics* $\Gamma(\Sigma(M))$ *of* $\Sigma(M)$ *such that* $(e_0,e_1)_*\nu$ *is an optimal coupling of* $(e_0)_*\nu = \mu_0, (e_1)_*\nu = \mu_1$*, where* $e_a\colon \Gamma(\Sigma(M)) \to \Sigma(M)$ *denotes the evaluation map. According to Theorem 3.3 in [11] every optimal path measure* ν *will give no mass to* d_t*-geodesics through the poles. Hence we can omit the* d_t*-geodesics through the poles without changing the* W_t*-geodesics. Since the punctured cone* $(\Sigma_0, g_t)_{t\in I}$ *is a Ricci flow, and in particular a super-Ricci flow in the sense of Definition 2.4 in [59], the metric measure space* $(\Sigma(M), d_{\Sigma(M)}(t), \hat{m}_t)_{t\in I}$ *is a super-Ricci flow as well.*

Let us emphasize that for each $t \in [0,1/8)$ *the sectional curvature of the punctured spherical cone* Σ_0 *is neither bounded from below nor from above. Indeed, for* $x,y \in S^2(1/\sqrt{3})$ *and* $0 < r < \pi$ *an orthonormal basis of the tangent space* $T_{(x,y,r)}\Sigma_0$ *is given by* $\{\hat{u}_1,\hat{u}_2,\hat{v}_1,\hat{v}_2,\hat{w}\}$ *where*

$$\hat{u}_i = \frac{1}{\sin r}(u_i,0,0), \quad \hat{v}_i = \frac{1}{\sin r}(0,v_i,0), \quad \hat{w} = (0,0,1)$$

and u_1,u_2 *is an orthonormal basis of* $T_x(S^2(1/\sqrt{3}))$ *and* v_1,v_2 *is an orthonormal basis of* $T_y(S^2(1/\sqrt{3}))$*. Then for the sectional curvature we find*

$$\mathrm{Sec}_{(x,y,r)}(\hat{u}_1,\hat{u}_2) = \frac{3-\cos^2 r}{\sin^2 r}, \quad \mathrm{Sec}_{(x,y,r)}(\hat{u}_1,\hat{v}_1) = -\frac{\cos^2 r}{\sin^2 r}$$
$$\mathrm{Sec}_{(x,y,r)}(\hat{u}_1,\hat{v}_2) = -\frac{\cos^2 r}{\sin^2 r}, \quad \mathrm{Sec}_{(x,y,r)}(\hat{u}_1,\hat{w}) = 1,$$

and analogously if we replace $\hat{u}_1$ *by the vectors* $\hat{u}_2,\hat{v}_1,\hat{v}_2$*. This implies in particular that* $\mathrm{Ric}_{(x,y,r)}(\xi,\xi) = 4$*, but for* $r \to 0$ *and* $r \to \pi$*,* $\mathrm{Sec}_{(x,y,r)}(\hat{u}_1,\hat{u}_2) \to +\infty$ *and* $\mathrm{Sec}_{(x,y,r)}(\hat{u}_1,\hat{v}_i) \to -\infty$*.*

Example. *Consider a surface of revolution with piecewise constant negative curvature* $\mathrm{Ric} = -K$ *for some* $K > 0$ *depicted in Figure 1. Under the evolution of a Ricci flow the curvature of the surface where* $\mathrm{Ric} = -K$ *will increase, while the curvature of the "edges" (*$\mathrm{Ric} = +\infty$*) will decrease. In this sense the region of negative curvature will inflate, while the edges will smooth out. Under the evolution of a super-Ricci flow the surface inflates as well but it may keep the edges.*

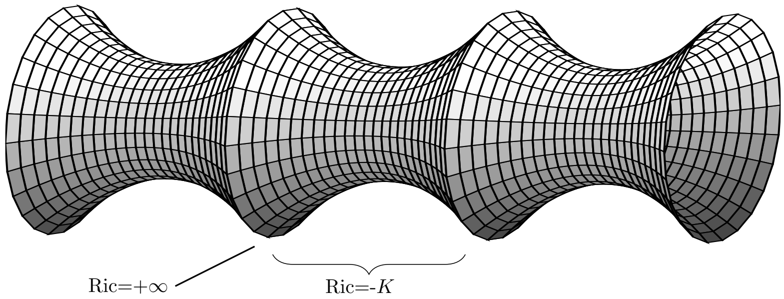

Figure 1: Surface of revolution of a piecewise hyperbolic space

Finally, let us briefly comment on the a priori assumption that each of the static spaces satisfies a Riemannian curvature-dimension condition.

Lemma 2.8. *Given a time-dependent mm-space* $(X, d_t, m_t)_{t\in I}$ *which satisfies all the assumptions mentioned in the beginning of this chapter but no Riemannian curvature-dimension condition is requested. Instead of that, each static mm-space* (X, d_t, m_t) *is merely assumed to be infinitesimally Hilbertian and* S_t *is requested to be absolutely continuous along* W_t*-geodesics.*

Then assertion $(\mathbf{I}_N)$ *of the Main Theorem 2.5 implies that for a.e.* $t \in I$ *the static space*

$$(X, d_t, m_t) \quad \textit{satisfies a } RCD^*(-L, N) \textit{ condition.}$$

Proof. $(\mathbf{I}_N)$ together with the log-Lipschitz bound (12) implies that along all W_t-geodesics

$$\partial_a^+ S_t(\mu^a)\big|_{a=1-} - \partial_a^- S_t(\mu^a)\big|_{a=0+} \quad \geq \quad -L \cdot W_t^2(\mu^0, \mu^1) + \frac{1}{N}\big|S_t(\mu^0) - S_t(\mu^1)\big|^2.$$

In combination with the absolute continuity of $a \mapsto S_t(\mu^a)$ this yields the $\mathrm{RCD}^*(-L, N)$-condition, cf. [59]. $\square$

Preliminary remarks.

We use ∂_t as a short hand notation for $\frac{d}{dt}$. Moreover, we put $\partial_t^+ u(t) = \limsup_{s\to t} \frac{1}{t-s}(u(t) - u(s))$ and $\partial_t^- u(t) = \liminf_{s\to t} \frac{1}{t-s}(u(t) - u(s))$.

In the sequel, r, s, t always denote 'time' parameters whereas a, b denote 'curve' parameters.

2.2 The Heat Equation for Time-dependent Dirichlet Forms

2.3 The Heat Equation

Let us choose here a setting which is slightly more general than for the rest of the chapter. We assume that we are given a Polish space X and a σ-finite reference measure $m_\diamond$ on it which is assumed to have full topological support. Moreover, we assume that we are given a strongly local Dirichlet form $\mathcal{E}_\diamond$ with domain $\mathcal{F} = Dom(\mathcal{E}_\diamond)$ on $\mathcal{H} = L^2(X, m_\diamond)$ and with square field operator $\Gamma_\diamond$ such that $\mathcal{E}_\diamond(u) = \int_X \Gamma_\diamond(u,u)\, dm_\diamond$ for all functions $u \in \mathcal{F}$. These objects will be regarded as *reference measure* and *reference Dirichlet form*, resp., in the subsequent definitions and discussions. The spaces $\mathcal{H}$ and $\mathcal{F}$ will be regarded as a Hilbert space equipped with the scalar products $\int uv dm_\diamond$ and $\mathcal{E}_\diamond(u,v) + \int uv dm_\diamond$, resp. We identify $\mathcal{H}$ with its own dual; the dual of $\mathcal{F}$ is denoted by $\mathcal{F}^*$. Thus we have $\mathcal{F} \subset \mathcal{H} \subset \mathcal{F}^*$ with continuous and dense embeddings.

Let $I \subset \mathbb{R}$ be a bounded open interval, say $I = (0,T)$ for simplicity. In order to deal with time-dependent evolutions we consider for $0 \le s < \tau \le T$ the Hilbert spaces

$$\mathcal{F}_{(s,\tau)} = L^2\big((s,\tau) \to \mathcal{F}\big) \cap H^1\big((s,\tau) \to \mathcal{F}^*\big)$$

equipped with the respective norms $\big(\int_s^\tau \|u_t\|_{\mathcal{F}}^2 + \|\partial_t u_t\|_{\mathcal{F}^*}^2\, dt\big)^{1/2}$. According to [52], Lemma 10.3, the embeddings $\mathcal{F}_{(s,\tau)} \subset \mathcal{C}\big([s,\tau] \to \mathcal{H}\big)$ hold true which guarantee that values at $t = s$ and $t = \tau$ are well defined.

Moreover, assume that we are given a one-parameter family $(m_t)_{t\in(0,T)}$ of measures on X such that $m_t = e^{-f_t} m_\diamond$ for some bounded measurable function f on $I \times X$ with $f_t \in \mathcal{F}$ and $\exists C$ s.t. $\forall t, x$

$$\Gamma_\diamond(f_t)(x) \le C. \tag{22}$$

The basic ingredient will be a 1-parameter family $(\Gamma_t)_{t\in(0,T)}$ of

- symmetric, positive semidefinite bilinear forms Γ_t on $\mathcal{F}$, each of which has the diffusion property

$$\Gamma_t(\Psi(u_1,\ldots,u_k), v) = \sum_{i=1}^k \Psi_i(u_1,\ldots,u_k)\Gamma_t(u_i, v)$$

  $(\forall k \in \mathbb{N}, \forall v, u_1, \ldots, u_k \in \mathcal{F} \cap L^\infty(X, m_\diamond), \forall \Psi \in \mathcal{C}^1(\mathbb{R}^k)$ with $\Psi(0) = 0)$,

- and all of them being uniformly comparable ('uniformly elliptic') w.r.t. the reference form $\Gamma_\diamond$ on $\mathcal{F}$, i.e. $\exists C$ s.t. $\forall t \in (0,T), \forall u \in \mathcal{F}, \forall x \in X$

$$\frac{1}{C}\,\Gamma_\diamond(u)(x) \le \Gamma_t(u)(x) \le C\,\Gamma_\diamond(u)(x). \tag{23}$$

For each $t \in (0,T)$ we define a (strongly local, densely defined, symmetric) Dirichlet form $\mathcal{E}_t$ on $L^2(X, m_t)$ with domain $Dom(\mathcal{E}_t) = \mathcal{F}$ and a self-adjoint, non-positive operator A_t on $L^2(X, m_t)$ with domain $Dom(A_t) \subset \mathcal{F}$ uniquely determined through the relations

$$\int_X \Gamma_t(u,v)\, dm_t = \mathcal{E}_t(u,v) = -\int_X A_t u\, v\, dm_t$$

for $u,v\in\mathcal{F}$. Recall that $u\in Dom(A_t)$ if and only if $u\in\mathcal{F}$ and $\exists C'$ such that $\mathcal{E}_t(u,v)\le C'\cdot\|v\|_{L^2(m_t)}$ for all $v\in\mathcal{F}$.

Definition 2.9. *A function u is called solution to the heat equation*

$$A_t u=\partial_t u\qquad \text{on } (s,\tau)\times X$$

if $u\in\mathcal{F}_{(s,\tau)}$ and if for all $w\in\mathcal{F}_{(s,\tau)}$

$$-\int_s^\tau \mathcal{E}_t(u_t,w_t)dt=\int_s^\tau \langle\partial_t u_t, w_t e^{-f_t}\rangle_{\mathcal{F}^*,\mathcal{F}}\,dt \tag{24}$$

where $\langle\cdot,\cdot\rangle_{\mathcal{F}^*,\mathcal{F}}=\langle\cdot,\cdot\rangle$ denotes the dual pairing. Note that thanks to (22)*, $w\in L^2\big((s,\tau)\to\mathcal{F}\big)$ if and only if $we^{-f}\in L^2\big((s,\tau)\to\mathcal{F}\big)$.*

Since $u_t\in Dom(A_t)$ (and thus $\partial_t u_t\in L^2$) for almost every t by virtue of Theorem 2.20 we may equivalently rewrite the right hand side of the above equation as

$$\int_s^\tau \langle\partial_t u_t, w_t e^{-f_t}\rangle_{\mathcal{F}^*,\mathcal{F}}\,dt=\int_s^\tau\int_X \partial_t u_t\cdot(w_t e^{-f_t})\,dm_\diamond\,dt=\int_s^\tau\int_X \partial_t u_t\cdot w_t\,dm_t\,dt$$

which allows for a more intuitive, alternative formulation of (24) as follows:

$$-\int_s^\tau \mathcal{E}_t(u_t,w_t)dt=\int_s^\tau\int_X \partial_t u_t\cdot w_t\,dm_t\,dt.$$

Theorem 2.10. *For all $0\le s<\tau\le T$ and each $h\in\mathcal{H}$ there exists a unique solution $u\in\mathcal{F}_{(s,\tau)}$ of the heat equation on $(s,\tau)\times X$ with $u_s=h$ (or equivalently with $\lim_{t\searrow s}u_t=h$).*

Proof. For each t the bilinear form $\mathcal{E}_t^\diamond$ on $\mathcal{F}$ is defined by

$$\begin{aligned}\mathcal{E}_t^\diamond(u,v)&=-\int_X A_t u\,v\,dm_\diamond\\&=\int_X \Gamma_t(u,ve^{f_t})e^{-f_t}\,dm_\diamond\\&=\int_X\left[\Gamma_t(u,v)+v\Gamma_t(u,f_t)\right]\,dm_\diamond\end{aligned}$$

for $u,v\in\mathcal{F}$. It immediately follows that $u\in\mathcal{F}_{(s,\tau)}$ is a solution to the heat equation if and only if for all $w\in\mathcal{F}_{(s,\tau)}$

$$-\int_s^\tau \mathcal{E}_t^\diamond(u_t,w_t)dt=\int_s^\tau\int_X \partial_t u_t\cdot w_t\,dm_\diamond\,dt.$$

(Indeed, we simply have to replace the test function w_t by $w_te^{f_t}$.)

Our assumptions on Γ_t and f_t guarantee that $\mathcal{E}_t^\diamond$ for each t is a closed coercive form with domain $\mathcal{F}=Dom(\mathcal{E}_\diamond)$ on $\mathcal{H}=L^2(X,m_\diamond)$, uniformly comparable to $\mathcal{E}_\diamond$. For each t , the operator A_t is a bounded linear operator from $\mathcal{F}$ to $\mathcal{F}^*$.

Indeed,

$$
\begin{aligned}
\|A_t\|_{\mathcal{F},\mathcal{F}^*} &= \sup_{u,v\in\mathcal{F}} \frac{\big|\mathcal{E}_t^\diamond(u,v)\big|}{\|u\|_{\mathcal{F}}^{1/2}\cdot\|v\|_{\mathcal{F}}^{1/2}}\\
&\le \sup_{u,v\in\mathcal{F}} \frac{1}{\|u\|_{\mathcal{F}}^{1/2}\cdot\|v\|_{\mathcal{F}}^{1/2}} \int_X |\Gamma_t(u,v)|\, dm_\diamond\\
&+ \sup_{u,v\in\mathcal{F}} \frac{1}{\|u\|_{\mathcal{F}}^{1/2}\cdot\|v\|_{\mathcal{F}}^{1/2}} \int_X |v\Gamma_t(u,f_t)|\, dm_\diamond\\
&\le C\left(1+\|\Gamma(f_t)\|_\infty^{1/2}\right)
\end{aligned}
$$

if C is chosen such that $|\Gamma_t(u,v)| \le C\cdot\Gamma_\diamond(u)^{1/2}\cdot\Gamma_\diamond(v)^{1/2}$ for all u,v and t. Thus we may apply the general existence result for solutions to time-dependent operator equations $\partial_t u = A_t u$ on a fixed Hilbert space $\mathcal{H}$. For this, we refer to [42], Chapter III, Theorem 4.1 and Remark 4.3, see also [52], Theorem 10.3. (Note, however, that the latter assumes a continuity of $t\mapsto A_t$ in operator norm which is not really necessary.) $\square$

Remark 2.11. *We denote this solution by $u_t(x) = P_{t,s}h(x)$. Then $(P_{t,s})_{0<s\le t<T}$ is a family of bounded linear operators on $\mathcal{H}$ which has the* propagator property

$$P_{t,r} = P_{t,s}\circ P_{s,r}$$

for all $r\le s\le t$. For fixed s and h the function $t\mapsto P_{t,s}h$ is continuous in $\mathcal{H}$ (due to the embedding $\mathcal{F}_{(s,T)}\subset\mathcal{C}\big([s,T]\to\mathcal{H}\big)$). And by construction the function $(t,x)\mapsto P_{t,s}h(x)$ is a solution to the (forward) heat equation $\partial_t u = A_t u$ on $(s,T)\times X$. That is, for all $h\in\mathcal{H}$

$$\partial_t P_{t,s}h = A_t P_{t,s}h. \tag{25}$$

Note that the operator $P_{t,s}:\mathcal{H}\to\mathcal{H}$ in the general time-dependent case is not symmetric – neither with respect to $m_\diamond$ nor with respect to m_t nor with respect to m_s.

2.4 The Adjoint Heat Equation

Definition 2.12. *Given $0\le\sigma<t\le T$, a function v is called solution to the adjoint heat equation*

$$-A_s v + \partial_s f\cdot v = \partial_s v \qquad \text{on } (\sigma,t)\times X$$

if $v\in\mathcal{F}_{(\sigma,t)}$ and if for all $w\in\mathcal{F}_{(\sigma,t)}$

$$\int_\sigma^t \mathcal{E}_s(v_s,w_s)ds + \int_\sigma^t\int_X v_s\cdot w_s\cdot\partial_s f_s\, dm_s\, ds = \int_\sigma^t\int_X \partial_s v_s\cdot w_s\, dm_s\, ds.$$

Theorem 2.13. *Assume* (22) *and*

$$|f_t(x)-f_s(x)|\le L\,|t-s|. \tag{26}$$

(i) Given $0 \leq \sigma < t \leq T$, for each $g \in \mathcal{H}$ there exists a unique solution $v \in \mathcal{F}_{(\sigma,t)}$ of the adjoint heat equation on $(\sigma, t) \times X$ with $v_t = g$.

(ii) This solution can be represented as

$$v_s = P^*_{t,s} g$$

in terms of a family $(P^*_{t,s})_{s \leq t}$ of linear operators on $\mathcal{H}$ satisfying the 'adjoint propagator property'

$$P^*_{t,r} = P^*_{s,r} \circ P^*_{t,s} \qquad (\forall r \leq s \leq t).$$

(iii) The operators $P_{t,s}$ and $P^*_{t,s}$ are in duality w.r.t. each other:

$$\int P_{t,s} h \cdot g \, dm_t = \int h \cdot P^*_{t,s} g \, dm_s \qquad (\forall g, h \in \mathcal{H}).$$

Proof. (i), (ii) The assumption implies that the same arguments used before to prove existence and uniqueness of solutions to the heat equation $\partial_t u = A_t u$ can now be applied to prove existence and uniqueness of solutions to the adjoint heat equation $-\partial_s v = A_s v - (\partial_s f_s) v$.

(iii) Put $u_t = P_{t,s} h$ and $v_s = P^*_{t,s} g$. Then

$$\begin{aligned}
&\int u_t v_t \, dm_t - \int u_s v_s \, dm_s \\
&= \int_s^t \int \partial_r u_r \, v_r \, dm_r \, dr + \int_s^t \int u_r \, \partial_r v_r \, dm_r \, dr - \int_s^t \int u_r \, v_r \, \partial_r f_r \, dm_r \, dr \\
&= \int_s^t \mathcal{E}_r(u_r, v_r) \, dr - \int_s^t \mathcal{E}_r(u_r, v_r) \, dr = 0.
\end{aligned}$$

□

Note, however, that – even under the assumption $m_\diamond(X) < \infty$ – in general constants will not be solutions to the adjoint heat equation. Instead of preserving constants, the adjoint heat flow preserves integrals of nonnegative densities.

Lemma 2.14. *For each fixed t, the operators A_t and $A^*_t : u \mapsto A_t u - \partial_t f_t \cdot u$ on $L^2(X, m_t)$ have the same domains: $Dom(A_t) = Dom(A^*_t)$*

Proof. Recall that $v \in Dom(A^*_t)$ if and only if $v \in Dom(\mathcal{E}_t)$ and if there exists a constant C such that for all $u \in Dom(\mathcal{E}_t)$

$$\mathcal{E}_t(u, v) + \int u \, v \, \partial_t f \, dm_t \leq C \cdot \|u\|_{L^2(m_t)}.$$

Boundedness of $\partial_t f$ implies that this is equivalent to $v \in Dom(A_t)$. □

In contrast to the form domains, the operator domains $Dom(A_t)$ in general will depend on t.

Example 2.15. *Consider* $\mathcal{H} = L^2(\mathbb{R}, dx)$ *with* $m_t(dx) = dx$ *and*

$$\Gamma_t(u)(x) = \big[1 + t \cdot 1_{\mathbb{R}_+}(x)\big] \cdot |u'(x)|^2$$

for $t \in I = (0, 1)$*. Then*

$$Dom(A_t) = \Big\{u \in W^{1,2}(\mathbb{R}) \cap W^{2,2}(\mathbb{R}_-) \cap W^{2,2}(\mathbb{R}_+) : \ u'(0-) = (1+t) \cdot u'(0+)\Big\}.$$

Thus $Dom(A_s) \neq Dom(A_t)$ *for all* $s \neq t$*.*

Proof. Obviously, $u \in Dom(A_t)$ if and only if $u \in W^{1,2}(\mathbb{R})$ and $[1 + t \cdot 1_{\mathbb{R}_+}]u' \in W^{1,2}(\mathbb{R})$. □

A basic quantity for the subsequent considerations will be the time-dependent Boltzmann entropy. Here we put $S_t(v) := \int_X v \cdot \log v \, dm_t$ and consider it as a time-dependent functional on the space of (not necessarily normalized) measurable functions $v : X \to [0, \infty]$.

Proposition 2.16. *(i) For all solutions* $u \geq 0$ *to the heat equation and all* $s < t$

$$S_t(u_t) \leq e^{L(t-s)} \cdot S_s(u_s).$$

(ii) For all solutions $v \geq 0$ *to the adjoint heat equation and all* $s < t$

$$S_s(v_s) \leq S_t(v_t) + L \int_s^t \int_X v_r \, dm_r \, dr.$$

Note that $\int_X v_r dm_r$ *is independent of* r *if* $m_\diamond(X) < \infty$*.*

Proof. In both cases, straightforward calculations yield

$$\begin{aligned} e^{Lt}\partial_t \left[e^{-Lt} \int u_t \log u_t \, dm_t\right] \quad &\leq \quad \int (\log u_t + 1)\partial_t u_t \, dm_t \\ = -\int \Gamma_t(\log u_t) \, u_t \, dm_t \leq 0 \end{aligned}$$

and

$$\begin{aligned} \partial_s \int v_s \log v_s \, dm_s \quad &= \quad \int (\log v_s + 1)\partial_s v_s \, dm_s - \int v_s \log v_s \cdot \partial_s f_s \, dm_s \\ &= \quad \int \Gamma_s(\log v_s) \, v_s \, dm_s + \int v_s \cdot \partial_s f_s \, dm_s \geq -L \int v_s \, dm_s. \end{aligned}$$

□

2.5 Energy Estimates

Throughout this section, assume (22) as well as (26) and in addition

$$|\Gamma_t(u) - \Gamma_s(u)| \leq 2L \cdot \int_s^t \Gamma_r(u) dr \tag{27}$$

for all $u \in \mathcal{F}$ and all $s < t$.

Recall that by definition each solution u to the heat equation on $(s,\tau)\times X$ satisfies $u\in L^2\big((s,\tau)\to\mathcal{F}\big)\cap H^1\big((s,\tau)\to\mathcal{F}^*\big)\subset\mathcal{C}\big((s,\tau)\to\mathcal{H}\big)$ and

$$\int_s^\tau \mathcal{E}_t(u_t)\,dt \le \frac{1}{2}\|u_s\|^2_{L^2(m_s)}. \tag{28}$$

We are now going to prove that these assertions can be improved by one order of (spatial) differentiation. To do so, we first define a self-adjoint, non-positive operator $\tilde{A}_t$ on $L^2(X,m_\diamond)$ by

$$-\int_X \tilde{A}_t u\, v\, dm_\diamond = \tilde{\mathcal{E}}_t(u,v) := \int_X \Gamma_t(u,v)\, dm_\diamond$$

for all $u,v\in\mathcal{F}$. Then $Dom(\tilde{A}_t)=Dom(A_t)$ and

$$\tilde{A}_t u = A_t u + \Gamma_t(u,f_t).$$

Indeed, $-\int A_t u\, v\, dm_\diamond = \int \Gamma_t(u, ve^{f_t})e^{-f_t}dm_\diamond = -\int \tilde{A}_t u\, v\, dm_\diamond + \int \Gamma_t(u,f_t)v\, dm_\diamond$. Next, consider the Hille-Yosida approximation $\tilde{A}^\delta_t := (I-\delta\tilde{A}_t)^{-1}\tilde{A}_t$ of $\tilde{A}_t$ on $L^2(X,m_\diamond)$, put $\tilde{\mathcal{E}}^\delta_t(u,v) := -\int \tilde{A}^\delta_t u\, v\, dm_\diamond$ and recall the well-known fact that $\tilde{\mathcal{E}}^\delta_t(u,u)\nearrow\tilde{\mathcal{E}}_t(u,u)$ for each $u\in\mathcal{F}$ as $\delta\searrow 0$. More generally,

Lemma 2.17. *For all $\alpha,\beta>0$ with $\beta-\alpha\le\frac{1}{2}$: $\mathcal{F}\subset Dom((I-\delta\tilde{A}_t)^{-\alpha}\tilde{A}^\beta_t)$ and for all $u\in\mathcal{F}$:*

$$u\in Dom(\tilde{A}^\beta_t) \quad\Longleftrightarrow\quad \sup_{\delta>0}\Big\|(I-\delta\tilde{A}_t)^{-\alpha}\tilde{A}^\beta_t u\Big\|_{L^2}<\infty$$

with $\Big\|(I-\delta\tilde{A}_t)^{-\alpha}\tilde{A}^\beta_t u\Big\|_{L^2}\nearrow\Big\|\tilde{A}^\beta_t u\Big\|_{L^2}$ for $\delta\searrow 0$.

Proof. For fixed t we apply the spectral theorem to the non-negative self-adjoint operator $-\tilde{A}_t$ on $\mathcal{H}$ which yields the representation $-\tilde{A}_t=\int_0^\infty \lambda\, E_\lambda$ in terms of projection operators. For each continuous semi-bounded $\Phi:\mathbb{R}_+\to\mathbb{R}$

$$Dom\left(\Phi(-\tilde{A}_t)\right) = \left\{u\in\mathcal{H}:\ \int_0^\infty |\Phi(\lambda)|^2 dE_\lambda(u,u)\right\}$$

and $(\Phi(-\tilde{A}_t)u,v)_\mathcal{H}=\int_0^\infty\Phi(\lambda)dE_\lambda(u,v)$. Thus, in particular,

$$\mathcal{F}=\left\{u\in\mathcal{H}:\ \int_0^\infty \lambda dE_\lambda(u,u)\right\}$$

and

$$Dom\left((I-\delta\tilde{A}_t)^{-\alpha}\tilde{A}^\beta_t\right)=\left\{u\in\mathcal{H}:\ \int_0^\infty\left|\frac{\lambda^\beta}{(1+\delta\lambda)^\alpha}\right|^2 dE_\lambda(u,u)\right\}.$$

Moreover, by monotone convergence as $\delta\searrow 0$

$$\begin{aligned}&\Big\|(I-\delta\tilde{A}_t)^{-\alpha}\tilde{A}^\beta_t u\Big\|^2_{L^2}\\ &=\int_0^\infty\left|\frac{\lambda^\beta}{(1+\delta\lambda)^\alpha}\right|^2 dE_\lambda(u,u)\ \nearrow\ \int_0^\infty \lambda^{2\beta}dE_\lambda(u,u)=\Big\|\tilde{A}^\beta_t u\Big\|^2_{L^2}.\end{aligned}$$

□

Lemma 2.18. *For all $\delta > 0$ and all $u, v \in \mathcal{F}$ the map $t \mapsto \tilde{\mathcal{E}}_t^\delta(u,v)$ is absolutely continuous with*

$$\left|\partial_t \tilde{\mathcal{E}}_t^\delta(u,v)\right| \le \frac{L}{2}\left[\tilde{\mathcal{E}}_t(u,u) + \tilde{\mathcal{E}}_t(v,v)\right].$$

Proof. For all δ, u, v as above, put $u_t^\delta = (I - \delta\tilde{A}_t)^{-1}u$ and $v_t^\delta = (I - \delta\tilde{A}_t)^{-1}v$. Then

$$\begin{aligned}
\partial_t \tilde{\mathcal{E}}_t^\delta(u,v) &= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\int\left[(I-\delta\tilde{A}_{t+\epsilon})^{-1}\tilde{A}_{t+\epsilon}u - (I-\delta\tilde{A}_t)^{-1}\tilde{A}_t u\right]\cdot v\,dm_\diamond \\
&= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\int\left[(I-\delta\tilde{A}_{t+\epsilon})^{-1}(\tilde{A}_{t+\epsilon}-\tilde{A}_t)(1-\delta\tilde{A}_t)^{-1}u\right]\cdot v\,dm_\diamond \\
&= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\tilde{\mathcal{E}}_t(u_t^\delta, v_{t+\epsilon}^\delta) - \tilde{\mathcal{E}}_{t+\epsilon}(u_t^\delta, v_{t+\epsilon}^\delta)\right] \\
&\le \frac{L}{2}\lim_{\epsilon\to 0}\left[\tilde{\mathcal{E}}_t(u_t^\delta, u_t^\delta) + \tilde{\mathcal{E}}_{t+\epsilon}(v_{t+\epsilon}^\delta, v_{t+\epsilon}^\delta)\right] \\
&\le \frac{L}{2}\lim_{\epsilon\to 0}\left[\tilde{\mathcal{E}}_t(u,u) + \tilde{\mathcal{E}}_{t+\epsilon}(v,v)\right] = \frac{L}{2}\left[\tilde{\mathcal{E}}_t(u,u) + \tilde{\mathcal{E}}_t(v,v)\right].
\end{aligned}$$

Here we also used the fact that $\tilde{\mathcal{E}}_t(u_t^\delta, u_t^\delta) \nearrow \tilde{\mathcal{E}}_t(u_t, u_t)$ as $\delta \to 0$. □

Lemma 2.19. *There exists a constant C such that for all $0 < s < \tau < T$, for all solutions $u \in \mathcal{F}_{(s,\tau)}$ to the heat equation on $(s,\tau)\times X$ and for all $\delta > 0$*

$$\int_s^\tau \int_X \left|(I-\delta\tilde{A}_t)^{-1/2}\tilde{A}_t u_t\right|^2 dm_\diamond\,dt \le C\cdot\left[\mathcal{E}_s(u_s) + \|u_s\|^2_{L^2(m_s)}\right]. \tag{29}$$

Thus, in particular, if $u_s \in \mathcal{F}$ then $u_t \in Dom(\tilde{A}_t)$ for a.e. $t \in (s,\tau)$ and

$$\int_s^\tau \int_X \left|\tilde{A}_t u_t\right|^2 dm_\diamond\,dt \le C\cdot\left[\mathcal{E}_s(u_s) + \|u_s\|^2_{L^2(m_s)}\right]. \tag{30}$$

Proof. For any $\delta > 0$ and $u \in \mathcal{F}$

$$\begin{aligned}
\tilde{\mathcal{E}}_s(u_s) &\ge \tilde{\mathcal{E}}_s^\delta(u_s) \ge -\int_s^\tau \partial_t\tilde{\mathcal{E}}_t^\delta(u_t)\,dt \ge -2\int_s^\tau \mathcal{E}_t^\delta(u_t, \partial_t u_t)\,dt - o_1 \\
&= 2\int_s^\tau\int_X (I-\delta\tilde{A}_t)^{-1}\tilde{A}_t u\cdot A_t u_t\,dm_\diamond\,dt - o_1 \\
&= 2\int_s^\tau\int_X (I-\delta\tilde{A}_t)^{-1}\tilde{A}_t u\cdot \tilde{A}_t u_t\,dm_\diamond\,dt \\
&\qquad -2\int_s^\tau\int_X (I-\delta\tilde{A}_t)^{-1}\tilde{A}_t u\cdot \Gamma_t(u_t, f_t)\,dm_\diamond\,dt - o_1 \\
&\ge \int_s^\tau\int_X \left|(I-\delta\tilde{A}_t)^{-1/2}\tilde{A}_t u\right|^2 dm_\diamond\,dt - o_1 - o_2.
\end{aligned}$$

Here

$$o_1 \quad := \quad \int_s^\tau \partial_r\mathcal{E}_r^\delta(u_t)\Big|_{r=t}dt \le L\int_s^\tau \mathcal{E}_t(u_t)dt \le \frac{L}{2}\|u_s\|^2_{L^2(m_s)}$$

according to the previous Lemma and

$$\begin{aligned}
o_2 &:= \int_s^\tau\int_X \left|(I-\delta\tilde{A}_t)^{-1/2}\Gamma_t(u_t, f_t)\right|^2 dm_\diamond\,dt \\
&\le C'\int_s^\tau\int_X \Gamma_t(u_t)\,e^{-f_t}\,dm_\diamond\,dt \le \frac{C'}{2}\|u_s\|^2_{L^2(m_s)}
\end{aligned}$$

for $C' = \sup_t \|\Gamma_t(f_t)e^{f_t}\|_{L^\infty(m_t)}$. Moreover, $\tilde{\mathcal{E}}_s(u_s) \le C''\mathcal{E}_s(u_s)$ for $C'' = \sup_t \|e^{f_t}\|_{L^\infty(m_t)}$. Thus the claim follows with $C = \max\{C'', \frac{L+C'}{2}\}$. □

Theorem 2.20. *For all $0 < s < \tau < T$ and for all solutions $u \in \mathcal{F}_{(s,T)}$ to the heat equation*

(i) $u_t \in Dom(A_t)$ for a.e. $t \in (s, \tau)$.

(ii) If the initial condition $u_s \in \mathcal{F}$ then

$$u \in L^2\big((s,\tau) \to Dom(A_.)\big) \cap H^1\big((s,\tau) \to \mathcal{H}\big).$$

More precisely,

$$e^{-3L\tau}\mathcal{E}_\tau(u_\tau) + 2\int_s^\tau e^{-3Lt}\int_X \big|A_t u_t\big|^2\, dm_t\, dt \le e^{-3Ls} \cdot \mathcal{E}_s(u_s). \tag{31}$$

(iii) For all solutions v to the adjoint heat equation on $(\sigma, t) \times X$ and all $s \in (\sigma, t)$

$$\mathcal{E}_s(v_s) + \|v_s\|^2_{L^2(m_s)} \le e^{3L(t-s)} \cdot \Big[\mathcal{E}_t(v_t) + \|v_t\|^2_{L^2(m_t)}\Big].$$

Moreover, $v_s \in Dom(A_s)$ for a.e. $s \in (\sigma, t)$.

Proof. (i): In the case $u_s \in \mathcal{F}$, this follows from the previous Lemma and the fact that $Dom(A_t) = Dom(\tilde{A}_t)$. In the general case $u_s \in \mathcal{H}$, by the very definition of the heat equation it follows that $u_\sigma \in \mathcal{F}$ for a.e. $\sigma \in (s, \tau)$. Applying the previous argument now with σ in the place of s yields that $u_t \in Dom(A_t)$ for a.e. $t \in (\sigma, \tau)$ and thus the latter finally holds for a.e. $t \in (s, \tau)$.

(ii): The log-Lipschitz bound (27) states $|\partial_t \Gamma_t(.)| \le 2L \cdot \Gamma_t(.)$. Together with (26) this implies $\partial_s \mathcal{E}_s(u_t)\big|_{s=t} \le 3L \cdot \mathcal{E}_t(u_t)$. Therefore,

$$e^{3Lt}\partial_t\left[e^{-3Lt}\mathcal{E}_t(u_t)\right] \quad \le \quad \partial_s \mathcal{E}_t(u_s)\big|_{s=t} = -2\int |A_t u_t|^2 dm_t$$

where the last equality is justified according to (i).

(iii) Similarly as we did in the previous Lemmas, we can construct a regularization for the adjoint heat equation which will allow to prove that $v_s \in Dom(A_s)$ for a.e. $s \in (\sigma, t)$. Therefore, we may conclude

$$\begin{aligned} \partial_s \mathcal{E}_s(v_s) \quad &\ge \quad 2\int |A_s v_s|^2 dm_s - 3L \cdot \mathcal{E}_s(v_s) - 2\int A_s v_s \cdot v_s \cdot \partial_s f_s\, dm_s \\ &\ge \quad -3L \cdot \mathcal{E}_s(v_s) - \frac{L}{2}\int v_s^2\, dm_s \end{aligned}$$

and thus

$$\begin{aligned} \partial_s\Big[\mathcal{E}_s(v_s) + \|v_s\|^2_{L^2(m_s)}\Big] \quad &\ge \quad -3L \cdot \mathcal{E}_s(v_s) - \frac{L}{2}\int v_s^2\, dm_s \\ &\qquad +2\int \big[\Gamma_s(v_s) + v_s^2 \cdot \partial_s f_s\big] dm_s - \int v_s^2 \cdot \partial_s f_s\, dm_s \\ &\ge \quad -3L \cdot \Big[\mathcal{E}_s(v_s) + \|v_s\|^2_{L^2(m_s)}\Big]. \end{aligned}$$

□

Remark 2.21. *For fixed s and a.e. $\sigma > s$ the operator $P_{\sigma,s}$ maps $\mathcal{H}$ into $Dom(\mathcal{E})$ and then for a.e. $t > \sigma$ the operator $P_{t,\sigma}$ maps $Dom(\mathcal{E})$ into $Dom(A_t)$. Thus by composition, for a.e. $t > s$ the operator $P_{t,s}$ maps $\mathcal{H}$ into $Dom(A_t)$.*

A simple restatement of the assertions of the subsequent Proposition 2.22 will yield that for all $s \le t$ and all $h \in \mathcal{H}$

- $0 \le h \le 1 \quad \Rightarrow \quad 0 \le P_{t,s}h \le 1$
- $P_{t,s}1 = 1$ provided $m_\diamond(X) < \infty$
- $\big(P_{t,s}h\big)^2 \le P_{t,s}\big(h^2\big)$.

Proposition 2.22. *The following holds true.*

(i) For all solutions u to the heat equation on $(s,\tau) \times X$ and all $t > s$

$$u_s \ge 0 \text{ a.e. on } X \qquad \Longrightarrow \qquad u_t \ge 0 \text{ a.e. on } X.$$

More generally, for any $M \ge 0$

$$u_s \le M \text{ a.e. on } X \qquad \Longrightarrow \qquad u_t \le M \text{ a.e. on } X.$$

If $m_\diamond(X) < \infty$ then this implication holds for all $M \in \mathbb{R}$.

(ii) For all solutions v to the adjoint heat equation on $(\sigma, t) \times X$ and all $s < t$

$$v_t \ge 0 \text{ a.e. on } X \qquad \Longrightarrow \qquad v_s \ge 0 \text{ a.e. on } X.$$

More generally, for any $M \ge 0$

$$v_t \le M \text{ a.e. on } X \qquad \Longrightarrow \qquad v_s \le e^{L(t-s)}M \text{ a.e. on } X.$$

If $m_\diamond(X) < \infty$ then this implication holds for all $M \in \mathbb{R}$.

(iii) For all solutions u to the heat equation on $(s,\tau) \times X$, all $t > s$ and all $p \in [1,\infty]$

$$\|u_t\|_{L^p(m_t)} \le e^{L/p\cdot(t-s)} \cdot \|u_s\|_{L^p(m_s)}.$$

In particular, $\int u_t\, dm_t \le e^{L(t-s)} \int u_s\, dm_s$ for nonnegative solutions.

(iv) For all solutions u, g to the heat equation on $(s,\tau) \times X$ and all $t > s$

$$u_s^2 \le g_s \text{ a.e. on } X \qquad \Longrightarrow \qquad u_t^2 \le g_t \text{ a.e.on } X.$$

Proof. (i) Assume that u solves the heat equation. Put $w = (u - M)_+$. Then for each t, strong locality of the Dirichlet form $\mathcal{E}_t$ implies

$$\mathcal{E}_t\big(u_t, (u_t - M)_+\big) = \mathcal{E}_t\big((u_t - M)_+, (u_t - M)_+\big).$$

The chain rule applied to $\Phi(x) = (x)_+$ implies that a.e on $(s,T) \times X$

$$\partial_t u_t \cdot (u_t - M)_+ = \partial_t (u_t - M)_+ \cdot (u_t - M)_+.$$

Therefore, for a.e. t

$$\begin{aligned}
0 &\leq \mathcal{E}_t\big((u_t - M)_+, (u_t - M)_+\big) = \mathcal{E}_t\big(u_t, (u_t - M)_+\big) \\
&= -\int \partial_t u_t, (u_t - M)_+ e^{-f_t}\, dm_\diamond = -\int \partial_t (u_t - M)_+ (u_t - M)_+ e^{-f_t}\, dm_\diamond \\
&\leq -\frac{1}{2} e^{Lt} \cdot \partial_t \left[e^{-Lt} \int_X (u_t - M)_+^2 dm_t \right],
\end{aligned}$$

where we used (26) in the last inequality. Thus $u_s \leq M$ will imply $u_t \leq M$ for all $t > s$.

In the case, $m_\diamond(X) < \infty$, the constants will be in $\mathcal{H}$ and solve the heat equation. Thus the previous argument can also be applied to $u \pm M$ which yields the claim.

(ii) Assume that v solves the adjoint heat equation. Then with a similar calculation as before we obtain for a.e. s

$$\begin{aligned}
&\frac{1}{2}\partial_s \int (v_s - e^{L(t-s)}M)_+^2\, dm_s \\
=& \int (v_s - e^{L(t-s)}M)_+ \partial_s (v_s - e^{L(t-s)}M)_+\, dm_s - \frac{1}{2}\int (v_s - e^{L(t-s)}M)_+^2 \partial_s f_s\, dm_s \\
=& \int (v_s - e^{L(t-s)}M)_+ (\partial_s v_s + Le^{L(t-s)}M)_+\, dm_s - \frac{1}{2}\int (v_s - e^{L(t-s)}M)_+^2 \partial_s f_s\, dm_s \\
=& \mathcal{E}_s(v_s, (v_s - e^{L(t-s)}M)_+) + \int v_s (v_s - e^{L(t-s)}M)_+ \partial_s f_s\, dm_s \\
&+ \int (v_s - e^{L(t-s)}M)_+ (Le^{L(t-s)}M)_+\, dm_s - \frac{1}{2}\int (v_s - e^{L(t-s)}M)_+^2 \partial_s f_s\, dm_s \\
\geq& -\frac{3}{2} L \int (v_s - e^{L(t-s)}M)_+^2\, dm_s.
\end{aligned}$$

Applying Gronwall's inequality yields

$$\int (v_s - e^{L(t-s)}M)_+^2\, dm_s \leq e^{3L(t-s)} \int (v_t - M)_+^2\, dm_t,$$

which proves the claim.

(iii) Assume $p \in (1, \infty)$. (The case $p = \infty$ follows from (i), and the case $p = 1$ follows from (ii) by duality.) Then, by the previous arguments the linear operator

$$P_{t,s} \colon L^1(m_s) + L^\infty(m_s) \to L^1(m_t) + L^\infty(m_t)$$

maps $L^1(m_s)$ boundedly into $L^1(m_t)$ and $L^\infty(m_s)$ boundedly into $L^\infty(m_t)$. Then, by the Riesz-Thorin interpolation theorem $P_{t,s}$ maps $L^p(m_s)$ boundedly into $L^p(m_t)$ with quantitative estimate

$$||P_{t,s}u||_{L^p(m_t)} \leq e^{L(t-s)/p} ||u||_{L^p(m_s)}.$$

(iv) Choose $w = (u^2 - g)_+$. Then, again by the chain rule and since u and

g are solutions to the heat equation, we find for a.e. t

$$\begin{aligned}\frac{1}{2}e^{Lt}\cdot\partial_t\left[e^{-Lt}\int_X w_t^2 dm_t\right] &\leq \int \partial_t(u_t^2-g_t)w_t\,dm_t \\ &= \int \partial_t u_t(2u_tw_t)\,dm_t - \int \partial_t g_t w_t\,dm_t \\ &= -\mathcal{E}_t(u_t,2u_tw_t)+\mathcal{E}_t(g_t,w_t) \\ &= -\mathcal{E}_t(u_t^2-g_t,w_t)-2\int_X \Gamma_t(u_t,u_t)w_t\,dm_t \\ &= -\mathcal{E}_t(w_t,w_t)-2\int_X \Gamma_t(u_t,u_t)w_t\,dm_t \leq 0,\end{aligned}$$

where we applied the strong locality in the last equation. Thus

$$\int w_t^2 dm_t \leq e^{L(t-s)}\int w_s^2 dm_s$$

for all $t>s$. This proves the claim.

□

As a direct consequence we obtain the following corollary.

Corollary 2.23. *For all $s<t$*

(i) $\|P_{t,s}\|_{L^\infty(m_s)\to L^\infty(m_t)}\leq 1, \qquad \|P^*_{t,s}\|_{L^1(m_t)\to L^1(m_s)}\leq 1,$

(ii) $\|P_{t,s}\|_{L^1(m_s)\to L^1(m_t)}\leq e^{L(t-s)}, \qquad \|P^*_{t,s}\|_{L^\infty(m_t)\to L^\infty(m_s)}\leq e^{L(t-s)},$

(iii) $\|P_{t,s}\|_{L^2(m_s)\to L^2(m_t)}\leq e^{L(t-s)/2}, \qquad \|P^*_{t,s}\|_{L^2(m_t)\to L^2(m_s)}\leq e^{L(t-s)/2}.$

The next result yields that the heat flow is a dynamic EVI$(-L/2,\infty)$-flow for $\frac{1}{2}$ times the Dirichlet energy $\frac{1}{2}\mathcal{E}_t$ on $L^2(X,m_t)$. For the definition of dynamic EVI-flows we refer to Section 2.10.

Theorem 2.24. *(i) Then the heat flow is a* dynamic forward EVI$(-L/2,\infty)$-flow for $\frac{1}{2}\times$ the Dirichlet energy *on $L^2(X,m_t)_{t\in I}$, see section 2.10. More precisely, for all solutions $(u_t)_{t\in(s,\tau)}$ to the heat equation, for all $\tau\leq T$ and all $w\in Dom(\mathcal{E})$*

$$-\frac{1}{2}\partial_s^+\big\|u_s-w\big\|_{s,t}^2\Big|_{s=t}+\frac{L}{4}\cdot\big\|u_t-w\big\|_t^2 \geq \frac{1}{2}\mathcal{E}_t(u_t)-\frac{1}{2}\mathcal{E}_t(w) \tag{32}$$

where $\|.\|_{s,t}$ is defined according to Definition 2.71 with $d_t(v,w)=\big\|v-w\big\|_t=(\int|v-w|^2dm_t)^{1/2}$.

(ii) The heat flow is uniquely characterized by this property. For all $t>s$ and all solutions to the heat equation $\|u_t\|_t\leq e^{L(t-s)/2}\|u_s\|_s$.

Proof. (i) Assumption (26) implies $\partial_t\big\|v\big\|_t^2\leq L\big\|v\big\|_t^2$ as well as (following the argumentation from Proposition 2.72)

$$\partial_s\big\|v\big\|_{s,t}^2\big|_{s=t}\leq\frac{L}{2}\big\|v\big\|_t^2$$

for all v and t. Therefore, we can estimate

$$\begin{aligned}\frac{1}{2}\partial_s^+\left\|u_s - w\right\|_{s,t}^2\Big|_{s=t} &\leq \limsup_{s\to t}\frac{1}{2(s-t)}\Big(\left\|u_s - w\right\|_t^2 - \left\|u_t - w\right\|_t^2\Big)\\ &\quad + \limsup_{s\to t}\frac{1}{2(s-t)}\Big(\left\|u_s - w\right\|_{s,t}^2 - \left\|u_s - w\right\|_t^2\Big)\\ &\leq \langle u_t - w, \partial_t u_t\rangle_t + \frac{L}{4}\left\|u_t - w\right\|_t^2\\ &= -\mathcal{E}_t(u,u) + \mathcal{E}_t(w,u) + \frac{L}{4}\left\|u_t - w\right\|_t^2\\ &\leq -\frac{1}{2}\mathcal{E}_t(u,u) + \frac{1}{2}\mathcal{E}_t(w,w) + \frac{L}{4}\left\|u_t - w\right\|_t^2.\end{aligned}$$

(ii) Uniqueness and the growth estimate immediately follow from the EVI-property. Indeed, the distance $\left\|.\right\|_t$ and the function $\mathcal{E}$ on the time-dependent geodesic space $L^2(X, m_t)_{t\in I}$ satisfy all assumptions mentioned in Section 2.10 on EVI-flows. In particular, the distance is log-Lipschitz: $\partial_t\left\|v\right\|_t^2 \leq L\left\|v\right\|_t^2$ and the energy satisfies the growth bound $\mathcal{E}_s \leq C_0\,\mathcal{E}_t$. □

The next lemma states semicontinuity of the heat flow and the adjoint heat flow with respect to the seminorm $\sqrt{\mathcal{E}}$.

Lemma 2.25. *Let $u, g \in Dom(\mathcal{E})$, $0 < r \leq t < T$. Then*

$$\lim_{s\nearrow t} P_{t,s}^* g = g \quad \text{in } (Dom(\mathcal{E}), \sqrt{\mathcal{E}}),$$
$$\lim_{s\searrow r} P_{s,r} u = u \quad \text{in } (Dom(\mathcal{E}), \sqrt{\mathcal{E}}).$$

Proof. Since $P_{t,s}^* g \to g$ in $L^2(X)$ and the Dirichlet energy is lower semicontinuous we have

$$\mathcal{E}_t(g) \leq \liminf_{s\nearrow t}\mathcal{E}_t(P_{t,s}^* g).$$

On the other hand from Theorem 2.20(iii)

$$\mathcal{E}_s(P_{t,s}^* g) + ||P_{t,s}^* g||_{L^2(m_s)} \leq e^{L(t-s)}(\mathcal{E}_t(g) + ||g||_{L^2(m_t)}),$$

for every $s < t$. Hence, again since $P_{t,s}^* g \to u$ in $L^2(X)$,

$$\begin{aligned}\mathcal{E}_t(g) &\geq \limsup_{s\nearrow t} e^{-L(t-s)}(\mathcal{E}_s(P_{t,s}^* g) + ||P_{t,s}^* g||_{L^2(m_s)}) - ||g||_{L^2(m_t)}\\ &\geq \limsup_{s\nearrow t}\mathcal{E}_s(P_{t,s}^* g) = \limsup_{s\nearrow t}\mathcal{E}_t(P_{t,s}^* g),\end{aligned}$$

where the last identity follows from the Lipschitz property of the metrics and the logarithmic densities. Then, since $\mathcal{E}_t$ is a bilinear form, the parallelogram identity yields

$$\begin{aligned}\limsup_{s\nearrow t}\mathcal{E}_t(P_{t,s}^* g - g) &= \limsup_{s\nearrow t}(2\mathcal{E}_t(g) + 2\mathcal{E}_t(P_{t,s}^* g) - \mathcal{E}_t(u + P_{t,s}^* g))\\ &\leq 4\mathcal{E}_t(g) - \liminf_{s\nearrow t}\mathcal{E}_t(g + P_{t,s}^* g)) \leq 4\mathcal{E}_t(g) - \mathcal{E}_t(2g)\\ &= 0,\end{aligned}$$

where the last inequality is a consequence of the lower semicontinuity of $\mathcal{E}_t$.

The second assertion follows along the same lines replacing Theorem 2.20(iii) by Theorem 2.20(ii). $\square$

2.5.1 The Commutator Lemma

In the static case, generator and semigroup commute. In the dynamic case, this is no longer true. However, we can estimate the error

$$\left| \int_X \left[A_t(P_{t,s}u) - P_{t,s}(A_s u) \right] v \, dm_t \right|.$$

To guarantee well-definedness of all the expressions, we avoid 'Laplacians' and use 'gradients' instead.

Lemma 2.26. *For all $\sigma < \tau$, all solutions $u \in \mathcal{F}_{(\sigma,\tau)}$ to the heat equation, and all solutions $v \in \mathcal{F}_{(\sigma,\tau)}$ to the adjoint heat equation*

$$|\mathcal{E}_t(u_t, v_t) - \mathcal{E}_s(u_s, v_s)| \le C(u_s, v_t) \cdot |t-s|^{1/2} \tag{33}$$

for a.e. $s, t \in (\sigma, \tau)$ with $s < t$ where

$$C(u_s, v_t) = C \cdot \Big[\mathcal{E}_s(u_s) + \mathcal{E}_t(v_t) + \|v_t\|^2_{L^2(m_t)} \Big] \tag{34}$$

with $C := Le^{3(L+1)T}$.

In other words, the commutator lemma states

$$\left| \int_X \left[A_t(P_{t,s}u_s) - P_{t,s}(A_s u_s) \right] v_t \, dm_t \right| \le C(u_s, v_t) \cdot |t-s|^{1/2}. \tag{35}$$

Proof. Obviously, the function $r \mapsto \mathcal{E}_r(u_r, v_r)$ is finite (even locally bounded) and measurable on (σ, τ). Therefore, by Lebesgue's density theorem for a.e. $s, t \in (\sigma, \tau)$

$$\mathcal{E}_t(u_t, v_t) = \lim_{\delta \searrow 0} \frac{1}{\delta} \int_{t-\delta}^t \mathcal{E}_r(u_r, v_r) \, dr, \quad \mathcal{E}_s(u_s, v_s) = \lim_{\delta \searrow 0} \frac{1}{\delta} \int_s^{s+\delta} \mathcal{E}_r(u_r, v_r) \, dr$$

and thus

$$\mathcal{E}_t(u_t, v_t) - \mathcal{E}_s(u_s, v_s) \quad = \quad \lim_{\delta \searrow 0} \int_s^{t-\delta} \frac{1}{\delta} \Big(\mathcal{E}_{r+\delta}(u_{r+\delta}, v_{r+\delta}) - \mathcal{E}_r(u_r, v_r) \Big) \, dr.$$

To proceed, we decompose the integrand into three terms

$$\begin{aligned}
\frac{1}{\delta} \left[\mathcal{E}_{r+\delta}(u_{r+\delta}, v_{r+\delta}) - \mathcal{E}_r(u_r, v_r) \right] \quad &= \quad \frac{1}{\delta} \left[\mathcal{E}_{r+\delta}(u_{r+\delta}, v_{r+\delta}) - \mathcal{E}_{r+\delta}(u_r, v_{r+\delta}) \right] \\
&\quad + \frac{1}{\delta} \left[\mathcal{E}_{r+\delta}(u_r, v_{r+\delta}) - \mathcal{E}_r(u_r, v_{r+\delta}) \right] \\
&\quad + \frac{1}{\delta} \left[\mathcal{E}_r(u_r, v_{r+\delta}) - \mathcal{E}_r(u_r, v_r) \right] \\
&=: \quad \alpha_r(\delta) + \beta_r(\delta) + \gamma_r(\delta).
\end{aligned}$$

Let us first estimate the second term

$$\begin{aligned}
\beta_r(\delta) &= \frac{1}{4\delta}\left[\mathcal{E}_{r+\delta}(u_r+v_{r+\delta})+\mathcal{E}_{r+\delta}(u_r-v_{r+\delta})-\mathcal{E}_r(u_r+v_{r+\delta})-\mathcal{E}_r(u_r-v_{r+\delta})\right]\\
&\le \frac{3L}{4}\,e^{3L\delta}\left[\mathcal{E}_r(u_r+v_{r+\delta})+\mathcal{E}_r(u_r-v_{r+\delta})\right]\\
&\le \frac{3L}{2}\,e^{6L\delta}\left[\mathcal{E}_r(u_r)+\mathcal{E}_{r+\delta}(v_{r+\delta})\right]
\end{aligned}$$

due to the fact that $|\partial_r\mathcal{E}_r(w)| \le 3L\,\mathcal{E}_r(w)$ for each $w\in\mathcal{F}$. According to Theorem 2.20, the final expressions can be estimated (uniformly in δ) in terms of $\mathcal{E}_s(u_s)$ and $\mathcal{E}_t(v_t)+\|v_t\|^2_{L^2(m_t)}$. Thus we finally obtain

$$\begin{aligned}
\lim_{\delta\searrow 0}\int_s^{t-\delta}\beta_r(\delta)\,dr &\le \frac{3L}{2}\int_s^t\left[\mathcal{E}_r(u_r)+\mathcal{E}_r(v_r)\right]dr\\
&\le (t-s)\,\frac{3L}{2}\,e^{3L(t-s)}\Big[\mathcal{E}_s(u_s)+\mathcal{E}_t(v_t)+\|v_t\|^2_{L^2(m_t)}\Big].
\end{aligned}$$

Now let us consider jointly the first and third terms

$$\begin{aligned}
\int_s^{t-\delta}\left[\alpha_r(\delta)+\gamma_r(\delta)\right]dr &= \frac{1}{\delta}\int_s^{t-\delta}\left[\mathcal{E}_{r+\delta}\big((u_{r+\delta}-u_r),v_{r+\delta}\big)+\mathcal{E}_r\big(u_r,(v_{r+\delta}-v_r)\big)\right]dr\\
&= -\frac{1}{\delta}\int_s^{t-\delta}\int_X\Big[(u_{r+\delta}-u_r)\cdot A_{r+\delta}v_{r+\delta}\cdot e^{-f_{r+\delta}}\\
&\qquad +A_ru_r\cdot(v_{r+\delta}-v_r)\cdot e^{-f_r}\Big]\,dm_\diamond\,dr\\
&= -\frac{1}{\delta}\int_0^\delta\int_s^{t-\delta}\int_X\Big[A_{r+\epsilon}u_{r+\epsilon}\cdot A_{r+\delta}v_{r+\delta}\cdot e^{-f_{r+\delta}}+\\
&\qquad A_ru_r\cdot(-A_{r+\epsilon}v_{r+\epsilon}+\dot f_{r+\epsilon}v_{r+\epsilon})\cdot e^{-f_r}\Big]\,dm_\diamond\,dr\,d\epsilon
\end{aligned}$$

Integrability of $|A_ru_r|^2$ w.r.t. $dm_r\,dr$ implies that $\int_{t-\delta}^t|A_ru_r|^2dm_r\,dr\to 0$ as $\delta\to 0$ as well as $\int_s^{s+\delta}|A_ru_r|^2dm_r\,dr\to 0$. Thus together with Lipschitz continuity of $t\mapsto f_t$ this implies

$$\frac{1}{\delta}\int_0^\delta\int_s^{t-\delta}\int_X\Big[A_{r+\epsilon}u_{r+\epsilon}\cdot A_{r+\delta}v_{r+\delta}\cdot e^{-f_{r+\delta}}+-A_ru_r\cdot A_{r+\epsilon}v_{r+\epsilon}\cdot e^{-f_r}\Big]\,dm_\diamond\,dr\,d\epsilon\to 0$$

as $\delta\to 0$. Thus (since $\dot f$ is bounded by L and since $r\mapsto\|v_r\|_{L^2(m_r)}$ is non-decreasing)

$$\begin{aligned}
\lim_{\delta\to 0}\left|\int_s^{t-\delta}\left[\alpha_r(\delta)+\gamma_r(\delta)\right]dr\right| &\le -\frac{1}{\delta}\int_0^\delta\int_s^{t-\delta}\int_X\left|A_ru_r\cdot\dot f_{r+\epsilon}v_{r+\epsilon}\right|dm_r\,dr\,d\epsilon\\
&\le L\cdot|t-s|^{1/2}\cdot\left(\int_s^t\left|A_ru_r\right|^2dm_r\,dr\right)^{1/2}\cdot\|v_t\|_{L^2(m_t)}\\
&\le L\cdot|t-s|^{1/2}\cdot\left(\frac{1}{2}e^{3L(t-s)}\mathcal{E}_s(u_s)\right)^{1/2}\cdot\|v_t\|_{L^2(m_t)}.
\end{aligned}$$

To summarize, we have

$$\begin{aligned}
\big|\mathcal{E}_t(u_t,v_t)-\mathcal{E}_s(u_s,v_s)\big| &= \lim_{\delta\searrow 0}\left|\int_s^{t-\delta}\big(\alpha_r(\delta)+\beta_r(\delta)+\gamma_r(\delta)\big)\,dr\right| \\
&\le |t-s|\,\frac{3L}{2}\,e^{3L(t-s)}\Big[\mathcal{E}_s(u_s)+\mathcal{E}_t(v_t)+\|v_t\|^2_{L^2(m_t)}\Big] \\
&\quad +L\cdot|t-s|^{1/2}\cdot\left(\frac{1}{2}e^{3L(t-s)}\mathcal{E}_s(u_s)\right)^{1/2}\cdot\|v_t\|_{L^2(m_t)} \\
&\le C\cdot|t-s|^{1/2}\cdot\Big[\mathcal{E}_s(u_s)+\mathcal{E}_t(v_t)+\|v_t\|^2_{L^2(m_t)}\Big]
\end{aligned}$$

with $C:=Le^{3(L+1)T}$ according to the energy estimates of the previous Theorem.

□

2.6 Heat Flow and Optimal Transport on Time-dependent Metric Measure Spaces

We are now going to define, construct, and analyze the heat equation on *time-dependent metric measure spaces* $\big(X,d_t,m_t\big)_{t\in I}$.

2.6.1 The Setting

Here and for the rest of the chapter, our setting is as follows:

The 'state space' X is a Polish space and the 'parameter set' $I\subset\mathbb{R}$ will be a bounded open interval; for convenience we assume $I=(0,T)$. For each t under consideration, d_t will be a complete separable geodesic metric on X and m_t will be a σ-finite Borel measure on X. We always assume that there exist constants $C,K,L,N'\in\mathbb{R}$ such that

- the metrics d_t are uniformly bounded and equivalent to each other with
$$\left|\log\frac{d_t(x,y)}{d_s(x,y)}\right|\le L\cdot|t-s| \tag{36}$$
for all s,t and all x,y ('log Lipschitz continuity in t');
- the measures m_t are mutually absolutely continuous with bounded, Lipschitz continuous logarithmic densities; more precisely, choosing some reference measure $m_\diamond$ the measures can be represented as $m_t=e^{-f_t}m_\diamond$ with functions f_t satisfying $|f_t(x)|\le C$, $|f_t(x)-f_t(y)|\le C\cdot d_t(x,y)$ and
$$|f_s(x)-f_t(x)|\le L\cdot|s-t| \tag{37}$$
for all s,t and all x,y;
- for each t the static space (X,d_t,m_t) is infinitesimally Hilbertian and satisfies a curvature-dimension condition $\mathrm{CD}(K,N')$ in the sense of [57], [43], [5].

In terms of the metric d_t for given t, we define the L^2*-Kantorovich-Wasserstein metric* W_t on the space of probability measures on X:

$$W_t(\mu,\nu)=\inf\left\{\int_{X\times X}d_t^2(x,y)\,dq(x,y):\ q\in\mathrm{Cpl}(\mu,\nu)\right\}^{1/2}$$

where $\mathrm{Cpl}(\mu,\nu)$ as usual denotes the set of all probability measures on $X\times X$ with marginals μ and ν. In general, it is not really a metric but just a pseudo metric. Denote by $\mathcal{P}=\mathcal{P}(X)$ the set of all probability measures μ on X (equipped with its Borel σ-field) with $W_t(\mu,\delta_z)<\infty$ or some/all $z\in X$ and $t\in I$.

The log-Lipschitz bound (36) implies that for all $s,t\in I$ and all $\mu,\nu\in\mathcal{P}$

$$\left|\log\frac{W_t(\mu,\nu)}{W_s(\mu,\nu)}\right|\le L\cdot|t-s|. \tag{38}$$

Note that the latter is equivalent to weak differentiability of $t\mapsto W_t(\mu,\nu)$ and $|\partial_t W_t(\mu,\nu)|\le L\cdot W_t(\mu,\nu)$ for all $\mu,\nu\in\mathcal{P}$.

A powerful tool is the dual representation of W_t^2:

$$\frac{1}{2}W_t^2(\mu,\nu)=\sup\left\{\int\varphi d\mu+\int\psi d\nu:\varphi(x)+\psi(y)\le\frac{1}{2}d_t^2(x,y)\right\},$$

where the supremum is taken among all continuous and bounded functions φ,ψ. Closely related to this is the d_t-Hopf-Lax semigroup defined on bounded Lipschitz functions φ by

$$Q_a^t\varphi(x):=\inf_{y\in X}\left\{\varphi(y)+\frac{1}{2a}d_t^2(x,y)\right\},\quad a>0,\ x\in X.$$

The map $(a,x)\mapsto Q_a^t\varphi(x)$ satisfies the Hamilton-Jacobi equation

$$\partial_a Q_a^t\varphi(x)=-\frac{1}{2}(\mathrm{lip}_t Q_a^t\varphi)^2(x),\quad \lim_{a\to 0}Q_a^t\varphi(x)=\varphi(x). \tag{39}$$

In addition, since (X,d_t) is assumed to be geodesic,

$$\mathrm{Lip}(Q_a^t\varphi)\le 2\mathrm{Lip}(\varphi),\quad \mathrm{Lip}(Q_\cdot^t f(x))\le 2[\mathrm{Lip}(\varphi)]^2.$$

See for instance [7, Section 3] for these facts.

For $\mu,\nu\in\mathcal{P}(X)$ the Kantorovich duality can be written as

$$\frac{1}{2}W_t^2(\mu_0,\mu_1)=\sup_\phi\left\{\int Q_1^t\varphi d\mu_1-\int\varphi d\mu_0\right\}. \tag{40}$$

We say that a curve $\mu\colon J\to\mathcal{P}(X)$ belongs to $AC^p(J;\mathcal{P}(X))$ if

$$W_t(\mu^a,\mu^b)\le\int_a^b g(r)dr\quad\forall a<b\in J$$

for some $g\in L^p(J)$. We will exclusively treat the case $p=2$ and call μ a *2-absolutely continuous* curve. Recall that there exists a minimal function g, called *metric speed* and denoted by $|\dot\mu_a|_t$ such that

$$|\dot\mu^a|_t:=\lim_{b\to a}\frac{W_t(\mu^a,\mu^b)}{|b-a|}.$$

See for example [4, Theorem 1.1.2]. For continuous curves $\mu\in\mathcal{C}([0,1],\mathcal{P}(X))$ satisfying $\mu^a=u^a m$ with $u^a\le R$, μ belongs to $AC^2([0,1],\mathcal{P}(X))$ if and only if for each $t \in (0,T)$ there exists a velocity potential $(\Phi_t^a)_a$ such that $\int_0^1 \int \Gamma_t(\Phi_t^a) d\mu^a da < \infty$ and

$$\int \varphi d\mu^{a_1} - \int \varphi d\mu^{a_0} = \int_{a_0}^{a_1} \int \Gamma_t(\varphi, \Phi_t^a) d\mu^a da, \text{ for every } \varphi \in Dom(\mathcal{E}). \quad (41)$$

Moreover we can express the metric speed in the following way

$$|\dot{\mu}^a|_t^2 = \int \Gamma_t(\Phi_t^a) d\mu^a. \quad (42)$$

See section 6 and 8 in [8] for a detailed discussion.

Occasionally, we have to measure the 'distance' between points $x, y \in X$ which belong to different time sheets. In this case, for $s, t \in I$ and $\mu, \nu \in \mathcal{P}(X)$ we define

$$W_{s,t}(\mu,\nu) := \inf \lim_{h \to 0} \sup_{\substack{0=a_0<\cdots<a_n=1,\\ a_i - a_{i-1} \le h}} \left\{ \sum_{i=1}^n (a_i - a_{i-1})^{-1} W^2_{\vartheta(a_{i-1})}(\mu^{a_{i-1}}, \mu^{a_i}) \right\}^{1/2}$$

where the infimum runs over all 2-absolutely continuous curves $\mu\colon [0,1] \to \mathcal{P}(X)$ with $\mu_0 = \mu$, $\mu_1 = \nu$. See Section 6.1 for a detailed discussion and in particular for the equivalent characterization

$$W_{s,t}(\mu,\nu) = \inf \left\{ \int_0^1 |\dot{\mu}^a|^2_{W_{s+a(t-s)}} da \right\}^{1/2} \quad (43)$$

where the infimum runs over all 2-absolutely continuous curves $(\rho^a)_{a\in[0,1]}$ in $\mathcal{P}(X)$ connecting μ and ν.

In the following we will make frequently use of the concept of regular curves, which already has been successfully used in [7, 24, 8]. We use the refined version of [8].

Definition 2.27. *For fixed $t \in [0,T]$, let $\rho^a = u^a m_t \in \mathcal{P}(X)$, $a \in [0,1]$. We say that the curve ρ is regular (w.r.t. m_t) if:*

1. $u \in \mathcal{C}^1([0,1], L^1(X)) \cap \mathrm{Lip}([0,1], \mathcal{F}^*)$,
2. *there exists a constant $R > 0$ such that $u^a \le R$ m-a.e. for every $a \in [0,1]$,*
3. *there exists a constant $E > 0$ such that $\mathcal{E}_t(\sqrt{u^a}) \le E$ for every $a \in [0,1]$.*

Remark. *Due to our assumptions on the measures, $(\rho^a)_a$ is a regular curve w.r.t m_t if and only if it is also a regular curve w.r.t m_s. In this case, it is also a regular curve w.r.t m_ϑ, where ϑ is a function belonging to $\mathcal{C}^1([0,1],\mathbb{R})$. So we will just say regular curve.*

We will use the following approximation result which is a combination of [8, Lemma 12.2] and [24, Lemma 4.11].

Lemma 2.28. *Let X be a RCD(K,∞) space. Let $\rho_0, \rho_1 \in \mathcal{P}(X)$ and $(\rho^a)_{a\in[0,1]}$ be the W_t-geodesic connecting them. Then there exists a sequence of regular curves $(\rho_n^a)_{a\in[0,1]}$, $n \in \mathbb{N}$, such that*

$$W_t(\rho_n^a, \rho_a) \to 0 \text{ for every } a \in [0,1], \quad (44)$$

$$\limsup_{n\to\infty} \int_0^1 |\dot{\rho}_n^a|_t^2 da \le W_t^2(\rho_0, \rho_1). \quad (45)$$

If we additionally impose that $\rho_0, \rho_1 \in Dom(S)$, then

$$S_t(\rho_a^n) \to S_t(\rho_a) \text{ for every } a \in [0,1], \tag{46}$$

and

$$\limsup_{n\to\infty} \sup_{a\in[0,1]} S_t(\rho_a^n) \leq \sup_{a\in[0,1]} S_t(\rho_a) = \max_{a\in[0,1]} S_t(\rho_a). \tag{47}$$

Proof. We follow the argumentation in [8, Lemma 12.2] and approximate ρ_0, ρ_1 by two sequences of measures $\{\sigma_i^n\}_n$ with bounded densities. Then as in [7, Proposition 4.11] one employs a threefold regularization procedure to the W_t-geodesic $(\nu_a^n)_a$ connecting σ_0^n and σ_1^n: Given $k \in \mathbb{N}$, we first define $\rho_a^{n,k,1} = H_{1/k}^t \nu_a^n$, where H^t denotes the static semigroup. Then we set

$$\rho_a^{n,k,2} = \int_{\mathbb{R}} \rho_{a-a'}^{n,k,1} \chi_k(a') da',$$

where $\chi_k(a) = k\chi(ka)$ for some smooth kernel $\chi \in \mathcal{C}_c(\mathbb{R})$. Finally we set $\rho_a^{n,k} = h^{1/k,t}\rho_a^{n,k,2}$, where $h^{1/k,t}$ denotes the mollification of the static t-semigroup. Then by a standard diagonal argument one obtains a sequence of regular curves in the sense of Definition 2.27 satisfying (44) and (45).

In order to show (46) and (47) note that since X is a RCD(K,∞) space we have that $a \mapsto S_t(\rho_a)$ is K-convex, where (ρ_a) denotes the W_t geodesic. Together with the lower semicontinuity of the entropy the map $a \mapsto S_t(\rho_a)$ is continuous. Using the convexity properties we follow the argumentation in [24, Lemma 4.11] and insert the explicit formulas of the regularization (ρ_a^n) to obtain

$$\begin{aligned} S_t(\rho_a^n) \leq S_t(\rho_a^{n,2}) &\leq \int_{\mathbb{R}} \chi_n(a') S_t(\rho_{a-a'}) da' \\ &\leq S_t(\rho_a) + \int_{\mathbb{R}} \chi_n(a') |S_t(\rho_{a-a'}) - S_t(\rho_a)| da'. \end{aligned} \tag{48}$$

Since $a \mapsto S_t(\rho_a)$ is uniformly continuous by compactness, the last term vanishes as $n \to \infty$. Thus we obtain $\limsup_{n\to\infty} S_t(\rho_a^n) \leq S_t(\rho_a)$. The lower semicontinuity in turn implies (46).
One obtains (47) from (48) by exploiting the uniform continuity of the entropy along geodesics on compact intervals once more. $\square$

Later on in this chapter (Section 2.7.2), we will see that there is an easier construction of regular curves based on the 'dual heat flow' to be introduced next.

2.6.2 The Heat Equation on Time-dependent Metric Measure Spaces

Due to the CD(K,N')-condition for each of the static spaces (X, d_t, m_t), the detailed analysis of energies, gradients and heat flows on mm-spaces due to Ambrosio, Gigli and Savaré [4, 5, 6, 7] applies. In particular, for each t there is a well-defined energy functional

$$\mathcal{E}_t(u) = \int_X |\nabla_t u|^2 dm_t = \liminf_{\substack{v\to u \text{ in } L^2(X,m_t) \\ v\in \mathrm{Lip}(X,d_t)}} \int_X (\mathrm{lip}_t v)^2 \, dm_t \tag{49}$$

for $u \in L^2(X, m_t)$ where $\mathrm{lip}_t u(x)$ denotes the pointwise Lipschitz constant (w.r.t. the metric d_t) at the point x and $|\nabla_t u|$ denotes the minimal weak upper gradient (again w.r.t. d_t). Since (X, d_t, m_t) is assumed to be infinitesimally Hilbertian, for each t under consideration $\mathcal{E}_t$ is a quadratic form. Indeed, it is a strongly local, regular Dirichlet form with intrinsic metric d_t and square field operator

$$\Gamma_t(u) = |\nabla_t u|^2.$$

In the sequel, we freely switch between these two notations of the same object.

The Laplacian Δ_t is defined as the generator of $\mathcal{E}_t$, i.e. as the unique non-positive self-adjoint operator on $L^2(X, m_t)$ with domain $\mathcal{D}(\Delta_t) \subset \mathcal{D}(\mathcal{E}_t)$ and

$$-\int_X \Delta_t u\, v\, dm_t = \mathcal{E}_t(u, v) \qquad (\forall u \in \mathcal{D}(\Delta_t), v \in \mathcal{D}(\mathcal{E}_t)).$$

Thanks to the $\mathrm{RCD}(K, \infty)$-condition, for each t the domain of the Laplacian coincides with the domain of the Hessian [27], i.e. $Dom(\Delta_t) = W^{2,2}(X, d_t, m_t)$. Indeed, the 'self-improved Bochner inequality' implies that

$$\Gamma_{2,t}(u) \geq K\, |\nabla_t u|^2 + |\nabla_t^2 u|^2_{HS}$$

which after integration w.r.t. m_t, integration by parts, and application of Cauchy-Schwarz inequality gives

$$\|\nabla_t^2 u\|^2 \leq (1 + K_-/2) \cdot \Big(\|\Delta_t u\|^2 + \|u\|^2 \Big) \tag{50}$$

with $K_- := \max\{-K, 0\}$ and $\|.\|^2 := \|.\|^2_{L^2(m_t)}$.

Note that in general, $Dom(\Delta_t)$ may depend on t, see Example 2.15.

Due to our assumptions that the measures are uniformly equivalent and that the metrics are uniformly equivalent, the sets $L^2(X, m_t)$ and $W^{1,2}(X, d_t, m_t) := \mathcal{D}(\mathcal{E}_t)$ do not depend on t and the respective norms for varying t are equivalent to each other. We put $\mathcal{H} = L^2(X, m_\diamond)$ and $\mathcal{F} = \mathcal{D}(\mathcal{E}_\diamond)$ as well as

$$\mathcal{F}_{(s,\tau)} = L^2\big((s,\tau) \to \mathcal{F}\big) \cap H^1\big((s,\tau) \to \mathcal{F}^*\big) \subset \mathcal{C}\big([s,\tau] \to \mathcal{H}\big)$$

for each $0 \leq s < \tau \leq T$. For the definition of 'solution to the heat equation' and for the existence of the heat propagator we refer to the previous chapter.

Theorem 2.29. *(i) For each $0 \leq s < \tau \leq T$ and each $h \in \mathcal{H}$ there exists a unique solution $u \in \mathcal{F}_{(s,\tau)}$ to the heat equation $\partial_t u_t = \Delta_t u_t$ on $(s,\tau) \times X$ with $u_s = h$.*

(ii) The heat propagator $P_{t,s} : h \mapsto u_t$ admits a kernel $p_{t,s}(x,y)$ w.r.t. m_s, i.e.

$$P_{t,s} h(x) = \int p_{t,s}(x,y) h(y)\, dm_s(y). \tag{51}$$

If X is bounded, for each $(s', y) \in (s, T) \times X$ the function $(t,x) \mapsto p_{t,s}(x,y)$ is a solution to the heat equation on $(s', T) \times X$.

(iii) All solutions $u : (t,x) \mapsto u_t(x)$ to the heat equation on $(s,\tau) \times X$ are Hölder continuous in t and x. All nonnegative solutions satisfy a scale invariant parabolic Harnack inequality of Moser type.

(iv) The heat kernel $p_{t,s}(x,y)$ is Hölder continuous in all variables, it is Markovian

$$\int p_{t,s}(x,y)\,dm_s(y) = 1 \qquad (\forall s<t, \forall x)$$

and has the propagator property

$$p_{t,r}(x,z) = \int p_{t,s}(x,y)\,p_{s,r}(y,z)\,dm_s(y) \qquad (\forall r<s<t, \forall s,z).$$

Proof. (i) It remains to verify the boundedness and regularity assumptions on f_t and Γ_t which were made for Theorem 2.10. Choose a reference point $t_0 \in I$ and put $\Gamma_\diamond = \Gamma_{t_0}$. Then $\mathcal{E}_\diamond(u) = \int \Gamma_{t_0}(u)e^{-f_{t_0}}dm_\diamond$. The uniform bounds on f_t and on $\Gamma_\diamond(f_t)$ are stated as assumption (37). The log Lipschitz bound (36) on d_t implies the requested uniform bound on Γ_t. The claim thus follows from Theorem 2.10.

(ii), (iii), (iv) The RCD-condition with finite N' implies scale invariant Poincaré inequalities and doubling properties for each of the static spaces (X, d_t, m_t) with uniform constants. Together with the uniform bounds on f_t, $\Gamma_t(.)$ and $\Gamma_t(f_t)$ this allows to apply results of [41] which provides all the assertions of the Theorem. $\square$

Remark 2.30. *The formula* (51) *allows to give a pointwise definition for $P_{t,s}h(x)$ for each $h \in L^2(X, m_\diamond)$ (or, in other words, to select a 'nice' version) and, moreover, it allows to extend its definition to $h \in L^1 \cup L^\infty$.*

Recall, however, that in general the operator $P_{t,s}$ is not symmetric w.r.t. any of the involved measures (m_t, m_s or $m_\diamond$) and that in general the operator norm in L^p for $p \neq \infty$ will not be bounded by 1.

2.6.3 The Dual Heat Equation

By duality, the propagator $(P_{t,s})_{s\le t}$ acting on bounded continuous functions induces a *dual propagator* $(\hat{P}_{t,s})_{s\le t}$ acting on probability measures as follows

$$\int u\,d(\hat{P}_{t,s}\mu) = \int (P_{t,s}u)d\mu \qquad (\forall u \in \mathcal{C}_b(X), \forall \mu \in \mathcal{P}(X)). \tag{52}$$

It obviously has the 'dual propagator property' $\hat{P}_{t,r} = \hat{P}_{s,r} \circ \hat{P}_{t,s}$. Whereas the time-dependent function $v_t(x) = P_{t,s}u(x)$ is a solution to the heat equation

$$\partial_t v = \Delta_t v, \tag{53}$$

the time-dependent measure $\nu_s(dy) = \hat{P}_{t,s}\mu(dy)$ is a solution to the *dual heat equation*

$$-\partial_s \nu = \hat{\Delta}_s \nu.$$

Here again $\hat{\Delta}_s$ is defined by duality: $\int u\,d(\hat{\Delta}_s\mu) = \int \Delta_s u\,d\mu \quad (\forall u, \forall \mu)$.

If we define Markov kernels $p_{t,s}(x,dy)$ for $s \le t$ by $p_{t,s}(x,dy) = p_{t,s}(x,y)\,dm_s(y)$ then

$$P_{t,s}u(x) = \int u(y)p_{t,s}(x,dy) = \int u(y)p_{t,s}(x,y)\,dm_s(y)$$

and the dual propagator is given by

$$(\hat{P}_{t,s}\mu)(dy) = \int p_{t,s}(x, dy)\, d\mu(x) = \left[\int p_{t,s}(x, y)\, d\mu(x)\right] dm_s(y).$$

In particular, $(\hat{P}_{t,s}\delta_x)(dy) = p_{t,s}(x, dy)$. Note that $\hat{P}_{t,s}\mu(X) = \int P_{t,s}1(x) d\mu(x) = 1$.

Theorem 2.31. *(i) For each $0 \le \sigma < t \le T$ and each $g \in \mathcal{H}$ there exists a unique solution $v \in \mathcal{F}_{(0,t)}$ to the adjoint heat equation $\partial_s v_s = -\Delta_s v_s + (\partial_s f_s) v_s$ on $(\sigma, t) \times X$ with $v_t = g$.*

(ii) This solution is given as $v_s(y) = P^*_{t,s} g(y)$ in term of the adjoint heat propagator

$$P^*_{t,s} g(y) = \int p_{t,s}(x, y) g(x)\, dm_t(x). \tag{54}$$

If X is bounded, for each $(t', x) \in (0, t) \times X$ the function $(s, y) \mapsto p_{t,s}(x, y)$ is a solution to the adjoint heat equation on $(0, t') \times X$.

(iii) All solutions $v : (s, y) \mapsto v_s(y)$ to the adjoint heat equation on $(\sigma, t) \times X$ are Hölder continuous in s and y. All nonnegative solutions satisfy a scale invariant parabolic Harnack inequality of Moser type.

Proof. The assumption on Lipschitz continuity of $t \mapsto f_t$ implies that all the regularity assumptions requested in [41] also hold for the time-dependent operators $\Delta_s - (\partial_s f_s)$ (which then are just the operators Δ_s perturbed by multiplication operators in terms of bounded functions). Thus all the previous results apply without any changes. □

Corollary 2.32. *For all $g, h \in L^1(X)$*

$$\int h \cdot P^*_{t,s} g\, dm_s = \int P_{t,s} h \cdot g\, dm_t$$

and

$$\hat{P}_{t,s}\big(g \cdot m_t\big) = \big(P^*_{t,s} g\big) \cdot m_s. \tag{55}$$

Lemma 2.33. *(i) $\hat{P}_{t,s}$ is continuous on $\mathcal{P}(X)$ w.r.t. weak convergence.*

(ii) The dual heat flow $s \mapsto \mu_s = \hat{P}_{t,s}\mu$ is uniformly Hölder continuous (w.r.t. any of the metrics $W_\tau, r \in I$, see next section). More precisely, there exists a constant C such that for all $s, s' < t$, all τ and all μ

$$W^2_\tau(\mu_s, \mu_{s'}) \le C \cdot |s - s'|. \tag{56}$$

(iii) If X is compact then for each $s < t$

$$\hat{P}_{t,s} : \mathcal{P}(X) \to \mathcal{D}$$

where $\mathcal{D} = \{\mu \in \mathcal{P}(X) :\ \mu = u\, m_\diamond,\ u \in \mathcal{F} \cap L^\infty,\ 1/u \in L^\infty\}$.

(iv) For $\mu \in \mathcal{P}(X)$ such that $\mu \in Dom(S)$, the dual heat flow $(\hat{P}_{t,s}\mu)_{s<t}$ belongs to
$AC^2([0, t], \mathcal{P}(X))$.

Proof. (i) For each bounded continuous u on X the function $P_{t,s}u$ is bounded continuous. Thus $\mu_n \to \mu$ implies

$$\int u\, d\hat{P}_{t,s}\mu_n = \int P_{t,s}u\, d\mu_n \to \int P_{t,s}u\, d\mu = \int u\, d\hat{P}_{t,s}\mu$$

which proves the requested convergence $\hat{P}_{t,s}\mu_n \to \hat{P}_{t,s}\mu$.

(ii) Given $\mu_s = \hat{P}_{t,s}\mu$ and $\mu_{s'} = \hat{P}_{t,s'}\mu$ for $s < s' < t$. Then

$$W_\tau^2(\mu_s, \mu_{s'}) \quad \le \quad \int\int d_\tau^2(x,y)\, p_{s',s}(x,y)\, dm_s(y)\, d\mu_{s'}(x).$$

According to [56, 41], the heat kernel admits upper Gaussian estimates of the form

$$p_{s',s}(x,y) \le \frac{C}{m_\tau(B_\tau(\sqrt{\sigma}, x))} \cdot \exp\Big(- \frac{d_\tau^2(x,y)}{C\sigma}\Big)$$

with $\sigma := |s-s'|$ and $B_\tau(r,x)$ denoting the ball of radius r around x in the metric space (X, d_τ). Moreover, Bishop-Gromov volume comparison in RCD(K,N)-spaces provides an upper bound for the volume of spheres

$$A(R,x) \le \Big(\frac{R}{r}\Big)^{N-1} \cdot e^{R\sqrt{|K|(N-1)}} \cdot A(r,x)$$

for $R \ge r$ where $A(r,x) = \partial_{r+} m_\tau(B_\tau(r,x))$ and thus (by integrating from 0 to $\sqrt{\sigma}$)

$$A(R,x) \le N \frac{R^{N-1}}{\sigma^{N/2}} \cdot e^{R\sqrt{|K|(N-1)}} \cdot m_\tau(B_\tau(\sqrt{\sigma}, x))$$

for $R \ge \sqrt{\sigma}$. Hence, we finally obtain

$$\begin{aligned}
&W_\tau^2(\mu_s, \mu_{s'}) \le \int\int d_\tau^2(x,y)\, p_{s',s}(x,y)\, dm_s(y)\, d\mu_{s'}(x) \\
&\le \int_X \Big[\frac{C}{m_\tau(B_\tau(\sqrt{\sigma}, x))} \cdot \int_X d_\tau^2(x,y) \cdot \exp\Big(- \frac{d_\tau^2(x,y)}{C\sigma}\Big) dm_\tau(y)\Big] d\mu_{s'}(x) \\
&\le C\sigma + C \int_X \int_{\sqrt{\sigma}}^\infty R^2 \cdot \exp\Big(- \frac{R^2}{C\sigma}\Big) N \frac{R^{N-1}}{\sigma^{N/2}} \cdot e^{R\sqrt{|K|(N-1)}}\, dR\, d\mu_{s'}(x) \\
&\le C' \cdot \sigma.
\end{aligned}$$

(iii) By definition of solution to the adjoint heat equation, the densities u_s of $\hat{P}_{t,s}\mu$ (w.r.t. m_s) lie in $Dom(\mathcal{E})$. Parabolic Harnack inequality implies continuity and positivity. Together with compactness of X this yields upper and lower bounds (away from 0) for u.

(iv) In a similar calculation as in Proposition 2.16, we find for $\mu = v m_t$, $\mu_s = \hat{P}_{t,s}\mu$ since the dual heat flow is mass preserving,

$$\begin{aligned}
\int_s^t \int \Gamma_r(\log v_r) d\mu_r dr &= S_t(\mu) - S_s(\mu_s) - \int_s^t \int v_r \partial_r f_r dm_r dr \\
&\le S_t(\mu) + m_t(X) + L(t-s).
\end{aligned}$$

Now choose $\phi \in Dom(\mathcal{E})$ with $\phi, \Gamma(\phi) \in L^\infty(X)$. Then

$$
\begin{aligned}
&\left| \int \phi v_t dm_t - \int \phi v_s dm_s \right| = \left| \int_s^t \mathcal{E}_r(\phi, v_r) dr \right| \\
&\leq \int_s^t \left(\int \Gamma_r(\phi) v_r dm_r \right)^{1/2} \left(\int \Gamma_r(\log v_r) v_r dm_r \right)^{1/2} dr \\
&\leq \int_s^t \left(\int \Gamma_t(\phi) v_r dm_r \right)^{1/2} \left(e^{2L(s-t)} \int \Gamma_r(\log v_r) v_r dm_r \right)^{1/2} dr
\end{aligned}
$$

Then, Theorem 7.3 in [1] yields

$$
|\dot{\mu}_r|_t^2 \leq e^{2L(s-t)} \int \Gamma_r(\log v_r) v_r dm_r \in L^1_{loc}((0,t)),
$$

where the last conclusion is due to our previous calculation. $\square$

Lemma 2.34. *Let $u, g \in Dom(\mathcal{E})$ and $t \in (0,T)$ with $g \in L^1(X, m_t)$. Then,*

$$
\lim_{h \searrow 0} \frac{1}{h} \left(\int u g dm_t - \int u P^*_{t,t-h} g dm_{t-h} \right) = \int \Gamma_t(u, g) dm_t.
$$

Proof. Without loss of generality assume that $g \geq 0$ and $\int g \, dm_t = 1$. The general case can be obtained by considering the positive and negative parts separately and normalization. We first prove that for $g \in Dom(\mathcal{E})$ and $u \in \mathrm{Lip}(X)$

$$
\frac{1}{h} \left(\int u g dm_t - \int u P^*_{t,t-h} g dm_{t-h} \right) = \int_0^1 \int \Gamma_{t-rh}(u, P^*_{t,t-rh} g) dm_{t-rh} dr. \tag{57}
$$

Note that for $0 \leq r_1 \leq r_2 \leq 1$

$$
\begin{aligned}
&\left| \int u P^*_{t,t-r_2 h} g dm_{t-r_2 h} - \int u P^*_{t,t-r_1 h} g dm_{t-r_1 h} \right| \\
&\qquad \leq \mathrm{Lip}(u) W_2(\hat{P}_{t,t-r_2 h}(g m_t), \hat{P}_{t,t-r_1 h}(g m_t)),
\end{aligned}
$$

and hence, as a consequence of Lemma 2.33(ii), the map $r \mapsto \int u P^*_{t,t-rh} g dm_{t-rh}$ is absolutely continuous. Thus

$$
\begin{aligned}
&\frac{1}{h} \left(\int u g dm_t - \int u P^*_{t,t-h} g dm_{t-h} \right) = -\frac{1}{h} \int_0^1 \partial_r \int u P^*_{t,t-rh} g dm_{t-rh} dr \\
&= -\frac{1}{h} \int_0^1 \int u e^{-f_{t-rh}} \partial_r P^*_{t,t-rh} g dm_\diamond - \frac{1}{h} \int_0^1 \int u P^*_{t,t-rh} g \partial_r e^{-f_{t-rh}} dm_\diamond dr \\
&= \int_0^1 \mathcal{E}^\diamond_{t-rh}(P^*_{t,t-rh} g, u e^{-f_{t-rh}}) dr + \int_0^1 \int P^*_{t,t-rh} g u e^{-f_{t-rh}} \partial_r f_{t-rh} dm_\diamond dr \\
&- \int_0^1 \int P^*_{t,t-rh} g u e^{-f_{t-rh}} \partial_r f_{t-rh} dm_\diamond dr \\
&= \int_0^1 \mathcal{E}^\diamond_{t-rh}(P^*_{t,t-rh} g, u e^{-f_{t-rh}}) dr = \int_0^1 \mathcal{E}_{t-rh}(P^*_{t,t-rh} g, u) dr,
\end{aligned}
$$

where we used that $r \mapsto P^*_{t,t-rh} g$ is a rescaled solution to the adjoint heat equation.

Since we assume that the space has a lower Riemannian Ricci bound, we obtain equation (57) for every $u \in Dom(\mathcal{E})$ by approximating with Lipschitz functions u_n, satisfying $u_n \to u$ strongly in $(Dom(\mathcal{E}), \sqrt{||\cdot||^2_{L^2(X)} + \mathcal{E}(\cdot)})$, see [6, Proposition 4.10]. Hence

$$\begin{aligned}\lim_{h\searrow 0} \frac{1}{h}\left(\int u g dm_t - \int u P^*_{t,t-h} g dm_{t-h}\right) &= \lim_{h\searrow 0}\int_0^1 \int \Gamma_{t-rh}(u, P^*_{t,t-rh} g) dm_{t-rh} dr \\ &= \int_0^1 \lim_{h\searrow 0} \int \Gamma_{t-rh}(u, P^*_{t,t-rh} g) dm_{t-rh} dr \\ &= \int \Gamma_t(u,g) dm_t,\end{aligned}$$

where the third inequality directly follows from Lemma 2.25 and the second equality follows from dominated convergence. □

To summarize:

- ▷ Given any $h \in L^2(X, m_s)$ the function $(t,x) \mapsto u_t(x) = P_{t,s} h(x)$ solves the *heat equation* $\partial_t u_t = \Delta_t u_t$ in $(s,T) \times X$ with initial condition $u_s = h$. In Markov process theory, this is the *Kolmogorov backward equation* (in reverse time direction).
- ▷ By duality we obtain the *dual propagator* $\hat{P}_{t,s}$ acting on probability measures. Given any $\nu \in (\mathcal{P}(X), W_t)$, the probability measures $(s,y) \mapsto \mu_s = \hat{P}_{t,s}\nu$ solve the *dual heat equation* $-\partial_s \mu_s = \hat{\Delta}_s \mu_s$ in $[0,t) \times X$ with terminal condition $\mu_t = \nu$.
- ▷ Their densities $v_s = \frac{d\mu_s}{dm_s}$ solve the *Fokker-Planck equation* or *Kolmogorov forward equation* (in reverse time direction)

  $$-\partial_s v_s = \Delta_s v_s - \partial_s f_s \cdot v_s$$

  in $(0,t) \times X$. The latter is also called *adjoint heat equation.*

2.7 Towards Transport Estimates

In the sequel, N always will denote an extended number in $[1,\infty]$. The assumptions from section 2.6.1 will always be in force (in particular, we assume $\mathrm{RCD}^*(K,N')$ and the bounds (36) and (37)). Moreover, X will be assumed to be bounded (and thus compact).

2.7.1 From Dynamic Convexity to Transport Estimates

Definition 2.35. *We say that the time-dependent mm-space $(X, d_t, m_t)_{t\in I}$ is a* super-N-Ricci flow *if the Boltzmann entropy S is* dynamical N-convex *on $I \times \mathcal{P}$ in the following sense: for a.e. $t \in I$ and every W_t-geodesic $(\mu^a)_{a\in[0,1]}$ in $\mathcal{P}$ with $\mu^0, \mu^1 \in Dom(S)$*

$$\partial_a^+ S_t(\mu^a)\big|_{a=1-} - \partial_a^- S_t(\mu^a)\big|_{a=0+} \geq -\frac{1}{2}\partial_t^- W^2_{t-}(\mu^0,\mu^1) + \frac{1}{N}\Big|S_t(\mu^0) - S_t(\mu^1)\Big|^2. \tag{58}$$

N-super Ricci flows in the case $N=\infty$ are simply called super Ricci flows.

Recall that $\mathcal{D}=\{\mu\in\mathcal{P}(X):\ \mu=u\,m_\diamond,\ u\in\mathcal{F}\cap L^\infty,\ 1/u\in L^\infty\}$.

Proposition 2.36. *Given probability measures $\mu,\nu\in\mathcal{D}\subset\mathcal{P}$, then the W_t-geodesic $(\rho^a)_{a\in[0,1]}$ connecting μ and ν has uniformly bounded densities $\frac{d\rho^a}{dm_t}\le C$ and there exist W_t-Kantorovich potentials ϕ from μ to ν and ψ from ν to μ (both conjugate to each other) such that*

$$\partial_a S_t(\rho^a)\big|_{a=0+}\ge -\mathcal{E}_t(\phi,u),\qquad \partial_a S_t(\rho^a)\big|_{a=1-}\le +\mathcal{E}_t(\psi,v).$$

Proof. This result uses only properties of the static mm-space (X,d_t,m_t). It can be found as estimate (6.19) in the proof of Theorem 6.5 in [3]. Note that due to our (upper and lower) boundedness assumption on u,v, no extra regularization is requested. □

Proposition 2.37. *Given $\tau\le T$ and $\mu,\nu\in\mathcal{D}\subset\mathcal{P}$, put $\mu_t=\hat P_{t,\tau}\mu$ and $\nu_t=\hat P_{t,\tau}\nu$. For each $t\in(0,\tau)$, let ϕ_t and ψ_t be any conjugate W_t-Kantorovich potentials from μ_t to ν_t and vice versa. Then for every $t\in(0,\tau)$*

$$\frac{1}{2}\partial_r^- W_t^2(\mu_r,\nu_r)|_{r=t+}\ge \mathcal{E}_t(\phi_t,u_t)+\mathcal{E}_t(\psi_t,v_t) \tag{59}$$

whereas

$$\frac{1}{2}\partial_r^+ W_t^2(\mu_r,\nu_r)|_{r=t-}\le \mathcal{E}_t(\phi_t,u_t)+\mathcal{E}_t(\psi_t,v_t). \tag{60}$$

Here u_t and v_t denote the densities of μ_t and ν_t, resp., w.r.t. m_t.

Proof. We closely follow the argumentation of the proof of Theorem 6.3 in [3]. According to Proposition 2.20, $u_t,v_t\in Dom(\mathcal{E})$. Moreover, due to boundedness of X, the Kantorovich potentials ϕ_t and ψ_t are Lipschitz and thus also lie in $Dom(\mathcal{E})$. Since ϕ_t and ψ_t are conjugate W_t-Kantorovich potentials from μ_t to ν_t and vice versa, we get

$$\frac{1}{2}W_t^2(\mu_t,\nu_t)=\int\phi_t d\mu_t+\int\psi_t d\nu_t$$

whereas

$$\frac{1}{2}W_t^2(\mu_r,\nu_r)\ge\int\phi_t d\mu_r+\int\psi_t d\nu_r$$

for $r\neq t$. Thus

$$\begin{aligned}&\frac{1}{2}\liminf_{r\searrow t}\frac{1}{r-t}\left[W_t^2(\mu_r,\nu_r)-W_t^2(\mu_t,\nu_t)\right]\\ &\quad\ge\ \liminf_{r\searrow t}\frac{1}{r-t}\left[\int\phi_t[d\mu_r-d\mu_t]+\int\psi_t[d\nu_r-d\nu_t]\right]\\ &\quad=\ \mathcal{E}_t(\phi_t,u_t)+\mathcal{E}_t(\psi_t,v_t).\end{aligned}$$

Similarly, we obtain

$$\begin{aligned}&\frac{1}{2}\limsup_{r\nearrow t}\frac{1}{t-r}\left[W_t^2(\mu_t,\nu_t)-W_t^2(\mu_r,\nu_r)\right]\\ &\quad\le\ \limsup_{r\nearrow t}\frac{1}{t-r}\left[\int\phi_t[d\mu_t-d\mu_r]+\int\psi_t[d\nu_t-d\nu_r]\right]\\ &\quad=\ \mathcal{E}_t(\phi_t,u_t)+\mathcal{E}_t(\psi_t,v_t).\end{aligned}$$

□

Theorem 2.38. *Assume that* $\big(X, d_t, m_t\big)_{t\in(0,T)}$ *is a super-Ricci flow and that* $(\mu_t)_{t\le\tau}$ *and* $(\nu_t)_{t\le\tau}$ *are dual heat flows started in probability measures* $\mu_\tau, \nu_\tau \in \mathcal{D}$. *Then*

$$\partial^-_{t+} W_t^2(\mu_t, \nu_t) \ge 0.$$

Proof. The assumptions on the densities are preserved by the dual heat flow, that is, μ_t and ν_t will have densities in $Dom(\mathcal{E})$ which are bounded from above and bounded away from 0, uniformly in t. According to Proposition 2.36

$$\partial_a S_t(\eta^{1-}) - \partial_a S_t(\eta^{0+}) \le \mathcal{E}_t(\phi_t, u_t) + \mathcal{E}_t(\psi_t, v_t)$$

with ϕ_t and ψ_t being suitable W_t-Kantorovich potentials from μ_t to ν_t and vice versa. Proposition 2.37 yields

$$\mathcal{E}_t(\phi_t, u_t) + \mathcal{E}_t(\psi_t, v_t) \le \frac{1}{2}\partial^-_r W_t^2(\mu_r, \nu_r)\big|_{r=t+}.$$

Being a super-Ricci flow implies

$$-\frac{1}{2}\partial^-_r W_r^2(\mu_t, \nu_t)\big|_{r=t-} \le \partial_a S(\eta^{1-}) - \partial_a S(\eta^{0+}) \tag{61}$$

for every W_t-geodesic $(\eta^b)_{b\in[0,1]}$ connecting μ_t and ν_t. Summing up these inequalities (and multiplying by 2), we arrive at

$$\partial^-_r W_t^2(\mu_r, \nu_r)\big|_{r=t+} + \partial^-_r W_r^2(\mu_t, \nu_t)\big|_{r=t-} \ge 0, \tag{62}$$

which seems to be almost the claim. However, applying the chain rule for (non continuously differentiable) functions which depend twice on the same variable requires some care. Note that the first term in the above inequality reads $\liminf_{\delta\searrow 0} \frac{1}{\delta}\big(W_t^2(\mu_{t+\delta}, \nu_{t+\delta}) - W_t^2(\mu_t, \nu_t)\big)$. To conclude $\partial^-_{t+} W_t^2(\mu_t, \nu_t) \ge 0$ we will have to replace the second term in the above inequality by

$$\liminf_{\delta\searrow 0} \frac{1}{\delta}\Big(W_{t+\delta}^2(\mu_{t+\delta}, \nu_{t+\delta}) - W_t^2(\mu_{t+\delta}, \nu_{t+\delta})\Big).$$

To do so, we pass to the integrated version (w.r.t. t). Using the absolute continuity of $t \mapsto W_t^2(\mu_t, \nu_t)$, we obtain for all $r < s$

$$\begin{aligned}
W_s^2(\mu_s, \nu_s) - W_r^2(\mu_r, \nu_r) &= \int_r^s \limsup_{\delta\searrow 0} \frac{1}{\delta}\Big[W_t^2(\mu_{t+\delta}, \nu_{t+\delta}) - W_t^2(\mu_t, \nu_t) \\
&\qquad + W_{t+\delta}^2(\mu_{t+\delta}, \nu_{t+\delta}) - W_t^2(\mu_{t+\delta}, \nu_{t+\delta})\Big] dt \\
&\ge \int_r^s \liminf_{\delta\searrow 0} \frac{1}{\delta}\Big(W_t^2(\mu_{t+\delta}, \nu_{t+\delta}) - W_t^2(\mu_t, \nu_t)\Big) dt \\
&+ \limsup_{\delta\searrow 0} \frac{1}{\delta} \int_r^s \Big(W_{t+\delta}^2(\mu_{t+\delta}, \nu_{t+\delta}) - W_t^2(\mu_{t+\delta}, \nu_{t+\delta})\Big) dt \\
&= \int_r^s \liminf_{\delta\searrow 0} \frac{1}{\delta}\Big(W_t^2(\mu_{t+\delta}, \nu_{t+\delta}) - W_t^2(\mu_t, \nu_t)\Big) dt \\
&+ \liminf_{\delta\searrow 0} \frac{1}{\delta} \int_r^s \Big(W_t^2(\mu_t, \nu_t) - W_{t-\delta}^2(\mu_t, \nu_t)\Big) dt \\
&\ge \int_r^s \Big[\liminf_{\delta\searrow 0} \frac{1}{\delta}\Big(W_t^2(\mu_{t+\delta}, \nu_{t+\delta}) - W_t^2(\mu_t, \nu_t)\Big) \\
&+ \liminf_{\delta\searrow 0} \frac{1}{\delta}\Big(W_t^2(\mu_t, \nu_t) - W_{t-\delta}^2(\mu_t, \nu_t)\Big)\Big] dt \ge 0,
\end{aligned}$$

where the last inequality is due to (62). This proves the claim. In the previous argumentation, we twice interchanged $\int \dots dr$ and $\liminf_\delta$ or $\limsup_\delta$ which is justified by Lebesgue's dominated convergence theorem since $\frac{1}{\delta}[W^2_{t+\delta} - W^2_t]$ is uniformly bounded (due to the log-Lipschitz bound on the distances). $\square$

Corollary 2.39. *Assume that $\big(X, d_t, m_t\big)_{t\in(0,T)}$ is a super-Ricci flow and that $(\mu_t)_{t\le\tau}$ and $(\nu_t)_{t\le\tau}$ are dual heat flows started in points μ_τ and $\nu_\tau \in \mathcal{P}$, resp., for some $\tau \in (0,T]$. Then for all $0 \le s < t \le \tau$*

$$W_s(\mu_s, \nu_s) \le W_t(\mu_t, \nu_t). \tag{63}$$

Proof. For measures μ_τ, ν_τ with densities in $Dom(\mathcal{E})$ which are bounded from above and bounded away from 0 the estimate (63) immediately follows from the previous theorem and the fact that the map $t \mapsto W_t(\mu_t, \nu_t)$ is absolutely continuous (Lemma 2.33).

The set of such probability measures is dense in $\mathcal{P}$ (w.r.t. weak topology) and according to Lemma 2.33, $\hat{P}_{t,s}$ is continuous on $\mathcal{P}$. Thus the estimate (63) carries over to all $\mu_\tau, \nu_\tau \in \mathcal{P}$. $\square$

Theorem 2.40 (“$\mathbf{(I}_N\mathbf{)} \Rightarrow \mathbf{(II}_N\mathbf{)}$”)**.** *Assume that $\big(X, d_t, m_t\big)_{t\in(0,T)}$ is a super-N-Ricci flow and that probability measures $\mu_\tau, \nu_\tau \in \mathcal{P}$ are given for some $\tau \in (0,T]$. Then the dual heat flows $(\mu_t)_{t\le\tau}$ and $(\nu_t)_{t\le\tau}$ starting in these points satisfy for all $0 \le s < t \le \tau$*

$$W_s^2(\mu_s, \nu_s) \le W_t^2(\mu_t, \nu_t) - \frac{2}{N} \int_s^t \left[S_r(\mu_r) - S_r(\nu_r)\right]^2 dr. \tag{64}$$

Proof. For measures μ_τ, ν_τ within the subset $\mathcal{D}$ we follow the proof of the previous Theorem 2.38 line by line and finally use the enforcement of the super Ricci flow property to deduce

$$\begin{aligned} -\frac{1}{2} \liminf_{\delta\to 0} \frac{1}{\delta} \Big[W^2_{t+\delta}(\mu_{t+\delta}, \nu_{t+\delta}) - W^2_t(\mu_{t+\delta}, \nu_{t+\delta})\Big] \quad &\le \quad \partial_a S_t(\overline{\eta}^{1-}_t) - \partial_a S_t(\overline{\eta}^{0+}_t) \\ &\quad -\frac{1}{N} \left[S_t(\mu_t) - S_t(\nu_t)\right]^2 \end{aligned}$$

instead of (61). Together with the other estimates from the proof of the previous theorem this gives

$$-\frac{1}{2} \liminf_{\delta\to 0} \frac{1}{\delta} \Big[W^2_{t+\delta}(\mu_{t+\delta}, \nu_{t+\delta}) - W^2_t(\mu_t, \nu_t)\Big] \quad \le \quad -\frac{1}{N} \left[S_t(\mu_t) - S_t(\nu_t)\right]^2.$$

Integrating this w.r.t. t yields the claim.

For general $\mu_\tau, \nu_\tau \in \mathcal{P}$ we apply the previous result to the pair $\mu_t, \nu_t \in \mathcal{D}$ (cf. Lemma 2.33) which already yields the claim for all $0 \le s < t < \tau$. The claim for $t = \tau$ now follows by approximation

$$\begin{aligned} W_s^2(\mu_s, \nu_s) \quad &\le \quad W_t^2(\mu_t, \nu_t) - \frac{2}{N} \int_s^t \left[S_r(\mu_r) - S_r(\nu_r)\right]^2 dr \\ &\to \quad W_\tau^2(\mu_\tau, \nu_\tau) - \frac{2}{N} \int_s^\tau \left[S_r(\mu_r) - S_r(\nu_r)\right]^2 dr \end{aligned}$$

as $t \uparrow \tau$. Here the convergence of the integrals is obvious. The convergence of the first term on the right-hand side follows from Lemma 2.33. $\square$

2.7.2 From Gradient Estimates to Transport Estimates

Theorem 2.41 **("$(\mathbf{III}_N) \Rightarrow (\mathbf{II}_N)$").** *Assume that* $\big(X, d_t, m_t\big)_{t\in(0,T)}$ *satisfies the Bakry-Ledoux gradient estimate* (III_N) *for the primal heat flow. Then the dual heat flow starting in arbitrary points* $\mu_\tau^0, \mu_\tau^1 \in \mathcal{P}(X)$ *satisfies for all* $0 < s < \tau < T$

$$W_s^2(\mu_s^0,\mu_s^1) \le W_\tau^2(\mu_\tau^0,\mu_\tau^1) - \frac{2}{N}\int_s^\tau \big[S_t(\mu_t^0) - S_t(\mu_t^1)\big]^2\, dt. \tag{65}$$

Proof. (i) Given $\tau \in I$ and a regular curve (see chapter 3) $(\mu_\tau^a)_{a\in[0,1]}$, define of each $t \le \tau$ the W_t-*action*

$$\mathcal{A}_t\big(\dot{\mu_t}\big) = \sup\left\{\sum_{i=1}^k \frac{1}{a_i - a_{i-1}}\, W_t^2\big(\mu_t^{a_{i-1}},\mu_t^{a_i}\big) :\ k\in\mathbb{N},\ 0 = a_0 < a_1 < \ldots < a_k = 1\right\}$$

of the curve $a \mapsto \mu_t^a = \hat{P}_{\tau,t}\mu_\tau^a$. Let $t \in (0,\tau]$ be given with $\mathcal{A}_t\big(\dot{\mu_t}\big) < \infty$. In other words, such that the curve $a \mapsto \mu_t^a$ is 2-absolutely continuous. (Obviously, this is true for $t = \tau$. The subsequent discussion indeed will show that this holds for all $t \le \tau$.) Let $(u_t^a)_{a\in[0,1]}$ and $(\Phi_t^a)_{a\in[0,1]}$ denote the densities and velocity potentials for the curve $(\mu_t^a)_{a\in[0,1]}$ (see [8, Theorem 8.2], or (41),(42)) in the static space (X, d_t, m_t). Then, in particular,

$$\mathcal{A}_t\big(\dot{\mu_t}\big) = \int_0^1 |\dot\mu_t^a|_{W_t}\, da = \int_0^1\int_X \big|\nabla_t\Phi_t^a\big|^2\, d\mu_t^a\, da.$$

Given $s \in (0,t)$ and $\epsilon > 0$ choose bounded Lipschitz functions $-\varphi_s^0, \varphi_s^1$ which are in W_s-duality to each other such that

$$W_s^2(\mu_s^0,\mu_s^1) \ \le\ 2\Big[\int_X \varphi_s^1 d\mu_s^1 - \int_X \varphi_s^0 d\mu_s^0\Big] + \epsilon(t-s)$$

and let $(\varphi_s^a)_{a\in[0,1]}$ denote the Hopf-Lax interpolation of φ_s^0, φ_s^1 in the static space (X, d_s, m_s).

Then applying the continuity equation (41) and the Hamilton-Jacobi equation (39) yields

$$\begin{aligned}
\epsilon &+ \frac{1}{t-s}\Big[\mathcal{A}_t(\dot{\mu_t}) - W_s^2(\mu_s^0,\mu_s^1)\Big]\\
&\ge \frac{1}{t-s}\int_0^1 \big|\dot\mu_t^a\big|^2 da - \frac{2}{t-s}\Big[\int_X \varphi_s^1 d\mu_s^1 - \int_X \varphi_s^0 d\mu_s^0\Big]\\
&= \frac{1}{t-s}\int_0^1 \Big[\int_X \big|\nabla_t\Phi_t^a\big|^2 d\mu_t^a - 2\partial_a \int_X P_{t,s}\varphi_s^a d\mu_t^a\Big] da\\
&= \frac{1}{t-s}\int_0^1\int_X \Big[\big|\nabla_t\Phi_t^a - \nabla_t P_{t,s}\varphi_s^a\big|^2 - \big|\nabla_t P_{t,s}\varphi_s^a\big|^2 + P_{t,s}\big|\nabla_s\varphi_s^a\big|^2\Big] d\mu_t^a da\\
&\ge \frac{1}{t-s}\int_0^1\int_X \big|\nabla_t\Phi_t^a - \nabla_t P_{t,s}\varphi_s^a\big|^2 d\mu_t^a\, da\\
&\qquad + \frac{2}{N(t-s)}\int_s^t\int_0^1\int_X \Big[P_{t,r}\Delta_r P_{r,s}\varphi_s^a\Big]^2 d\mu_t^a da\, dr \ \ge 0
\end{aligned}$$

where for the second last inequality we have used the Bakry-Ledoux gradient estimate (III_N).

In the case $N = \infty$ this already proves the claim. Indeed, since $\epsilon > 0$ was arbitrary it states that

$$W_s^2(\mu_s^0, \mu_s^1) \le \mathcal{A}_\tau(\dot{\mu_\tau})$$

for any regular curve $(\mu_\tau^a)_{a\in[0,1]}$. Given any $\mu_\tau^0, \mu_\tau^1 \in \mathcal{P}(X)$ we can choose regular curves $(\mu_{\tau,n}^a)_{a\in[0,1]}$ for $n \in \mathbb{N}$ such that $\mathcal{A}_\tau(\dot{\mu_{\tau,n}}) \to W_\tau^2(\mu_\tau^0, \mu_\tau^1)$ and $W_\tau(\mu_{\tau,n}^0, \mu_\tau^0) \to 0$ as well as $W_\tau(\mu_{\tau,n}^1, \mu_\tau^1) \to 0$ for $n \to \infty$. According to Lemma 2.33, the latter also implies $W_s(\mu_{s,n}^0, \mu_s^0) \to 0$ as well as $W_s(\mu_{s,n}^1, \mu_s^1) \to 0$ for $n \to \infty$ where $\mu_{s,n}^a := \hat{P}_{\tau,s}\mu_{\tau,n}^a$. Together with the previous estimate (applied with $t = \tau$ to the regular curves $(\mu_{\tau,n}^a)_{a\in[0,1]}$) we obtain

$$W_s^2(\mu_s^0, \mu_s^1) = \lim_{n\to\infty} W_s^2(\mu_{s,n}^0, \mu_{s,n}^1) \le \lim_{n\to\infty} \mathcal{A}_\tau(\dot{\mu_{\tau,n}}) = W_\tau^2(\mu_\tau^0, \mu_\tau^1).$$

This is the claim.

Moreover, applying this monotonicity result to each pair $\mu_\tau^{a_{i-1}}, \mu_\tau^{a_i}$ of points on the initial regular curve selected by an arbitrary partition $(a_i)_{i=1,\dots,k}$ yields

$$\mathcal{A}_s(\dot{\mu_s}) \le \mathcal{A}_\tau(\dot{\mu_\tau})$$

for all $s \le \tau$. In particular, this implies that the previous argumentation is valid for all $t \le \tau$.

(ii) Moreover, the previous estimates for given s, t, ϵ can be tightened up by choosing $k \in \mathbb{N}$ and $(a_i)_{i=1,\dots,k}$ as well as for $i = 1, \dots, k$ suitable bounded Lipschitz functions $-\varphi_s^{0,i}, \varphi_s^{1,i}$ which are in W_s-duality to each other and which are 'almost maximizers' of the dual representation of $W_s^2\big(\mu_s^{a_{i-1}}, \mu_s^{a_i}\big)$ such that

$$\begin{aligned}
\epsilon \quad &+ \quad \frac{1}{t-s}\Big[\mathcal{A}_t(\dot{\mu_t}) - \mathcal{A}_s(\dot{\mu_s})\Big] \\
&\ge \quad \epsilon/2 + \frac{1}{t-s}\Big[\mathcal{A}_t(\dot{\mu_t}) - \sum_{i=1}^k \frac{1}{a_i - a_{i-1}} W_s^2\big(\mu_s^{a_{i-1}}, \mu_s^{a_i}\big)\Big] \\
&\ge \quad \frac{1}{t-s}\int_0^1 \big|\dot{\mu}_t^a\big|^2 da - \frac{2}{t-s}\sum_{i=1}^k \frac{1}{a_i - a_{i-1}}\Big[\int_X \varphi_s^{1,i} d\mu_s^1 - \int_X \varphi_s^{0,i} d\mu_s^0\Big] \\
&= \quad \frac{1}{t-s}\int_0^1 \Big[\int_X \big|\nabla_t \Phi_t^a\big|^2 d\mu_t^a - 2\partial_a \int_X P_{t,s}\varphi_s^{a,k} d\mu_t^a\Big] da \\
&= \quad \frac{1}{t-s}\int_0^1 \int_X \Big[\big|\nabla_t \Phi_t^a - \nabla_t P_{t,s}\varphi_s^{a,k}\big|^2 - \big|\nabla_t P_{t,s}\varphi_s^{a,k}\big|^2 + P_{t,s}\big|\nabla_s \varphi_s^{a,k}\big|^2\Big] d\mu_t^a da \\
&\ge \quad \frac{1}{t-s}\int_0^1 \int_X \big|\nabla_t \Phi_t^a - \nabla_t P_{t,s}\varphi_s^{a,k}\big|^2 d\mu_t^a \, da \\
&\qquad + \frac{2}{N(t-s)}\int_s^t \int_0^1 \int_X \Big[P_{t,r}\Delta_r P_{r,s}\varphi_s^{a,k}\Big]^2 d\mu_t^a da\, dr \quad =: (\alpha)
\end{aligned}$$

The function $\varphi_s^{a,k}$ here is obtained for $a \in (a_{i-1}, a_i)$ by Hopf-Lax interpolation of the Lipschitz functions $\varphi_s^{a_{i-1}+,k} := \frac{1}{a_i - a_{i-1}}\varphi_s^{0,i}$ and $\varphi_s^{a_i-,k} := \frac{1}{a_i - a_{i-1}}\varphi_s^{1,i}$.

Now let us choose t to be a Lebesgue density point of $t \mapsto \int_0^1 \mathcal{E}_t(P_{t,s}\varphi_s^a, P^*_{\tau,t}u^a_\tau)\, da$. Then for s sufficiently close to t the commutator lemma (applied to time points r and t) implies that

$$\begin{aligned}
&\Big[\frac{1}{(t-s)}\int_s^t\int_0^1\int_X P_{t,r}\Delta_r P_{r,s}\varphi_s^{a,k} d\mu_t^a da\, dr\Big]^2\\
\geq&\Big[\frac{1}{(t-s)}\int_s^t\int_0^1\int_X \Delta_t P_{t,s}\varphi_s^{a,k} d\mu_t^a da\, dr\Big]^2 - \epsilon\cdot N/2.
\end{aligned}$$

Let us also briefly remark that the densities u^a_t of the measures μ^a_t are bounded away from 0, uniformly in a (due to the smooth dependence on a of the measures in the regularized curve we started with) and locally uniformly in t (due to the parabolic Harnack inequality for solutions to the adjoint heat equation). In particular, in the subsequent calculations the singularity of the logarithm at 0 does not matter. Thus

$$\begin{aligned}
(\alpha) \quad = \quad & \frac{1}{t-s}\int_0^1\int_X \big|\nabla_t\Phi_t^a - \nabla_t P_{t,s}\varphi_s^{a,k}\big|^2 d\mu_t^a\, da\\
&+\frac{2}{N}\Big|\int_0^1\int_X \nabla_t P_{t,s}\varphi_s^{a,k}\cdot\nabla_t \log u_t^a\, d\mu_t^a da\Big|^2 - \epsilon\\
\geq \quad & \frac{2}{N+\epsilon}\Big|\int_0^1\int_X \nabla_t \Phi_t^a\cdot\nabla_t \log u_t^a\, d\mu_t^a da\Big|^2 - \epsilon\\
&+\Big[\frac{1}{t-s} - \frac{2}{\epsilon}\int_X \big|\nabla_t \log u_t^a\big|^2 d\mu_t^a\, da\Big]\cdot\int_X \big|\nabla_t\Phi_t^a - \nabla_t P_{t,s}\varphi_s^{a,k}\big|^2 d\mu_t^a\, da\\
\geq \quad & \frac{2}{N+\epsilon}\Big|\int_0^1\int_X \nabla_t \Phi_t^a\cdot\nabla_t \log u_t^a\, d\mu_t^a da\Big|^2 - \epsilon \;=:\; (\beta)
\end{aligned}$$

provided s is sufficiently close to t. Finally, using the continuity equation for the curve $(\mu^a_t)_{a\in[0,1]}$ (and its velocity potentials Φ^a_t) we obtain

$$(\beta) \quad = \quad \frac{2}{N+\epsilon}\Big|S_t(\mu_t^1) - S_t(\mu_t^0)\Big|^2 - \epsilon.$$

Passing to the limit $s \nearrow t$ yields

$$\epsilon \quad + \quad \partial^-_{t-}\mathcal{A}_t(\mu^\cdot_t) \geq \frac{2}{N+\epsilon}\Big|S_t(\mu_t^1) - S_t(\mu_t^0)\Big|^2 - \epsilon$$

and thus (since $\epsilon > 0$ was arbitrary)

$$\partial^-_{t-}\mathcal{A}_t(\mu^\cdot_t) \geq \frac{2}{N}\Big|S_t(\mu_t^1) - S_t(\mu_t^0)\Big|^2. \tag{66}$$

Recall that this holds for a.e. $t \in (0,\tau)$. Moreover, note that $t \mapsto \mathcal{A}_t(\mu^\cdot_t)$ is absolutely continuous. Indeed, by Lemma 2.33 and the log-Lipschitz assumption

(36)

$$\begin{aligned}
\left|W^2_{t+\epsilon}(\mu^a_{t+\epsilon},\mu^b_{t+\epsilon}) - W^2_t(\mu^a_t,\mu^b_t)\right| &\le \left|W^2_{t+\epsilon}(\mu^a_{t+\epsilon},\mu^b_t) - W^2_t(\mu^a_t,\mu^b_t)\right| \\
&\quad + \left|W^2_t(\mu^a_{t+\epsilon},\mu^b_{t+\epsilon}) - W^2_t(\mu^a_t,\mu^b_t)\right| \\
&\le 2L\epsilon\, e^{2L\epsilon}\, W^2_t(\mu^a_t,\mu^b_t) + \frac{2\sqrt{\epsilon}}{1-2\sqrt{\epsilon}} W^2_t(\mu^a_t,\mu^b_t) \\
&\quad + \frac{1}{\sqrt{\epsilon}} W^2_t(\mu^a_{t+\epsilon},\mu^a_t) + \frac{1}{\sqrt{\epsilon}} W^2_t(\mu^b_{t+\epsilon},\mu^b_t) \\
&\le C_0\sqrt{\epsilon}\, W^2_t(\mu^a_t,\mu^b_t) + C_1\sqrt{\epsilon}.
\end{aligned}$$

Thus we may integrate (66) from any $s \in (0,\tau)$ to τ to obtain

$$\mathcal{A}_s(\dot{\mu_s}) \le \mathcal{A}_\tau(\dot{\mu_\tau}) - \frac{2}{N}\int_s^\tau \left[S_t(\mu^0_t) - S_t(\mu^1_t)\right]^2 dt. \tag{67}$$

Finally, given arbitrary $\mu^0_\tau, \mu^1_\tau \in \mathcal{P}(X)$ the subsequent lemma provides a construction of 2-absolutely continuous, regular curves $(\tilde{\mu}^a_\sigma)_{a\in[0,1]}$ connecting $\mu^0_\sigma, \mu^1_\sigma$ for a.e. $\sigma < \tau$ with

$$\mathcal{A}_\sigma(\dot{\tilde{\mu}_\sigma}) \to W^2_\tau(\mu^0_\tau,\mu^1_\tau)$$

as $\sigma \nearrow \tau$. Carrying out the previous estimations, finally resulting in (67), with $(\tilde{\mu}^a_\sigma)_{a\in[0,1]}$ in the place of $(\mu^a_\tau)_{a\in[0,1]}$ yields

$$\begin{aligned}
W^2_s(\mu^0_s,\mu^1_s) &\le \mathcal{A}_s(\dot{\tilde{\mu}_s}) \\
&\le \mathcal{A}_\sigma(\dot{\tilde{\mu}_\sigma}) - \frac{2}{N}\int_s^\sigma \left[S_t(\mu^0_t) - S_t(\mu^1_t)\right]^2 dt \\
&\to W^2_\tau(\mu^0_\tau,\mu^1_\tau) - \frac{2}{N}\int_s^\tau \left[S_t(\mu^0_t) - S_t(\mu^1_t)\right]^2 dt.
\end{aligned}$$

This proves the claim. □

Lemma 2.42. *(i) Assume* (III) *(with $N=\infty$) and let $(\mu^a)_{a\in[0,1]}$ be an arbitrary W_τ-geodesic in $\mathcal{P}(X)$. Let χ be a standard convolution kernel on $\mathbb{R}$. Then for a.e. $t<\tau$ and every $\delta>0$ the measures*

$$\mu^{a,\delta}_t := \int_{\mathbb{R}} \left(\hat{P}_{\tau,t}\mu^{\vartheta(a)+\delta b}\right)\chi(b)db = \hat{P}_{\tau,t}\left(\int_{\mathbb{R}} \mu^{\vartheta(a)+\delta b}\chi(b)db\right)$$

constitute a regular curve $(\mu^{a,\delta}_t)_{a\in[0,1]}$ (in the sense of Definition 2.27). Here $\vartheta(a)=0$ for $a\in[0,\delta]$, $\vartheta(a)=1$ for $a\in[1-\delta,1]$, and $\vartheta(a)=\frac{a-\delta}{1-2\delta}$ for $a\in[\delta,1-\delta]$.

Choosing $t_n \nearrow \tau$ and $\delta_n \searrow 0$ yields a sequence of regular curves satisfying (44) - (47)*. In addition, for these approximations the endpoints are simply given by the dual heat flow:*

$$\mu^{a,\delta_n}_{t_n} = \hat{P}_{\tau,t_n}\mu^a$$

for $a=0$ as well as $a=1$ and for all n.

Proof. The re-parametrization by means of ϑ forces the curve to be constant for some short interval around the endpoints and squeeze it in-between. The latter leads to a moderate increase of the metric speed. The former guarantees that the endpoints remain unchanged under the convolution. The convolution w.r.t. the kernel χ guarantees smooth dependence on a, i.e. (1) of Def 2.27. (44) follows from Lemma 2.33. Smoothness in a (thanks to the convolution) and Hölder continuity in (t,x) (being a solution to the adjoint heat equation) guarantee uniform boundedness of $u_t^a(x)$ for $(a,t,x) \in [0,1] \times (0,t] \times X$ for each $t < \tau$, i.e. (2) of Def 2.27. Moreover, $u_t^a(x)$ is uniformly bounded away from 0. Thus (3) of Def 2.27 is equivalent to a uniform bound for the energy $\mathcal{E}_t(u^a)$.

Boundedness of u_r^a for $r < \tau$ implies

$$\int_0^1 \int_0^r \mathcal{E}_t(u_t^a)\, dt\, da \leq \frac{1}{2} \int_0^1 \|u_r^a\|^2_{L^2(m_r)} da < \infty.$$

Thus for a.e. $t < \tau$

$$\int_0^1 \mathcal{E}_t(u_t^a) da < \infty \quad \text{and} \quad \mathcal{E}_t(u_t^0) < \infty, \quad \mathcal{E}_t(u_t^1) < \infty.$$

Convolution w.r.t. the kernel χ thus turns the integrable function $a \mapsto \mathcal{E}_t\left(u_t^{\vartheta(a)}\right)$ into a bounded function: $\int_{\mathbb{R}} \mathcal{E}_t\left(u_t^{\vartheta(a+\delta b)}\right) \chi(b) db \leq C$. Since the energy $u \mapsto \mathcal{E}_t(u)$ is convex, Jensen's inequality implies

$$\mathcal{E}_t\left(\int_{\mathbb{R}} u_t^{\vartheta(a+\delta b)}\, \chi(b) db\right) \leq \int_{\mathbb{R}} \mathcal{E}_t\left(u_t^{\vartheta(a+\delta b)}\right) \chi(b) db \leq C.$$

The action estimate (45) follows from part (i) of the previous proof. Indeed, the dual heat flow decreases the action. Also convolution in the a-parameter decreases the action. The re-parametrization increases the action by a factor bounded by $\frac{1}{(1-2\delta)^2}$.

The entropy estimates (46) and (47) follow as in the proof of Lemma 2.28 $\square$

2.7.3 Duality between Transport and Gradient Estimates in the Case $N = \infty$

In the subsequent chapter, we will prove the implication $(\mathbf{II}_N) \Rightarrow (\mathbf{III}_N)$ by composing the results $(\mathbf{II}_N) \Rightarrow (\mathbf{IV}_N)$ and $(\mathbf{IV}_N) \Rightarrow (\mathbf{III}_N)$. Partly, these arguments are quite involved. (And actually, for the last one, we freely make use of the subsequent Theorem 2.43).

Here we present a direct, much simpler proof in the particular case $N = \infty$. Indeed, this proof will yield a slightly stronger statement: the equivalence of the respective estimates for given pairs s, t. See also [37] for a related result.

Theorem 2.43 ("**(II)** $\Leftrightarrow$ **(III)**")**.** *For fixed* $0 < s < t < T$ *the following are equivalent:*

(II)$_{t,s}$ *For all* $\mu, \nu \in \mathcal{P}$

$$W_s(\hat{P}_{t,s}\mu, \hat{P}_{t,s}\nu) \leq W_t(\mu, \nu) \tag{68}$$

(III)$_{t,s}$ For all $u \in Dom(\mathcal{E})$

$$\Gamma_t(P_{t,s}u) \le P_{t,s}(\Gamma_s(u)) \quad m\text{-a.e. on } X. \tag{69}$$

Proof. "(II)$_{t,s}$ $\Rightarrow$ (III)$_{t,s}$": Given a bounded Lipschitz function u on X, points $x, y \in X$, and a d_t-geodesic $(\gamma^a)_{a\in[0,1]}$ connecting x and y, put $\mu_t^a = \delta_{\gamma^a}$ and $\mu_t^a = \hat{P}_{t,s}\mu_t^a$. The transport estimate $W_s(\mu_s^a, \mu_s^b) \le W_t(\mu_t^a, \mu_t^b)$ implies that

$$\big|\dot{\mu}_s\big|_{W_s} \le \big|\dot{\mu}_t\big|_{W_t} = \big|\dot{\gamma}\big|_{d_t} = d_t(x,y).$$

Thus following the argumentation from [6], Theorem 6.4, we obtain

$$\begin{aligned}
&\Big|P_{t,s}u(x) - P_{t,s}u(y)\Big| = \Big|\int u\, d\hat{P}_{t,s}\delta_x - \int u\, d\hat{P}_{t,s}\delta_y\Big| \\
\le\ & \int_0^1 \Big(\big|\nabla_s u\big|^2 d\mu_s^a\Big)^{1/2} \cdot \big|\dot{\mu}_s\big|_{W_s} da \\
\le\ & \int_0^1 \Big(P_{t,s}\big|\nabla_s u\big|^2(\gamma^a)\Big)^{1/2} \cdot \big|\dot{\gamma}\big|_{d_t} da \\
\le\ & d_t(x,y)\cdot \sup\Big\{P_{t,s}\big|\nabla_s u\big|^2(z):\ d_t(x,z)+d_t(z,y)=d_t(x,y)\Big\}.
\end{aligned}$$

The Hölder continuity of $z \mapsto P_{t,s}\big|\nabla_s u\big|^2(z)$, therefore, allows to conclude that $(P_{t,s}\big|\nabla_s u\big|^2)^{1/2}$ is an upper gradient for $P_{t,s}u$. This proves the claim for bounded Lipschitz functions. The extension to $u \in Dom(\mathcal{E})$ follows as in [6].

"(III)$_{t,s}$ $\Rightarrow$ (II)$_{t,s}$": previous Theorem. □

2.8 From Transport Estimates to Gradient Estimates and Bochner Inequality

As before, for the sequel a time-dependent mm-space $(X, d_t, m_t)_{t\in I}$ will be given such that

- for each $t \in I$ the static space satisfies the RCD$^*(K, N')$ condition for some finite numbers K and N'
- the distances are bounded and log-Lipschitz in t, that is, $|\partial_t d_t(x,y)| \le L \cdot d_t(x,y)$ for some L uniformly in t, x, y (existence of $\partial_t d_t$ for a.e. t)
- f is L-Lipschitz in t and x.

2.8.1 The Bochner Inequality

The Time-Derivative of the Γ-Operator

Definition 2.44. *Given an interval $J \subset I$ and $u \in \mathcal{F}_J$ with $\Gamma_r(u_r)(x) \le C$ uniformly in $(r,x) \in J \times X$. Then we define $\overset{\bullet}{\Gamma}_r(u_r)(x)$ as (one of the) weak subsequential limit(s) of*

$$\frac{1}{2\delta}\Big[\Gamma_{r+\delta}(u_r) - \Gamma_{r-\delta}(u_r)\Big](x) \tag{70}$$

in $L^2(J \times X)$ *for* $\delta \to 0$. *That is, for a suitable 0-sequence* $(\delta_n)_n$ *and all* $g \in L^2(J \times X)$

$$\frac{1}{2\delta_n} \int_J \int_X \Big[\Gamma_{r+\delta_n}(u_r) - \Gamma_{r-\delta_n}(u_r)\Big]\, g_r\, dm_r\, dr \to \int_J \int_X \overset{\bullet}{\Gamma}_r\,(u_r)\, g_r\, dm_r\, dr$$

as $n \to \infty$.

Actually, thanks to Banach-Alaoglu theorem, such a weak limit always exists since (70) – due to the log-Lipschitz continuity of the distances – defines a family of functions in $L^2(J \times X)$ with bounded norm. Thus in particular we will have

$$\begin{aligned} &\liminf_{\delta \to 0} \frac{1}{2\delta} \int_J \int_X \Big[\Gamma_{r+\delta}(u_r) - \Gamma_{r-\delta}(u_r)\Big]\, g_r\, dm_r\, dr \\ &\leq \int_J \int_X \overset{\bullet}{\Gamma}_r\,(u_r)\, g_r\, dm_r\, dr \\ &\leq \limsup_{\delta \to 0} \frac{1}{2\delta} \int_J \int_X \Big[\Gamma_{r+\delta}(u_r) - \Gamma_{r-\delta}(u_r)\Big]\, g_r\, dm_r\, dr. \end{aligned} \tag{71}$$

Remark 2.45. *All the subsequent statements involving* $\overset{\bullet}{\Gamma}_r\,(u_r)$ *will be independent of the choice of the sequence* $(\delta_n)_n$ *and of the accumulation point in* $L^2(J \times X)$. *For instance, the precise meaning of Theorem 2.3 is that each of the properties* **(I)**, **(II)** *or* **(III)** *will imply* **(IV)** *for* every *choice of the weak subsequential limit* $\overset{\bullet}{\Gamma}_r\,(u_r)$. *Conversely, if* **(IV)** *is satisfied for* some *choice of the weak subsequential limit* $\overset{\bullet}{\Gamma}_r\,(u_r)$ *then it implies properties* **(I)**, **(II)** *and* **(III)**. *Indeed, the only property of* $\overset{\bullet}{\Gamma}_r\,(u_r)$ *which enters the calculations is* (71).

Note that the log-Lipschitz continuity of the distances also immediately implies that

$$\Big|\overset{\bullet}{\Gamma}_r\,(u_r)\Big| \leq 2L \cdot \Gamma_r(u_r). \tag{72}$$

Lemma 2.46. *For every* $u \in \mathcal{F}_J$ *with* $\sup_{r,x} \Gamma_r(u_r)(x) < \infty$ *and every* $g \in L^\infty(J \times X)$

$$\int_J \int_X \overset{\bullet}{\Gamma}_r\,(u_r)\, g_r\, dm_r\, dr = \lim_{n \to \infty} \frac{1}{\delta_n} \int_J \int_X \Big[\Gamma_{r+\delta_n}(u_r, u_{r+\delta_n}) - \Gamma_r(u_r, u_{r+\delta_n})\Big]\, g_r\, dm_r\, dr.$$

In particular,

$$\begin{aligned} &\liminf_{\delta \searrow 0} \frac{1}{\delta} \int_J \int_X \Big[\Gamma_{r+\delta}(u_{r+\delta}, u_r) - \Gamma_r(u_{r+\delta}, u_r)\Big]\, g_r\, dm_r\, dr \\ &\leq \int_J \int_X \overset{\bullet}{\Gamma}_r\,(u_r)\, g_r\, dm_r\, dr \\ &\leq \limsup_{\delta \searrow 0} \frac{1}{\delta} \int_J \int_X \Big[\Gamma_{r+\delta}(u_{r+\delta}, u_r) - \Gamma_r(u_{r+\delta}, u_r)\Big]\, g_r\, dm_r\, dr. \end{aligned}$$

Proof.

$$\int_J\int_X \overset{\bullet}{\Gamma}_r\,(u_r)\,g_r\,dm_r\,dr = \lim_{n\to\infty}\Big(\frac{1}{2\delta_n}\int_J\int_X \Big[\Gamma_{r+\delta_n}(u_r)-\Gamma_r(u_r)\Big]\,g_r\,dm_r\,dr$$
$$+\frac{1}{2\delta_n}\int_J\int_X \Big[\Gamma_r(u_r)-\Gamma_{r-\delta_n}(u_r)\Big]\,g_r\,dm_r\,dr\Big)$$
$$=\lim_{n\to\infty}\Big(\frac{1}{2\delta_n}\int_J\int_X \Big[\Gamma_{r+\delta_n}(u_r)-\Gamma_r(u_r)\Big]\,g_r\,dm_r\,dr$$
$$+\frac{1}{2\delta_n}\int_J\int_X \Big[\Gamma_{r+\delta_n}(u_{r+\delta_n})-\Gamma_r(u_{r+\delta_n})\Big]\,g_r\,dm_r\,dr\Big)$$
$$=\lim_{n\to\infty}\Big(\frac{1}{\delta_n}\int_J\int_X \Big[\Gamma_{r+\delta_n}(u_r,u_{r+\delta_n})-\Gamma_r(u_r,u_{r+\delta_n})\Big]\,g_r\,dm_r\,dr$$
$$+\frac{1}{2\delta_n}\int_J\int_X \Big[\Gamma_{r+\delta_n}(u_{r+\delta_n}-u_r)-\Gamma_r(u_{r+\delta_n}-u_r)\Big]\,g_r\,dm_r\,dr\Big)$$
$$=\lim_{n\to\infty}\frac{1}{\delta_n}\int_J\int_X \Big[\Gamma_{r+\delta_n}(u_r,u_{r+\delta_n})-\Gamma_r(u_r,u_{r+\delta_n})\Big]\,g_r\,dm_r\,dr.$$

Here for the second equality we used index shift and Lusin's theorem (to replace $g_{r+\delta_n}dm_{r+\delta_n}$ again by $g_r dm_r$). The last equality follows from the log-Lipschitz continuity of $r\mapsto d_r$ which allows to estimate

$$\begin{aligned}
\frac{1}{\delta}\Big|\int_J\int_X &\Big[\Gamma_{r+\delta}(u_{r+\delta}-u_r)-\Gamma_r(u_{r+\delta}-u_r)\Big]\,g_r\,dm_r\,dr\Big|\\
&\le 2L\cdot\int_J\int_X \Gamma_r(u_{r+\delta}-u_r)\,g_r\,dm_r\,dr\\
&\le C'\cdot\int_J \mathcal{E}_r(u_{r+\delta}-u_r)dr\to 0
\end{aligned}$$

as $\delta\to 0$ since $r\mapsto u_r$, as a map from J to $\mathcal{F}$, is 'nearly continuous' (Lusin's theorem). $\square$

The Distributional Γ_2-Operator

Definition 2.47. *For $r\in(0,T)$ and $u\in Dom(\Delta_r)$ with $|\nabla_r u|\in L^\infty$ we define the distribution valued Γ_2-operator as a continuous linear operator*

$$\mathbf{\Gamma}_{2,r}(u):\mathcal{F}\cap L^\infty\to\mathbb{R}$$

by

$$\mathbf{\Gamma}_{2,r}(u)(g):=\int\Big[-\frac{1}{2}\Gamma_r\big(\Gamma_r(u),g\big)+(\Delta_r u)^2 g+\Gamma_r(u,g)\Delta_r u\Big]dm_r. \tag{73}$$

Note that

$$\begin{aligned}
\Big|\mathbf{\Gamma}_{2,r}(u)(g)\Big| &\le 2\|\nabla_r u\|_\infty\cdot\|\nabla_r^2 u\|_2\cdot\|\nabla_r g\|_2+\|g\|_\infty\cdot\|\Delta_r u\|_2^2\\
&\quad+\|\nabla_r u\|_\infty\cdot\|\nabla_r g\|_2\cdot\|\Delta_r u\|_2\\
&\le \|g\|_\infty\cdot\|\Delta_r u\|_2^2+C\cdot\|\nabla_r u\|_\infty\cdot\|\nabla_r g\|_2\cdot(\|\Delta_r u\|_2+\|u\|_2)
\end{aligned}$$

thanks to the fact that $\|\nabla_r^2 u\|_2^2\le(1+K_-)\cdot(\|\Delta_r u\|_2^2+\|u\|_2^2)$, cf. (50).

Also note that the assumptions on u will be preserved under the heat flow (at least for a.e. r) and the assumptions on g are preserved under the adjoint heat flow. If u is sufficiently regular (i.e. $\Delta u \in Dom(\mathcal{E}_r)$ and $|\nabla_r u|^2 \in Dom(\Delta_r)$) then obviously

$$\mathbf{\Gamma}_{2,r}(u)(g) = \int \Gamma_{2,r}(u) \cdot g \, dm_r$$

for all g under consideration where as usual $\Gamma_{2,r}(u) = \frac{1}{2}\Delta_r|\nabla_r u|^2 - \Gamma_r(u, \Delta_r u)$.

On the other hand, if $g \in Dom(\Delta_r)$ then in (73) we may replace the term $-\Gamma_r\big(\Gamma_r(u), g\big)$ by $\Gamma_r(u)\Delta_r g$.

The Bochner Inequality – Various Versions

Definition 2.48. *(i) We say that $(X, d_t, m_t)_{t\in I}$ satisfies the dynamic Bochner inequality with parameter $N \in (0, \infty]$ if for all $0 < s < t < T$ and for all $u_s, g_t \in \mathcal{F}$ with $g_t \geq 0$, $g_t \in L^\infty$, $u_s \in \mathrm{Lip}(X)$ and for a.e. $r \in (s,t)$*

$$\mathbf{\Gamma}_{2,r}(u_r)(g_r) \geq \frac{1}{2}\int \overset{\bullet}{\Gamma}_r (u_r) g_r dm_r + \frac{1}{N}\Big(\int \Delta_r u_r g_r dm_r\Big)^2 \tag{74}$$

where $u_r = P_{r,s}u_s$ and $g_r = P^*_{t,r}g_t$, cf. (21).

(ii) We say that $(X, d_t, m_t)_{t\in I}$ satisfies property $(\mathbf{IV}_N)$ *if it satisfies the dynamic Bochner inequality with parameter N as above and in addition the regularity assumption* (17) *is satisfied, i.e. $u_r \in \mathrm{Lip}(X)$ for all $r \in (s,t)$ with* $\sup_{r,x} \mathrm{lip}_r u_r(x) < \infty$.

Note that in the case $N = \infty$ inequality (74) simply states that

$$\mathbf{\Gamma}_{2,r}(u_r) \geq \frac{1}{2} \overset{\bullet}{\Gamma}_r (u_r) m_r$$

as inequality between distributions, tested against nonnegative functions g_r as above.

2.8.2 From Bochner Inequality to Gradient Estimates

Theorem 2.49 ("$(\mathbf{IV_N}) \Rightarrow (\mathbf{III_N})$"). *Suppose that the mm-space $(X, d_t, m_t)_{t\in I}$ satisfies the dynamic Bochner inequality* (74) *and the regularity assumption from Definition 2.48 (ii). Then for a.e. $x \in X$*

$$\Gamma_t(P_{t,s}u)(x) - P_{t,s}\Gamma_s(u)(x) \leq -\frac{2}{N}\int_s^t \big[P_{t,r}\Delta_r u_r(x)\big]^2 dr. \tag{75}$$

Proof. Given $s, t \in (0, T)$ as well as $u \in \mathrm{Lip}(X)$ and $g \in \mathcal{F} \cap L^\infty$ with $g \geq 0$, put $u_r = P_{r,s}u$, $g_r = P^*_{t,r}g$ for $r \in [s,t]$ and consider the function

$$h_r := \int g_r \Gamma_r(u_r) dm_r = \int \Gamma_r(u_r) d\mu_r$$

with $\mu_r := g_r \, m_r$.

(a) Choose $s \leq \sigma < \tau \leq t$ such that

$$h_\tau \leq \liminf_{\delta \searrow 0} \frac{1}{\delta}\int_{\tau-\delta}^{\tau} h_r dr \quad \text{and} \quad h_\sigma \geq \limsup_{\delta\searrow 0} \frac{1}{\delta}\int_\sigma^{\sigma+\delta} h_r dr. \tag{76}$$

Note that by Lebesgue's density theorem, the latter is true at least for a.e. $\sigma \geq s$ and for a.e. $\tau \leq t$. (Moreover, at the end of this proof (as part (b)) we will present an argument which allows to conclude that (76) holds for $\sigma = s, \tau = t$.) Then

$$\begin{aligned}
h_\tau - h_\sigma &\leq \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau-\delta} \big[h_{r+\delta} - h_r\big] dr \\
&\leq \limsup_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau-\delta} \int_X \Gamma_{r+\delta}(u_{r+\delta}) d(\mu_{r+\delta} - \mu_r)\, dr \\
&+ \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau-\delta} \int_X g_r \Big[\Gamma_{r+\delta}(u_{r+\delta}, u_r) - \Gamma_r(u_{r+\delta}, u_r)\Big] dm_r\, dr \\
&+ \limsup_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau-\delta} \int_X g_r \Big[\Gamma_{r+\delta}(u_{r+\delta}, u_{r+\delta} - u_r) + \Gamma_r(u_{r+\delta} - u_r, u_r)\Big] dm_r\, dr \\
&=: (I) + (II) + (III') + (III'').
\end{aligned}$$

Each of the four terms will be considered separately. Since $r \mapsto \mu_r$ is a solution to the dual heat equation, we obtain

$$\begin{aligned}
(I) &= \limsup_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau-\delta} \int_X \Gamma_{r+\delta}(u_{r+\delta}) \cdot \Big(- \int_r^{r+\delta} \Delta_q g_q\, dm_q\, dq\Big) dr \\
&= - \liminf_{\delta \searrow 0} \int_{\sigma+\delta}^{\tau} \int_X \Gamma_r(u_r) \Big(\frac{1}{\delta} \int_{r-\delta}^{r} \Delta_q g_q e^{-f_q}\, dq\Big) dm_\diamond\, dr \\
&= - \int_\sigma^\tau \int_X \Gamma_r(u_r) \cdot \Delta_r g_r\, dm_r\, dr
\end{aligned}$$

due Lebesgue's density theorem applied to $r \mapsto \Delta_r g_r e^{-f_r}$. Note that the latter function is in L^2 (Theorem 2.20) and the function $r \mapsto \Gamma_r(u_r)$ is in L^∞ thanks to Definition 2.48 (ii).

The second term can easily estimated in terms $\overset{\bullet}{\Gamma}_r$ according to Lemma 2.46:

$$\begin{aligned}
(II) &= \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau-\delta} \int_X g_r \Big[\Gamma_{r+\delta}(u_{r+\delta}, u_r) - \Gamma_r(u_{r+\delta}, u_r)\Big] dm_r\, dr \\
&\leq \int_\sigma^\tau \int_X g_r\ \overset{\bullet}{\Gamma}_r\ (u_r) dm_r dr.
\end{aligned}$$

The term (III') is transformed as follows

$$\begin{aligned}
&(III') \\
&= - \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau-\delta} \int_X \Big(\Gamma_{r+\delta}(g_r, u_{r+\delta}) + g_r\, \Delta_{r+\delta} u_{r+\delta}\Big) \cdot \Big(\int_r^{r+\delta} \Delta_q u_q\, dq\Big) dm_r\, dr \\
&= - \liminf_{\delta \searrow 0} \int_{\sigma+\delta}^{\tau} \int_X \Big(\Gamma_r(g_{r-\delta}, u_r) + g_{r-\delta}\, \Delta_r u_r\Big) \cdot \Big(\frac{1}{\delta} \int_{r-\delta}^{r} \Delta_q u_q\, dq\Big) dm_r\, dr \\
&= - \int_\sigma^\tau \int_X \Big(\Gamma_r(g_r, u_r) + g_r\, \Delta_r u_r\Big) \cdot \Delta_r u_r\, dm_r\, dr.
\end{aligned}$$

Here again we used Lebesgue's density theorem (applied to $r \mapsto \Delta_r u_r$) and the 'nearly continuity' of $r \mapsto g_r$ as map from (s,t) into $L^2(X,m)$ and as map into

$\mathcal{F}$ (Lusin's theorem). Moreover, we used the boundedness (uniformly in r and x) of g_r and of $\nabla_r u_r$ as well as the square integrability of $\Delta_r u_r$.

Similarly, the term (III'') will be transformed:

$$(III'') = -\liminf_{\delta\searrow 0} \frac{1}{\delta}\int_\sigma^{\tau-\delta}\int_X \Big(\Gamma_r(g_r,u_r)+g_r\,\Delta_r u_r\Big)\cdot\Big(\int_r^{r+\delta}\Delta_q u_q\,dq\Big)dm_r\,dr$$
$$= -\int_\sigma^\tau\int_X \Big(\Gamma_r(g_r,u_r)+g_r\,\Delta_r u_r\Big)\cdot\Big(\Delta_r u_r\Big)dm_r\,dr.$$

Summarizing and then using (74), we therefore obtain

$$h_\tau - h_\sigma = (I)+(II)+(III')+(III'')$$
$$\leq \int_\sigma^\tau\int_X\Big[-\Gamma_r(u_r)\cdot\Delta_r g_r + g_r\,\overset{\bullet}{\Gamma}_r(u_r) - 2\big(\Gamma_r(g_r,u_r)+g_r\,\Delta_r u_r\big)\,\Delta_r u_r\Big]dm_r\,dr$$
$$\leq -\frac{2}{N}\int_\sigma^\tau\Big[\int_X \Delta_r u_r\,g_r\,dm_r\Big]^2 dr = -\frac{2}{N}\int_\sigma^\tau\Big[\int_X P_{\tau,r}\Delta_r u_r\,g\,dm_\tau\Big]^2 dr.$$

Thus

$$\int_X \Gamma_\tau(P_{\tau,\sigma}u)g\,dm_\tau - \int_X P_{\tau,\sigma}\Gamma_\sigma(u)\,g\,dm_\tau \leq -\frac{2}{N}\int_\sigma^\tau\Big[\int_X P_{\tau,r}\Delta_r u_r\,g\,dm_\tau\Big]^2 dr. \tag{77}$$

(b) Recall that, given u and g, this holds for a.e. τ and a.e. σ. Now let us forget for the moment the term with N. Choosing g's from a dense countable set one may achieve that the exceptional sets for σ and τ in (77) do not depend on g. Next we may assume that $\sigma,\tau\in[s,t]$ with $\sigma<\tau$ is chosen such that (77) with $N=\infty$ simultaneously holds for all u from a dense countable set $\mathcal{C}_1$ in $\mathrm{Lip}(X)$. Approximating arbitrary $u\in\mathrm{Lip}(X)$ by $u_n\in\mathcal{C}_1$ yields

$$\int_X \Gamma_\tau(P_{\tau,\sigma}u)g\,dm_\tau - \int_X P_{\tau,\sigma}\Gamma_\sigma(u)\,g\,dm_\tau$$
$$\leq \liminf_n \int_X \Gamma_\tau(P_{\tau,\sigma}u_n)g\,dm_\tau - \lim_n\int_X P_{\tau,\sigma}\Gamma_\sigma(u_n)\,g\,dm_\tau \leq 0.$$

due to lower semicontinuity of the weighted energy on L^2. In other words, we have derived the gradient estimate **(III)** for almost all times σ and τ. Thanks to Theorem 2.43 this implies the transport estimate **(II)** for these time instances. But both sides of the transport estimate are continuous in time (thanks to the continuity of $r\mapsto W_r$ and the continuity of the dual heat flow). This implies that the transport estimate holds for all $\sigma,\tau\in[s,t]$ with $\sigma<\tau$. In particular, it holds for $\sigma=s$ and $\tau=t$. Again by Theorem 2.43 it yields the gradient estimate for given s and t and thus our initial assumption (76) is satisfied for the choice $\sigma=s$ and $\tau=t$.

(c) Taking this into account, we may conclude that (77) (for given N) holds with the choice $\sigma=s$ and $\tau=t$. Finally, choosing sequences of g's which approximate the Dirac distribution at a given $x\in X$ then implies that for all $u\in\mathrm{Lip}(X)$

$$\Gamma_t(P_{t,s}u)(x) - P_{t,s}\Gamma_s(u)(x) \leq -\frac{2}{N}\int_s^t\big[P_{t,r}\Delta_r u_r(x)\big]^2 dr \tag{78}$$

for a.e. $x\in X$. This proves the claim for bounded Lipschitz functions. The extension to $u\in Dom(\mathcal{E})$ follows as in [6]. $\square$

2.8.3 From Gradient Estimates to Bochner Inequality

In the previous chapter and the previous sections of this chapter, we have proven the implications $(\mathbf{III}_N) \Rightarrow (\mathbf{II}_N)$ and $(\mathbf{IV}_N) \Rightarrow (\mathbf{III}_N)$. Taking the subsequent section into account, where we show $(\mathbf{II}_N) \Rightarrow (\mathbf{IV}_N)$, we already have proven that $(\mathbf{III}_N) \Rightarrow (\mathbf{IV}_N)$. In the sequel, we will present another, more direct proof for this implication.

Theorem 2.50 ("$(\mathbf{III_N}) \Rightarrow (\mathbf{IV_N})$"). *Suppose that the mm-space $(X, d_t, m_t)_{t\in I}$ satisfies the gradient estimate* (75). *Then the dynamic Bochner inequality* (74) *holds true as well as the regularity assumption from Definition 2.48 (ii).*

Proof. Assume that the gradient estimate $(\mathbf{III}_N)$ holds true. It immediately implies the regularity assumption (17). To derive the dynamic Bochner inequality, let $s, t \in (0, T)$ as well as $u \in \mathrm{Lip}(X)$ and $g \in \mathcal{F} \cap L^\infty$ with $g \geq 0$ be given. Put $u_r = P_{r,s}u$, $g_r = P^*_{t,r}g$ for $r \in [s, t]$ and as before consider the function

$$h_r := \int g_r \Gamma_r(u_r) dm_r.$$

Then $(\mathbf{III}_N)$ implies that for all $s < \sigma < \tau < t$

$$\begin{aligned}
h_\tau - h_\sigma \leq & \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau-\delta} \big[h_{r+\delta} - h_r\big] dr \\
= & \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau-\delta} \int_X \Big[\Gamma_{r+\delta}(u_{r+\delta}) - P_{r+\delta,r}\Gamma_r(u_r)\Big] g_{r+\delta} dm_{r+\delta}\, dr \\
\leq & -\frac{2}{N} \limsup_{\delta \searrow 0} \int_\sigma^{\tau-\delta} \int_X \frac{1}{\delta} \int_r^{r+\delta} \big(P_{r+\delta,q}\Delta_q u_q\big)^2 dq\, g_{r+\delta} dm_{r+\delta}\, dr \\
\leq & -\frac{2}{N} \int_\sigma^\tau \liminf_{\delta \searrow 0} \Big(\int_X \frac{1}{\delta} \int_r^{r+\delta} P_{r+\delta,q}\Delta_q u_q\, dq\, g_{r+\delta} dm_{r+\delta}\Big)^2 \\
= & -\frac{2}{N} \int_\sigma^\tau \liminf_{\delta \searrow 0} \Big(\frac{1}{\delta} \int_r^{r+\delta} \int_X \Delta_q u_q\, g_q dm_q\, dq\Big)^2 dr \\
= & -\frac{2}{N} \int_\sigma^\tau \Big(\int_X \Delta_r u_r\, g_r dm_r\Big)^2 dr
\end{aligned}$$

according to Lebesgue's density theorem. On the other hand, similarly to the argumentation in the previous section, we have

$$\begin{aligned}
h_\tau - h_\sigma \geq & \limsup_{\delta \searrow 0} \frac{1}{\delta} \int_{\sigma-\delta}^{\tau} \big[h_{r+\delta} - h_r\big] dr \\
\geq & \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_{\sigma-\delta}^{\tau} \int_X \Gamma_{r+\delta}(u_{r+\delta}) d(\mu_{r+\delta} - \mu_r)\, dr \\
& + \limsup_{\delta \searrow 0} \frac{1}{\delta} \int_{\sigma-\delta}^{\tau} \int_X g_r \Big[\Gamma_{r+\delta}(u_{r+\delta}, u_r) - \Gamma_r(u_{r+\delta}, u_r)\Big] dm_r\, dr \\
& + \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_{\sigma-\delta}^{\tau} \int_X g_r \Big[\Gamma_{r+\delta}(u_{r+\delta}, u_{r+\delta} - u_r) + \Gamma_r(u_{r+\delta} - u_r, u_r)\Big] dm_r\, dr \\
=: & (I) + (II) + (III') + (III'').
\end{aligned}$$

Each of the four terms can be treated as before which then yields

$$
\begin{aligned}
& h_\tau - h_\sigma \geq (I) + (II) + (III') + (III'') \\
\geq & \int_\sigma^\tau \int_X \Big[-\Gamma_r(u_r) \cdot \Delta_r g_r + g_r \overset{\bullet}{\Gamma}_r (u_r) - 2\big(\Gamma_r(g_r, u_r) + g_r\, \Delta_r u_r\big)\, \Delta_r u_r \Big] dm_r\, dr \\
= & \int_\sigma^\tau \Big[-2\mathbf{\Gamma}_{2,r}(u_r)(g_r) + \int \overset{\bullet}{\Gamma}_r (u_r)\, g_r\, m_r \Big]\, dr.
\end{aligned}
$$

Combining this with the previous upper estimate and varying σ and τ, we thus have proven the dynamic Bochner inequality

$$
2\mathbf{\Gamma}_{2,r}(u_r)(g_r) \geq \int \overset{\bullet}{\Gamma}_r (u_r)\, g_r\, m_r + \frac{2}{N} \Big(\int_X \Delta_r u_r\, g_r dm_r \Big)^2
$$

for a.e. $r \in (s,t)$. □

2.8.4 From Transport Estimates to Bochner Inequality

Theorem 2.51 ("$(\mathbf{II_N}) \Rightarrow (\mathbf{IV_N})$"). *Suppose that the mm-space $(X, d_t, m_t)_{t \in I}$ satisfies the transport estimate* (19)=(64). *Then the dynamic Bochner inequality* (20)=(74) *with parameter N holds true as well as the regularity assumption* (17).

Proof of the regularity assumption. Thanks to Theorem 2.43, we already know that the transport estimate $(\mathbf{II}_N)$ implies the gradient estimate $(\mathbf{III}_N)$ in the case $N = \infty$. This proves the requested regularity. □

Proof of the dynamic Bochner inequality. We follow the argumentation from [18] with significant modifications due to time-dependence of functions, gradients, and operators and mainly because of lack of regularity.

Let $0 < s < t < T$ and $g_t \in \mathcal{F} \cap L^\infty$ with $g_t \geq 0$, $g_t \not\equiv 0$ as well as $u_s \in Lip(X)$ be given and fixed for the sequel. Without restriction $\int g_t dm_t = 1$. For $\tau \in (s,t)$, put $u_\tau = P_{\tau,s} u_s$ and $g_\tau = P^*_{t,\tau} g_t$. Note that – thanks to the parabolic Harnack inequality – g is uniformly bounded from above and bounded from below, away from 0, on $(s', t') \times X$ for each $s < s' < t' < t$. In the beginning, let us also assume that $||u_s||_\infty \leq 1/4$.

For each $\tau \in (s,t)$, define a Dirichlet form $\mathcal{E}^g_\tau$ on $L^2(X, g_\tau m_\tau)$ with domain $Dom(\mathcal{E}^g_\tau) := Dom(\mathcal{E})$ by

$$
\mathcal{E}^g_\tau(u) := \int \Gamma_\tau(u) g_\tau dm_\tau \qquad \text{for } u \in Dom(\mathcal{E}).
$$

Associated with the closed bilinear form $(\mathcal{E}^g_\tau, Dom(\mathcal{E}^g_\tau))$ on $L^2(X, g_\tau m_\tau)$, there is the self-adjoint operator Δ^g_τ and the semigroup $(H^{\tau,g}_a)_{a \geq 0}$, i.e. $u_a = H^{\tau,g}_a u$ solves

$$
\partial_a u_a = \Delta^g_\tau u_a \text{ on } (0,\infty) \times X, \qquad u_0 = u
$$

where $\Delta^g_\tau u = \Delta_\tau u + \Gamma_\tau(\log g_\tau, u)$. For fixed $\sigma \in (s, \tau)$, we define the path $(g^{\sigma,a}_\tau)_{a \geq 0}$ to be

$$
g^{\sigma,a}_\tau := g_\tau (1 + u_\sigma - H^{\tau,g}_a u_\sigma). \tag{79}
$$

Note that these are probability densities w.r.t. m_τ. Indeed, for all $a > 0$ and all $s < \sigma < \tau < t$

$$\int g_\tau^{\sigma,a} dm_\tau = 1 + \int u_\sigma (1 - H_a^{\tau,g} 1)\, g_\tau m_\tau = 1$$

thanks to conservativeness and symmetry of $H_a^{\tau,g}$ w.r.t. the measure $g_\tau m_\tau$. Moreover, $g_\tau^{\sigma,a} \geq 0$ for all a, σ and τ since the uniform bound $||u_s||_\infty \leq 1/4$ is preserved under the evolution of the time-dependent heat flow, thus $||u_\sigma||_\infty \leq ||P_{\sigma,s} u_s||_\infty \leq 1/4$, as well as under the heat flow in the static mm-space at fixed time τ, thus $||H_a^{\tau,g} u_\sigma||_\infty \leq ||u_\sigma||_\infty \leq 1/4$.

Now let us assume that the transport estimate $(\mathbf{II}_N)$ holds true and apply it to the probability measures $g_\tau m_\tau$ and $g_\tau^a m_\tau$. Then for all $s < \sigma < \tau < t$ and all $a > 0$

$$\begin{aligned} W_\sigma^2(\hat{P}_{\tau,s}(g_\tau m_\tau), \hat{P}_{\tau,\sigma}(g_\tau^{\sigma,a} m_\tau)) \leq\, & W_\tau^2(g_\tau m_\tau, g_\tau^{\sigma,a} m_\tau) \\ & - \frac{2}{N} \int_\sigma^\tau [S_r(\hat{P}_{\tau,r}(g_\tau m_\tau)) - S_r(\hat{P}_{\tau,r}(g_\tau^{\sigma,a} m_\tau))]^2 dr. \end{aligned}$$

Dividing by $2a^2$ and passing to the limit $a \searrow 0$, the subsequent Lemmata 2.52, 2.53 and 2.54 allow to estimate term by term. We thus obtain

$$\begin{aligned} & -\frac{1}{2} \int P_{\tau,\sigma}(\Gamma_\sigma(u_\sigma)) g_\tau dm_\tau + \int \Gamma_\tau(P_{\tau,\sigma} u_\sigma, u_\sigma) g_\tau dm_\tau \\ & \leq \frac{1}{2(1 - 2||u_\sigma||_\infty)} \int \Gamma_\tau(u_\sigma) g_\tau dm_\tau - \frac{1}{N} \int_\sigma^\tau \left[\int \Gamma_r\big(P_{\tau,r}(\log P_{\tau,r}^* g_\tau), u_\sigma\big)\, g_\tau dm_\tau \right]^2 dr. \end{aligned}$$

Replacing u_s by $\eta\, u_s$ for $\eta \in \mathbb{R}_+$ sufficiently small, we can get rid of the constraint $||u_s||_\infty \leq 1/4$. Then Lemma 2.52, Lemma 2.53 and Lemma 2.54 applied to ηu_s instead of u_s gives us

$$\begin{aligned} & -\frac{\eta^2}{2} \int P_{\tau,\sigma}(\Gamma_\sigma(u_\sigma)) g_\tau dm_\tau + \eta^2 \int \Gamma_\tau(P_{\tau,\sigma} u_\sigma, u_\sigma) g_\tau dm_\tau \\ & \leq \frac{\eta^2}{2(1 - 2\eta||u_\sigma||_\infty)} \int \Gamma_\tau(u_\sigma) g_\tau dm_\tau - \frac{\eta^2}{N} \int_\sigma^\tau \left[\int \Gamma_r\big(P_{\tau,r}(\log P_{\tau,r}^* g_\tau), u_\sigma\big)\, g_\tau dm_\tau \right]^2 dr. \end{aligned}$$

Dividing by η^2 and letting $\eta \to 0$ this inequality becomes

$$\begin{aligned} & -\frac{1}{2} \int P_{\tau,\sigma}(\Gamma_\sigma(u_\sigma)) g_\tau dm_\tau + \int \Gamma_\tau(P_{\tau,\sigma} u_\sigma, u_\sigma) g_\tau dm_\tau \\ & \leq \frac{1}{2} \int \Gamma_\tau(u_\sigma) g_\tau dm_\tau - \frac{1}{N} \int_\sigma^\tau \left[\int \Gamma_\tau\Big(P_{\tau,r}(\log P_{\tau,r}^* g_\tau), u_\sigma\Big)\, g_\tau dm_\tau \right]^2 dr. \end{aligned}$$

This can be reformulated into

$$\begin{aligned} & \frac{1}{2} \int \Gamma_\tau(u_\tau) g_\tau dm_\tau - \frac{1}{2} \int \Gamma_\sigma(u_\sigma) g_\sigma dm_\sigma \\ & - \frac{1}{2} \int \Gamma_\tau(u_\sigma) g_\tau dm_\tau - \frac{1}{2} \int \Gamma_\tau(u_\tau) g_\tau dm_\tau + \int \Gamma_\tau(u_\tau, u_\sigma) g_\tau dm_\tau \\ & \leq -\frac{1}{N} \int_\sigma^\tau \left[\int \Gamma_\tau\Big(P_{\tau,r}(\log P_{\tau,r}^* g_\tau), u_\sigma\Big)\, g_\tau dm_\tau \right]^2 dr. \end{aligned} \tag{80}$$

Now let us try to follow the argumentation from the proof of Theorem 2.50 and consider again the function

$$h_r := \int g_r \Gamma_r(u_r) dm_r$$

for $r \in (s,t)$. Recall that we already know from Theorem 2.43 that the transport estimate **(II$_N$)** implies the gradient estimate **(II)** ('without N'). Thus for all $s < \sigma < \tau < t$

$$\limsup_{\delta \searrow 0} \frac{1}{\delta} \int_{\sigma-\delta}^{\tau} \big(h_{r+\delta} - h_r\big) dr \le h_\tau - h_\sigma \le \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_{\sigma}^{\tau-\delta} \big(h_{r+\delta} - h_r\big) dr$$

Arguing as in the proof of Theorem 2.50 we get

$$h_\tau - h_\sigma \ge \int_\sigma^\tau \Big[-2\mathbf{\Gamma}_{2,r}(u_r)(g_r) + \int \overset{\bullet}{\Gamma}_r (u_r)\, g_r\, m_r \Big] dr.$$

On the other hand, applying the previous estimate (80) (with $r+\delta$, r and q in the place of τ, σ and r) we obtain

$$\begin{aligned} &h_\tau - h_\sigma \\ &\le \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau-\sigma} \Big[-\frac{2}{N} \int_r^{r+\delta} \Big[\int \Gamma_{r+\delta}\Big(P_{r+\delta,q}(\log P^*_{r+\delta,q} g_{r+\delta}), u_r\Big)\, g_{r+\delta} dm_{r+\delta} \Big]^2 dq \\ &\qquad + \int \Gamma_{r+\delta}(u_{r+\delta} - u_r) g_{r+\delta} dm_{r+\delta} \Big] dr. \end{aligned}$$

We estimate the term with the square from below using Young's inequality

$$\begin{aligned} &\Big[\int \Gamma_{r+\delta}\Big(P_{r+\delta,q}(\log P^*_{r+\delta,q} g_{r+\delta}), u_r\Big)\, g_{r+\delta} dm_{r+\delta} \Big]^2 \\ &\ge \frac{1}{1+\epsilon} \Big[\int \Gamma_r\Big(P_{r,q}(\log g_q), u_r\Big)\, g_r dm_r \Big]^2 \\ &- \frac{1}{\epsilon} \Big[\int \Gamma_{r+\delta}\Big(P_{r+\delta,q}(\log P^*_{r+\delta,q} g_{r+\delta}), u_r\Big)\, g_r dm_{r+\delta} - \int \Gamma_r\Big(P_{r,q}(\log g_q), u_r\Big)\, g_r dm_r \Big]^2, \end{aligned}$$

where $\epsilon > 0$ is arbitrary. Further estimating and using the log-Lipschitz continuity $r \mapsto \Gamma_r$ yields

$$
\begin{aligned}
&\Big[\int \Gamma_{r+\delta}\Big(P_{r+\delta,q}(\log P^*_{r+\delta,q} g_{r+\delta}), u_r\Big)\, g_{r+\delta} dm_{r+\delta} - \int \Gamma_r\Big(P_{r,q}(\log g_q), u_r\Big)\, g_r dm_r\Big]^2 \\
&\le 2\Big[\int \Gamma_{r+\delta}\Big(P_{r+\delta,q}(\log g_q), u_r\Big)\, g_{r+\delta} dm_{r+\delta} - \int \Gamma_r\Big(P_{r+\delta,q}(\log g_q), u_r\Big)\, g_{r+\delta} dm_{r+\delta}\Big]^2 \\
&+ 2\Big[\int \Gamma_r\Big(P_{r+\delta,q}(\log g_q), u_r\Big)\, g_{r+\delta} dm_{r+\delta} - \int \Gamma_r\Big(P_{r,q}(\log g_q), u_r\Big)\, g_{r+\delta} dm_{r+\delta}\Big]^2 \\
&+ 2\Big[\int \Gamma_r\Big(P_{r,q}(\log g_q), u_r\Big)\, g_{r+\delta} dm_{r+\delta} - \int \Gamma_r\Big(P_{r,q}(\log g_q), u_r\Big)\, g_r dm_r\Big]^2 \\
&\le 16L^2\delta^2 \cdot \\
&\Big[\int \Gamma_{r+\delta}\Big(P_{r+\delta,q}(\log g_q), u_r\Big)\, g_{r+\delta} dm_{r+\delta} + C\int \Gamma_{r+\delta}\Big(P_{r+\delta,q}(\log g_q) - u_r\Big)\, g_{r+\delta} dm_{r+\delta}\Big]^2 \\
&+ 2\Big[\int \Gamma_r\Big(P_{r+\delta,q}(\log g_q) - P_{r,q}(\log g_q), u_r\Big)\, g_{r+\delta} dm_{r+\delta}\Big]^2 \\
&+ 2\Big[\int \Gamma_r\Big(P_{r,q}(\log g_q), u_r\Big)\, d(g_{r+\delta} dm_{r+\delta} - g_r m_r)\Big]^2,
\end{aligned}
$$

which, after integration over $[r, r+\delta]$ and division by $\delta > 0$, converges to 0 as δ goes to 0. Indeed,

$$
\begin{aligned}
\delta \int_r^{r+\delta} \Big| \int \Gamma_{r+\delta}\Big(P_{r+\delta,q}(\log P^*_{r+\delta,q} g_{r+\delta}), u_r\Big)\, g_{r+\delta} dm_{r+\delta}\Big|^2 dq \\
\le C\delta\Big(\int_r^{r+\delta} \int \Gamma_q(\log g_q)\, g_q dm_q dr\Big)\mathcal{E}_r(u_r) \xrightarrow[\delta\to 0]{} 0,
\end{aligned}
$$

and Lemma 2.25 and Lebesgue differentiation theorem

$$
\frac{1}{\delta}\int_r^{r+\delta} \Big| \int \Gamma_r\Big(P_{r+\delta,q}(\log g_q) - P_{r,q}(\log g_q), u_r\Big)\, g_{r+\delta} dm_{r+\delta}\Big|^2 dq \xrightarrow[\delta\to 0]{} 0,
$$

while

$$
\frac{1}{\delta}\int_r^{r+\delta} \Big[\int \Gamma_r\Big(P_{r,q}(\log g_q), u_r\Big)\, d(g_{r+\delta} dm_{r+\delta} - g_r m_r)\Big]^2 dq \xrightarrow[\delta\to 0]{} 0.
$$

Thus, since ϵ is arbitrary, and from the Lebesgue differentiation theorem we get

$$
\begin{aligned}
\liminf_{\delta\to 0} \frac{1}{\delta}\int_r^{r+\delta} \Big[\int \Gamma_{r+\delta}\Big(P_{r+\delta,q}(\log P^*_{r+\delta,q} g_{r+\delta}), u_r\Big)\, g_{r+\delta} dm_{r+\delta}\Big]^2 dr \\
\ge \Big[\int \Gamma_r\Big(\log g_q, u_r\Big)\, g_r dm_r\Big]^2 = \Big[\int (\Delta_r u_r) g_r dm_r\Big]^2.
\end{aligned}
$$

Finally, with Corollary 2.23, the log-Lipschitz continuity of $r \mapsto \Gamma_r$, Lemma 2.25, and Lebesgue differentiation theorem applied to $r \mapsto \Delta_r u_r$, which is in

$L^2((s,t),\mathcal{H})$ thanks to Theorem 2.20,

$$
\begin{aligned}
&\limsup_{\delta\to 0}\frac{1}{\delta}\int_\sigma^{\tau-\delta}\int \Gamma_{r+\delta}(u_{r+\delta}-u_r)g_{r+\delta}dm_{r+\delta}\,dr\\
&\leq \limsup_{\delta\to 0}\frac{1}{\delta}\int_\sigma^{\tau-\delta}||g_{r+\delta}||_\infty\int \Gamma_{r+\delta}(u_{r+\delta}-u_r,u_{r+\delta})dm_{r+\delta}\,dr\\
&\leq \limsup_{\delta\to 0}\frac{1}{\delta}\int_\sigma^{\tau-\delta}e^{L|r+\delta-t|}||g_t||_\infty\cdot\\
&\qquad\Big(\int \Gamma_{r+\delta}(u_{r+\delta}-u_r,u_{r+\delta})dm_{r+\delta}-\int \Gamma_{r+\delta}(u_{r+\delta}-u_r,u_r)dm_{r+\delta}\Big)\,dr\\
&= \limsup_{\delta\to 0}\frac{1}{\delta}\int_\sigma^{\tau-\delta}e^{L|r+\delta-t|}||g_t||_\infty\cdot\\
&\qquad\Big(-\int\int_r^{r+\delta}\Delta_q u_q dq\Delta_{r+\delta}u_{r+\delta}dm_{r+\delta}-\int \Gamma_r(u_{r+\delta}-u_r,u_r)dm_r\Big)\,dr\\
&= \limsup_{\delta\to 0}\Big(\int_{\sigma+\delta}^{\tau}-e^{L|r-t|}||g_t||_\infty\int\frac{1}{\delta}\int_{r-\delta}^{r}\Delta_q u_q dq\Delta_r u_r dm_r dr\\
&\qquad+\int_\sigma^{\tau-\delta}e^{L|r+\delta-t|}||g_t||_\infty\int\frac{1}{\delta}\int_r^{r+\delta}\Delta_q u_q dq\Delta_r u_r dm_r\,dr\Big)\\
&\qquad=\int_\sigma^\tau e^{L|r-t|}||g_t||_\infty\Big(-\int(\Delta_r u_r)^2 dm_r+\int(\Delta_r u_r)^2dm_r\Big)=0.
\end{aligned}
$$

Combining the previous estimates we get

$$
h_\tau-h_\sigma\leq -\frac{2}{N}\int_\sigma^\tau\Big(\int \Delta_r u_r\, g_r dm_r\Big)^2 dr,
$$

and then

$$
-\frac{2}{N}\int_\sigma^\tau\Big(\int \Delta_r u_r\, g_r dm_r\Big)^2 dr\geq\int_\sigma^\tau\Big[-2\mathbf{\Gamma}_{2,r}(u_r)(g_r)+\int\overset{\bullet}{\Gamma}_r(u_r)\,g_r\,m_r\Big]\,dr,
$$

which proves the claim. □

Lemma 2.52. *For every $s<\sigma\leq\tau<t$,*

$$
\begin{aligned}
&\liminf_{a\to 0}\frac{W_\sigma^2(\hat P_{\tau,\sigma}(g_\tau^{\sigma,a}m_\tau),\hat P_{\tau,\sigma}(g_\tau m_\tau))}{2a^2}\\
&\geq-\int\frac{1}{2}P_{\tau,\sigma}(\Gamma_\sigma(u_\sigma))g_\tau dm_\tau+\int\Gamma_\tau(u_\tau,u_\sigma)g_\tau dm_\tau.
\end{aligned}
$$

Proof. We denote by Q_a^σ the Hopf-Lax semigroup with respect to the metric d_σ. Note that $aQ_a^\sigma(\phi)=Q_1^\sigma(a\phi)$, so the Kantorovich duality (40) can be written as

$$
\frac{W_\sigma^2(\nu_1,\nu_2)}{2a^2}=\frac{1}{a}\sup_\phi\Big[\int Q_a^\sigma\phi d\nu_1-\int\phi d\nu_2\Big].
$$

We deduce

$$\begin{aligned}
&\frac{W_\sigma^2(\hat{P}_{\tau,\sigma}(g_\tau^{\sigma,a}m_\tau), \hat{P}_{\tau,\sigma}(g_\tau m_\tau))}{2a^2} \geq \int \frac{Q_a^\sigma u_\sigma P^*_{\tau,\sigma}(g_\tau^{\sigma,a}) - u_\sigma P^*_{\tau,\sigma} g_\tau}{a} dm_s \\
&\geq \int \frac{Q_a u_\sigma - u_\sigma}{a} P^*_{\tau,\sigma}(g_\tau^{\sigma,a} - g_\tau) dm_\sigma + \int \frac{Q_a u_\sigma - u_\sigma}{a} P^*_{\tau,\sigma} g_\tau dm_\sigma \\
&+ \int u_\sigma \frac{P^*_{\tau,\sigma}(g_\tau^{\sigma,a} - g_\tau)}{a} dm_\sigma.
\end{aligned}$$

Note that, since u_s is a Lipschitz function, u_σ is a Lipschitz function as well. Indeed, from the dual representation of the Kantorovich-Rubinstein distance W_s^1 with respect to the metric d_s, we deduce

$$\begin{aligned}
&|u_\sigma(x) - u_\sigma(y)| = \left| \int u_s(z) d\hat{P}_{\sigma,s}(\delta_x)(z) - \int u_s(z) d\hat{P}_{t,s}(\delta_y)(z) \right| \\
&\leq \mathrm{Lip}_s(u_s) W_s^1(\hat{P}_{\sigma,s}(\delta_x), \hat{P}_{t,s}(\delta_y)) \leq \mathrm{Lip}_s(u_s) W_s(\hat{P}_{\sigma,s}(\delta_x), \hat{P}_{t,s}(\delta_y)) \\
&\leq \mathrm{Lip}_s(u_s) W_\sigma(\delta_x, \delta_y) = \mathrm{Lip}_s(u_s) d_\sigma(x,y),
\end{aligned}$$

where the last inequality is a consequence of Theorem 2.43

Since $0 \geq (Q_a^\sigma u_\sigma(x) - u_\sigma(x))/a \geq -2\mathrm{Lip}(u_\sigma)^2$ and $g_\tau^{\sigma,a} \to g_\tau$ in $L^2(X)$ the first integral vanishes. For the second integral we use (39) and estimate by Fatou's Lemma

$$\liminf_{a\to 0} \int \frac{Q_a^\sigma u_\sigma - u_\sigma}{a} P^*_{\tau,\sigma} g_\tau dm_\sigma \geq -\frac{1}{2} \int \mathrm{lip}_\sigma(u_\sigma)^2 P^*_{\tau,\sigma} g_\tau dm_\sigma.$$

For the last integral an argument similar to Lemma 2.34 for $H_a^{\tau,g}$ (compare Lemma 4.14 in [7]) yields

$$\lim_{a\to 0} \int \psi_\sigma \frac{P^*_{\tau,\sigma}(g_\tau^{\sigma,a} - g_\tau)}{a} dm_\sigma = \int \Gamma_\tau(P_{\tau,\sigma} u_\sigma, u_\sigma) g_\tau dm_\tau.$$

Combining the last two estimates we obtain

$$\begin{aligned}
&\liminf_{a\to 0} \frac{W_\sigma^2(\hat{P}_{\tau,\sigma}(g_\tau^{\sigma,a}m_\tau), \hat{P}_{\tau,\sigma}(g_\tau m_\tau))}{2a^2} \\
&\geq -\frac{1}{2} \int \mathrm{lip}_\sigma(u_\sigma)^2 P^*_{\tau,\sigma} g_\tau dm_\sigma + \int \Gamma_\tau(P_{\tau,\sigma} u_\sigma, u_\sigma) g_\tau dm_\tau \\
&= -\frac{1}{2} \int \Gamma_\sigma(u_\sigma) P^*_{\tau,\sigma} g_\tau dm_\sigma + \int \Gamma_\tau(P_{\tau,\sigma} u_\sigma, u_\sigma) g_\tau dm_\tau,
\end{aligned}$$

where the last inequality follows from our static RCD(K, N') assumption, which implies Poincaré inequality and doubling property for the static space (X, d_σ, m_σ), and the fact that u_σ is a Lipschitz function (cf. [20]). □

Lemma 2.53. *For every $s < \sigma \leq \tau < t$,*

$$\limsup_{a\to 0} \frac{W_\tau^2(g_\tau^{\sigma,a} m_\tau, g_\tau m_\tau)}{2a^2} \leq \frac{1}{2(1 - 2||\psi_\sigma||_\infty)} \int \Gamma_\tau(u_\sigma) g_\tau dm_\tau.$$

Proof. Let $(Q_a^\tau)_{a\geq 0}$ be the d_τ Hopf-Lax semigroup and fix a bounded Lipschitz function ϕ. Note that

$$\begin{aligned}
&\partial_a \int Q_a^\tau(\phi) g_\tau^{\sigma,a} dm_\tau \leq -\int \frac{1}{2}\mathrm{lip}_\tau(Q_a^\tau\phi)^2 g_\tau^{\sigma,a} dm_\tau + \int \Gamma_\tau(Q_a^\tau\phi, H_a^{\tau,g}u_\sigma) g_\tau dm_\tau \\
&= \int \left[-\frac{1}{2}\mathrm{lip}_\tau(Q_a^\tau\phi)^2(1+u_\sigma - H_a^{\tau,g}u_\sigma) + \Gamma_\tau(Q_a^\tau\phi, H_a^{\tau,g}u_\sigma)\right] g_\tau dm_\tau,
\end{aligned}$$

where the inequality follows from [4, Lemma 4.3.4] and dominated convergence. Applying the Cauchy-Schwartz inequality and that $\Gamma_\tau(\psi) \leq \mathrm{lip}_\tau(\psi)\ m_\tau$-a.e., we find

$$\begin{aligned}
&\int \Gamma_\tau(Q_a^\tau\phi, H_a^{\tau,g}u_\sigma) g_\tau dm_\tau \leq \sqrt{\mathcal{E}_g(Q_a^\tau\phi)\mathcal{E}_g(H_a^{\tau,g}u_\sigma)} \\
&\leq \sqrt{\int \mathrm{lip}_\tau(Q_a^\tau\phi)^2 g_\tau dm_\tau \mathcal{E}_g(H_a^{\tau,g}u_\sigma)}.
\end{aligned}$$

Then, since $1+u_\sigma - H_a^{\tau,g}u_\sigma \geq 1-2||u_\sigma||_\infty$, we obtain using Young's inequality

$$\begin{aligned}
\partial_a \int Q_a^\tau(\phi) g_\tau^{\sigma,a} dm_\tau &\leq \frac{1}{2(1-2||u_\sigma||_\infty)}\mathcal{E}_g(H_a^{\tau,g}u_\sigma) \leq \frac{1}{2(1-2||u_\sigma||_\infty)}\mathcal{E}_g(u_\sigma) \\
&= \frac{1}{2(1-2||u_\sigma||_\infty)}\int \Gamma_\tau(u_\sigma) g_\tau dm_\tau.
\end{aligned}$$

Integrating over $[0,a]$,

$$\int Q_a^\tau \phi g_\tau^{\sigma,\tau} dm_\tau - \int \phi g_\tau dm_\tau \leq \frac{a}{2(1-2||u_\sigma||_\infty)}\int \Gamma_\tau(u_\sigma) g_\tau dm_\tau,$$

and dividing by $a > 0$ proves the claim since the Kantorovich duality can be written as

$$\frac{W_\tau^2(\nu_1,\nu_2)}{2a^2} = \frac{1}{a}\sup_\phi \left[\int Q_a^\tau\phi d\nu_1 - \int \phi d\nu_2\right]$$

and ϕ was an arbitrary bounded Lipschitz function. □

Lemma 2.54.

$$\begin{aligned}
&\liminf_{a\to 0} \int_s^\tau \left[\frac{S_r(\hat{P}_{\tau,r}(g_\tau^{\sigma,a}m_\tau)) - S_r(\hat{P}_{\tau,r}(g_\tau m_\tau))}{a}\right]^2 dr \\
&\geq \int_s^\tau \left[\int \Gamma_\tau\big(P_{\tau,r}(\log g_r), u_\sigma\big) g_\tau dm_\tau\right]^2 dr.
\end{aligned}$$

Proof. With the same estimates as in [18] we have

$$\begin{aligned}
&[S_r(\hat{P}_{\tau,r}(g_\tau^{\sigma,a}m_\tau)) - S_r(\hat{P}_{\tau,r}(g_\tau m_\tau))]^2 \\
&\geq \frac{1}{(1+\delta)}\left[\int (P_{\tau,r}^*(g_\tau^{\sigma,a}) - g_r)\log g_r dm_r\right]^2 - \frac{1}{\delta}\left[\int \frac{(P_{\tau,r}^* g_\tau^{\sigma,a} - g_r)^2}{g_r} dm_r\right]^2.
\end{aligned}$$

Next we apply Jensen's inequality to the convex function $\alpha\colon \mathbb{R}\times\mathbb{R}_+ \to \mathbb{R}\cup\{+\infty\}$ defined by

$$\alpha(r,s) = \begin{cases} 0, & \text{if } r = 0 = s, \\ \frac{r^2}{s}, & \text{if } s \neq 0, \\ +\infty, & \text{if } s = 0 \text{ and } r \neq 0. \end{cases}$$

Recall that the map $dx \mapsto p_{\tau,r}(x,y)dm_\tau(x)$ is not Markovian, but Lemma 2.23 implies

$$0 \leq M_{\tau,r}(y) := \int_X p_{\tau,r}(x,y)dm_\tau(x) \leq e^{L(\tau-r)}.$$

Hence we can write

$$\begin{aligned}
&\int \alpha(P^*_{\tau,r}g^{\sigma,a}_\tau - P^*_{\tau,r}g_\tau, P^*_{\tau,r}g_\tau)dm_r \\
&\leq \int\int \frac{\alpha((g^{\sigma,a}_\tau(x) - g_\tau(x))M_{\tau,r}(y), g_\tau(x)M_{\tau,r}(y))}{M_{\tau,r}} p_{\tau,r}(x,y)dm_\tau(x)dm_r(y) \\
&= \int\int \alpha((g^{\sigma,a}_\tau(x) - g_\tau(x)), g_\tau(x))p_{\tau,r}(x,y)dm_\tau(x)dm_r(y) \\
&= \int \alpha((g^{\sigma,a}_\tau(x) - g_\tau(x)), g_\tau(x))dm_\tau(x) = \int g_\tau(\psi_\sigma - H^{\tau,g}_a u_\sigma)^2 dm_\tau,
\end{aligned}$$

where we applied Jensen's inequality in the second, Fubini in the third, and the definition of $g^{\sigma,a}_\tau$ in the last line. Dividing by a and taking the $\limsup$ we end up with

$$\begin{aligned}
&\limsup_{a\to 0} \frac{1}{a}\int \frac{(P^*_{\tau,r}g^{\sigma,a}_\tau - P^*_{\tau,r}g_\tau)^2}{P^*_{\tau,r}g_\tau} dm_r \leq \limsup_{a\to 0}\frac{1}{a}\int g_\tau(u_\sigma - H^{\tau,g}_a u_\sigma)^2 dm_\tau \\
&\leq \limsup_{a\to 0} 2||u_\sigma||_\infty \int g_\tau\left(\frac{H^{\tau,g}_a u_\sigma - u_\sigma}{a}\right)dm_\tau = -2||u_\sigma||_\infty \int g_\tau \Gamma_\tau(u_\sigma, 1)dm_\tau = 0.
\end{aligned}$$

The first equality follows from the fact that $\frac{1}{a}(H^{\tau,g}_a u_\sigma - u_\sigma) \to \Delta^g_\tau u_\sigma$ weakly in $\mathcal{F}^*$ (cf. Lemma 2.34 and [7, Lemma 4.14]).
Since $\delta > 0$ is arbitrary it suffices to show

$$\lim_{a\to 0}\frac{1}{a}\int P^*_{\tau,r}(g(H^{\tau,g}_a u_\sigma - u_\sigma))\log P^*_{\tau,r}g dm_r = \int \Gamma_\tau\big(P_{\tau,r}(\log P^*_{\tau,r}g), u_\sigma\big)g dm_\tau.$$

This, indeed, follows from the fact that $P_{\tau,r}(\log P^*_{\tau,r}g) \in \mathcal{F} = Dom(\mathcal{E}_\tau) = Dom(\mathcal{E}^g_\tau)$ (thanks to uniform boundedness of $P^*_{\tau,r}g$ from above and away from 0) and from the fact that $\frac{1}{a}(H^{\tau,g}_a u_\sigma - u_\sigma) \to \Delta^g_\tau u_\sigma$ weakly in $\mathcal{F}^*$ as $a \searrow 0$, more precisely (cf. Lemma 2.34)

$$\frac{1}{a}\int (H^{\tau,g}_a u_\sigma - u_\sigma)\phi g_\tau dm_\tau \to -\int \Gamma_\tau(u_\sigma, \phi)g_\tau dm_\tau$$

for all $\phi \in \mathcal{F}$ as $a \searrow 0$. $\square$

2.9 From Gradient Estimates to Dynamic EVI

In this section we will prove that the dual heat flow is a dynamic backward EVI-gradient flow presumed that the Bakry-Émery gradient estimate **(III)** holds for the ('primal') heat equation. We will present the argument only in the case $N = \infty$. That is, we now assume that for all $u \in Dom(\mathcal{E})$ and $0 < s < t < T$

$$\Gamma_t(P_{t,s}u) \leq P_{t,s}(\Gamma_s(u)) \quad m\text{-a.e. on } X. \tag{81}$$

For the notion of dynamic backward EVI$^{\pm}$-gradient flow we refer to section 2.10.

As in the previous chapters, the assumptions from section 2.6.1 will always be in force, in particular, we assume the RCD$^*(K, N')$-condition for each static mm-space (X, d_t, m_t) as well as boundedness and L-Lipschitz continuity (in t) for $\log d_t(x, y)$ and (in t and x) for $f_t(x)$.

2.9.1 Dynamic Kantorovich-Wasserstein Distances

For the subsequent discussions, let us fix $s, t \in I$ and – if not stated otherwise – $\vartheta : [0, 1] \to \mathbb{R}$ will always denote the linear interpolation

$$\vartheta(a) = (1 - a)s + ta, \tag{82}$$

In the following we introduce dynamic notions of the distance between two measures 'living in different time sheets'. The first notion seems to be natural and is defined via the length of curves, while the second one uses the approach of Hamilton Jacobi equations.

Definition 2.55. *For $s < t$ and a 2-absolutely continuous curve $(\mu^a)_{a\in[0,1]}$ we define the action*

$$\mathcal{A}_{s,t}(\mu) = \lim_{h\to 0} \sup \Big\{ \sum_{i=1}^{n} (a_i - a_{i-1})^{-1} W^2_{\vartheta(a_{i-1})}(\mu^{a_{i-1}}, \mu^{a_i}) \Big| \; 0 = a_0 < \cdots < a_n = 1, a_i - a_{i-1} \leq h \Big\}.$$

For two probability measures $\mu, \nu \in \mathcal{P}(X)$ we define

$$W^2_{s,t}(\mu, \nu) = \inf \Big\{ \mathcal{A}_{s,t}(\mu) \Big| \mu \in AC^2([0, 1], \mathcal{P}(X)) \text{ with } \mu_0 = \mu, \mu_1 = \nu \Big\}.$$

Lemma 2.56. *The following holds true.*

i) The action $\mu \mapsto \mathcal{A}_{s,t}(\mu)$ is lower semicontinuous, i.e. if $\mu^a_j \to \mu^a$ for every a as $j \to \infty$ we have

$$\mathcal{A}_{s,t}(\mu) \leq \liminf_{j\to\infty} \mathcal{A}_{s,t}(\mu_j).$$

ii) For every absolutely continuous curve μ

$$\mathcal{A}_{s,t}(\mu) = \lim_{h\to 0} \inf \Big\{ \sum_{i=1}^{n} (a_i - a_{i-1})^{-1} W^2_{\vartheta(a_{i-1})}(\mu^{a_{i-1}}, \mu^{a_i}) \Big| \; 0 = a_0 < \cdots < a_n = 1, a_i - a_{i-1} \leq h \Big\}.$$

Proof. Since $\mu_a^j \to \mu_a$ for every $a \in [0,1]$ in the Wasserstein sense we have for every partition $0 = a_0 < \cdots < a_n = 1$

$$\sum_{i=1}^{n}(a_i - a_{i-1})^{-1} W^2_{\vartheta(a_{i-1})}(\mu^{a_{i-1}}, \mu^{a_i}) = \lim_{j\to\infty} \sum_{i=1}^{n}(a_i - a_{i-1})^{-1} W^2_{\vartheta(a_{i-1})}(\mu_j^{a_{i-1}}, \mu_j^{a_i}),$$

and hence

$$\sum_{i=1}^{n}(a_i - a_{i-1})^{-1} W^2_{\vartheta(a_{i-1})}(\mu^{a_{i-1}}, \mu^{a_i}) \le \liminf_{j\to\infty} \mathcal{A}_{s,t}(\mu_j).$$

Taking the supremum over each partition and letting $h \to 0$ proves

$$\mathcal{A}_{s,t}(\mu) \le \liminf_{j\to\infty} \mathcal{A}_{s,t}(\mu_j).$$

We prove the second assertion by contradiction. Assume that there exists a sequence $h_j \to 0$, and a partition $0 = a_0^j < \cdots < a_{n^j}^j = 1$ such that

$$a_i^j - a_{i-1}^j \le h \quad \text{and} \quad \lim_{j\to\infty} \sum_{i=1}^{n}(a_i^j - a_{i-1}^j)^{-1} W^2_{\vartheta(a_{i-1}^j)}(\mu^{a_{i-1}^j}, \mu^{a_i^j}) < \mathcal{A}_{s,t}(\mu).$$

For every $j \in \mathbb{N}$ we define the curve $(\mu_j^a)_{a\in[0,1]}$ by

$$\mu_j^a = \mu^a_{a_{i-1}^j, a_i^j}, \text{ if } a \in [a_{i-1}^j, a_i^j],$$

where $(\mu^a_{a_{i-1}^j, a_i^j})_{a\in[a_{i-1}^j, a_i^j]}$ denotes the $W_{\vartheta(a_{i-1}^j)}$-geodesic connecting $\mu^{a_{i-1}^j}$ and μa_i^j. Note that for every partition $\{\bar{a}_i\}_{i=1}^N$ with $\bar{a}_i - \bar{a}_{i-1} \ll h_j$

$$\sum_{i=1}^{N}(\bar{a}_i - \bar{a}_{i-1})^{-1} W^2_{\vartheta(\bar{a}_{i-1})}(\mu_j^{\bar{a}_i}, \mu_j^{\bar{a}_{i-1}}) \le e^{2Lh_j} \sum_{i=1}^{n}(a_i^j - a_{i-1}^j)^{-1} W^2_{\vartheta(a_{i-1}^j)}(\mu^{a_i^j}, \mu^{a_{i-1}^j}),$$

since for every $a_{i-1}^j \le \bar{a}_{k-1} < \bar{a}_k \le a_i^j$

$$W^2_{\vartheta(a_{i-1}^j)}(\mu_j^{\bar{a}_k}, \mu_j^{\bar{a}_{k-1}}) \le \frac{(\bar{a}_k - \bar{a}_{k-1})^2}{(a_i^j - a_{i-1}^j)^2} W^2_{\vartheta(a_{i-1}^j)}(\mu^{a_{i-1}^j}, \mu^{a_i^j}).$$

Hence

$$\mathcal{A}_{s,t}(\mu_j) \le e^{2Lh_j} \sum_{i=1}^{n}(a_i^j - a_{i-1}^j)^{-1} W^2_{\vartheta(a_{i-1}^j)}(\mu^{a_i^j}, \mu^{a_{i-1}^j}).$$

This is a contradiction since $\mu_j^a \to \mu_a$ for every a and hence

$$\liminf_{j\to\infty} \mathcal{A}_{s,t}(\mu_j) \ge \mathcal{A}_{s,t}(\mu).$$

$\square$

Proposition 2.57. *For $s < t \in I$ and $\mu^0, \mu^1 \in \mathcal{P}$ we have*

$$W^2_{s,t}(\mu_0, \mu_1) = \inf\left\{\int_0^1 |\dot{\mu}^a|^2_{s+a(t-s)} da\right\} \tag{83}$$

where the infimum runs over all 2-absolutely continuous curves $(\mu^a)_{a\in[0,1]}$ in $\mathcal{P}$ connecting μ^0 and μ^1.

Proof. Choose an arbitrary partition $0 = a_0 < a_1 < \cdots < a_n = 1$ with $a_i - a_{i-1} \leq h$. Let $(\mu^a)_{a\in[0,1]} \in AC^2([0,1],\mathcal{P}(X))$. Then, from the absolute continuity of (μ^a), and the log Lipschitz property (36) we deduce

$$\begin{aligned}\sum_{i=1}^{n}(a_i - a_{i-1})^{-1}W^2_{\vartheta(a_{i-1})}(\mu^{a_{i-1}},\mu^{a_i}) &\leq \sum_{i=1}^{n}(a_i - a_{i-1})^{-1}\left(\int_{a_i}^{a_{i-1}}|\dot{\mu}^a|_{\vartheta(a_{i-1})}da\right)^2\\ &\leq \sum_{i=1}^{n}\int_{a_i}^{a_{i-1}}|\dot{\mu}^a|^2_{\vartheta(a_{i-1})}da\\ &\leq e^{2Lh}\int_0^1|\dot{\mu}^a|^2_{\vartheta(a)}da.\end{aligned}$$

Taking the supremum over all partitions and letting $h \to 0$ we obtain

$$\mathcal{A}_{s,t}(\mu) \leq \int_0^1 |\dot{\mu}^a|^2_{\vartheta(a)}da,$$

and consequently

$$W^2_{s,t}(\mu_0,\mu_1) \leq \inf\left\{\int_0^1|\dot{\mu}^a|^2_{s+a(t-s)}da\right\}.$$

To verify the other inequality, we fix again a curve $(\mu_a)_{a\in[0,1]} \in AC^2([0,1],\mathcal{P}(X))$ with finite energy $\mathcal{A}_{s,t}(\mu)$. For each $h > 0$ we consider the partition $0 = a_0 < a_1 < \cdots < a_n \leq 1 < a_{n+1}$ with $a_i = ih$ and $nh \leq 1$. We extend μ_a by μ_1 whenever $a > 1$. We define μ^h_a to be the $W_{\vartheta(a_{i-1})}$-geodesic connecting $\mu_{a_{i-1}}$ with μ_{a_i} whenever $a \in [a_{i-1}, a_i]$. Then we clearly have that $\mu^h \in AC^2([0,1],\mathcal{P}(X))$ and since μ is absolutely continuous, for each $a \in [0,1]$, $\mu^h_a \to \mu_a$ in $(\mathcal{P}(X), W)$. Note that $|\dot{\mu}^h_a|_{\vartheta(a)}$ is a uniformly bounded function in $L^2([0,1])$

$$\begin{aligned}\int_0^1|\dot{\mu}^h_a|^2_{\vartheta(a)}da &\leq e^{2Lh}\sum_{i=1}^{n+1}\int_{a_{i-1}}^{a_i}|\dot{\mu}^h_a|^2_{\vartheta(a_{i-1})}da\\ &\leq e^{2Lh}\sum_{i=1}^{n+1}(a_i-a_{i-1})^{-1}W^2_{\vartheta(a_{i-1})}(\mu_{a_{i-1}},\mu_{a_i}) < \infty,\end{aligned}$$

since μ^h_a is a piecewise geodesic and $\mathcal{A}_{s,t}(\mu) < \infty$. Then, by the Banach-Alaoglu Theorem there exists a subsequence (not relabeled) $h \to 0$, and a function $A \in L^2([0,1])$ such that $|\dot{\mu}^h|_{\vartheta(.)} \rightharpoonup A$ in $L^2([0,1])$. Hence from the convergence of $\mu^h_a \to \mu_a$ we get

$$\begin{aligned}W_{\vartheta(a)}(\mu_a,\mu_{a+\delta}) &= \lim_{h\to 0} W_{\vartheta(a)}(\mu^h_a,\mu^h_{a+\delta})\\ &\leq \liminf_{h\to 0}\int_a^{a+\delta}|\dot{\mu}_b|_{\vartheta(a)}db \leq \liminf_{h\to 0} e^{\delta(t-s)}\int_a^{a+\delta}|\dot{\mu}_b|_{\vartheta(b)}db\\ &= e^{\delta(t-s)}\int_a^{a+\delta}A(b)db,\end{aligned}$$

and hence

$$|\dot{\mu}_a|_{\vartheta(a)} \leq A(a) \text{ for a.e. } a \in [0,1].$$

Consequently,

$$\int_0^1 |\dot\mu_a|^2_{\vartheta(a)} da \le \int_0^1 A^2(a)da \le \liminf_{h\to 0} \int_0^1 |\dot\mu^h_a|^2_{\vartheta(a)} da$$
$$\le \liminf_{h\to 0} e^{2Lh} \sum_{i=1}^{n+1} \int_{a_{i-1}}^{a_i} |\dot\mu^h_a|^2_{\vartheta(a_{i-1})} da$$
$$\le \liminf_{h\to 0} e^{2Lh} \sum_{i=1}^{n+1} (a_i - a_{i-1})^{-1} W^2_{\vartheta(a_{i-1})}(\mu_{a_{i-1}}, \mu_{a_i}) \le \mathcal{A}_{s,t}(\mu),$$

which proves the claim.

□

To conclude this section we define a dynamic 'dual distance' inspired by the dual formulation of the Kantorovich distance. We introduce the function space HLS_ϑ defined by

$$HLS_\vartheta := \Big\{ \varphi \in \mathrm{Lip}_b([a_0,a_1]\times X) \Big| \; \partial_a \varphi_a \le -\frac{1}{2}\Gamma_{\vartheta(a)}(\varphi_a) \quad L^1 \times m \text{ a.e. in } (a_0,a_1)\times X \Big\}.$$

In particular for all nonnegative $\phi \in L^1(X)$ and $\varphi \in HLS_\vartheta$

$$\int \phi \varphi_{a_1} dm - \int \phi \varphi_{a_0} dm \le -\frac{1}{2} \int_{a_0}^{a_1} \int \phi \Gamma_{\vartheta(a)}(\varphi_a) dm da.$$

Definition 2.58. *Let $s < t$ and let $\vartheta \colon [a_0,a_1] \to [s,t]$ denote the linear interpolation. Define for two probability measures μ_0, μ_1*

$$\tilde W^2_\vartheta(\mu_0,\mu_1) := 2 \sup_\varphi \left\{ \int \varphi_{a_1} d\mu_1 - \int \varphi_{a_0} d\mu_0 \right\},$$

where the supremum runs over all maps $\varphi(a,x) = \varphi_a(x) \in HLS_\vartheta$.

Note that $\tilde W_\vartheta$ does not necessarily define a distance. It does not even have to be symmetric. The next Lemma collects two essential properties of $\tilde W_\vartheta$.

Lemma 2.59. *The following holds true.*

1. *$\tilde W_\vartheta$ is lower semicontinuous with respect to the weak-* topology on $\mathcal{P}(X) \times \mathcal{P}(X)$.*
2. *For every μ_0, μ_1*
$$W_s^2(\mu_0,\mu_1) \le e^{2L|s-t|}(a_1 - a_0)\tilde W^2_\vartheta(\mu_0,\mu_1). \tag{84}$$

Proof. To show the first assertion, let $\mu_0, \mu_1 \in \mathcal{P}(X)$ and choose $\varphi \in HLS_\vartheta$ almost optimal, i.e.

$$\frac{1}{2}\tilde W_\vartheta(\mu_0,\mu_1) \le \int \varphi_{a_1} d\mu_1 - \int \varphi_{a_0} d\mu_0 - \varepsilon,$$

where $\varepsilon > 0$. Let $\mu_0^n \to \mu_0$, $\mu_1^n \to \mu$ be two sequences converging in duality with continuous bounded functions on X. then, since φ_{a_1} and φ_{a_0} belong to $\mathcal{C}_b(X)$,

$$\begin{aligned}\frac{1}{2}\tilde{W}_\vartheta(\mu_0,\mu_1) &\leq \int \varphi_{a_1} d\mu_{a_1} - \int \varphi_{a_0} - \varepsilon \\ &= \lim_{n\to\infty} \left\{ \int \varphi_{a_1} d\mu_1^n - \int \varphi_{a_0} d\mu_0^n \right\} - \varepsilon \\ &\leq \frac{1}{2} \liminf_{n\to\infty} \tilde{W}_\vartheta(\mu_0^n,\mu_1^n) - \varepsilon.\end{aligned}$$

This proves, since $\varepsilon > 0$ was arbitrary, that $\tilde{W}_\vartheta$ is lower semicontinuous with respect to the weak-*topology on $\mathcal{P}(X) \times \mathcal{P}(X)$. The second statement follows from the Kantorovich duality. Indeed, let $\varphi \in \mathrm{Lip}_b(X)$. As already mentioned above the Hopf-Lax semigroup $\varphi_b := Q_b^s(\varphi)$ solves

$$\frac{d}{db}\varphi_b \leq -\frac{1}{2}\Gamma_s(\varphi_b) \leq -\frac{1}{2}e^{-2L|s-t|}\Gamma_{(1-b)s+bt}(\varphi_b) \quad L^1 \times m \text{ a.e. in}(0,1)\times X. \tag{85}$$

Set $\tilde{\varphi}_a := e^{-2L|s-t|}(a_1-a_0)^{-1}\varphi_{\gamma(a)}$, where $\gamma\colon [a_0,a_1] \to [0,1]$ with $\gamma(a) = \frac{a-a_0}{a_1-a_0}$. Then $\tilde{\varphi}$ solves

$$\frac{d}{da}\tilde{\varphi}_a \leq -\frac{1}{2}\Gamma_{\vartheta(a)}(\tilde{\varphi}_a) \text{ in } (a_0,a_1)\times X,$$

and

$$e^{-2L|s-t|}(a_1-a_0)^{-1}\left(\int \varphi_1 d\mu_1 - \int \varphi_0 d\mu_0\right) = \int \tilde{\varphi}_{a_1} d\mu_1 - \int \tilde{\varphi}_{a_0} d\mu_0.$$

Hence

$$e^{-2L|s-t|}(a_1-a_0)^{-1}\left(\int \varphi_1 d\mu_1 - \int \varphi_0 d\mu_0\right) \leq \frac{1}{2}\tilde{W}_\vartheta^2(\mu_0,\mu_1).$$

Taking the supremum among all φ the Kantorovich duality for the metric W_s implies

$$W_s^2(\mu_0,\mu_1) \leq e^{2L|s-t|}(a_1-a_0)\tilde{W}_\vartheta^2(\mu_0,\mu_1).$$

□

Proposition 2.60. *Let $\vartheta\colon [0,1] \to [s,t]$ be the linear interpolation. Then we have $\tilde{W}_\vartheta \leq W_{s,t}$.*

Proof. Fix $\varphi \in HJS_\vartheta$ and $(\mu)_{a\in[0,1]}$ 2-absolutely continuous curve. We subdivide $[0,1]$ into l intervals $[(k-1)/l, k/l]$ of length $\frac{1}{l}$. On each interval $[(k-1)/l, k/l]$ we approximate $(\mu_a)_{|[(k-1)/l,k/l]}$ by regular curves $(\rho_a^{n,k})_{a\in[(k-1)/l,k/l]}$. Obviously, for each k,n the map $[(k-1)/l, k/l] \ni a \mapsto \int \varphi_a d\rho_a^{k,n}$ is absolutely continuous;

$$\int \varphi_{a+h} d\rho_{a+h} - \int \varphi_a d\rho_a \leq \mathrm{Lip}(\varphi_{a+h})W(\rho_{a+h},\rho_a) + ||\varphi_{a+h}-\varphi_a||_\infty.$$

Let $u_a^{k,n}$ be the density of the regular curve $\rho_a^{k,n}$. Hence for fixed k,n

$$\frac{d}{da}\int \varphi_a u_a^{k,n} dm \leq \int \varphi_a \dot{u}_a^{k,n} dm - \frac{1}{2}\int u_a^{k,n}\Gamma_{\vartheta(a)}(\varphi_a) dm$$

From Lemma 87 we deduce

$$\int \dot{u}_a^{k,n}\varphi_a dm \le \frac{1}{2}|\dot{\rho}_a^{k,n}|^2_{\vartheta(k-1/l)} + \frac{1}{2}\int (\mathrm{lip}_{\vartheta(k-1/l)}\varphi_a)^2 d\rho_a^{k,n}.$$

Adding these two inequalites, integrating over $[(k-1)/l, k/l]$ and noting that

$$e^{-L\frac{|t-s|}{l}}(\mathrm{lip}_{\vartheta(k-1/l)}(\varphi_a))^2 \le \Gamma_{\vartheta(a)}(\varphi_a) \qquad m \text{ a.e.},$$

we obtain

$$\begin{aligned}
&\int \varphi_{k/l}u^{k,n}_{k/l}dm - \int \varphi_{k-1/l}u^{k,n}_{k-1/l}dm \\
&\le \frac{1}{2}\int_{k-1/l}^{k/l}|\dot{\rho}_a^{k,n}|^2_{\vartheta(k-1/l)}da + \frac{1}{2}(1-e^{-L\frac{|t-s|}{l}})\int_{k-1/l}^{k/l}\int (\mathrm{lip}_{\vartheta(k-1/l)}\varphi_a)^2 d\rho_a^{k,n}da \\
&\le \frac{1}{2}\int_{k-1/l}^{k/l}|\dot{\rho}_a^{k,n}|^2_{\vartheta(k-1/l)}da + \frac{C_1}{2l}(1-e^{-L\frac{|t-s|}{l}})
\end{aligned}$$

Taking the limit $n \to \infty$ (and taking the scaling into account) gives

$$\int \varphi_{k/l}d\mu_{k/l} - \int \varphi_{k-1/l}d\mu_{k-1/l} \le \frac{1}{2}lW^2_{\vartheta(k-1/l)}(\mu_{k-1/l},\mu_{k/l}) + \frac{C_1}{2l}(1-e^{-L\frac{|t-s|}{l}}).$$

Summing over each partition and noting that the left hand side is a telescoping sum yields

$$\int \varphi_1 d\mu_1 - \int \varphi_0 d\mu_0 \le \frac{1}{2}\sum_{k=1}^{l} lW^2_{\vartheta(k-1/l)}(\mu_{k-1/l},\mu_{k/l}) + \frac{C_1}{2}(1-e^{-L\frac{|t-s|}{l}}).$$

Letting $l \to \infty$ we obtain the desired estimate. $\square$

Corollary 2.61. *Let $s < t$ and $[0,1] \ni a \mapsto \vartheta(a) = (1-a)s + at$. Then for every $\mu_0, \mu_1 \in \mathcal{P}(X)$ we have*

$$W_{s,t}(\mu_0,\mu_1) = \tilde{W}_\vartheta(\mu_o,\mu_1).$$

Proof. We already know from Proposition 2.60 that $W_{s,t}(\mu_0,\mu_1) \ge W_{\vartheta^*}(\mu_o,\mu_1)$. Hence it remains to prove the other inequality.

For this let $(\varphi_a) \in HLS_\vartheta$, and (μ_a) an absolutely continuous curve connecting μ_0 and μ_1.

Consider the Partition $0 = a_0 < a_1 < \dots a_n = 1$ with $a_i - a_{i-1} \le h$ for some $h > 0$. Set

$$[a_{i-1},a_i] \ni a \mapsto \vartheta_i(a) = \frac{a_i - a}{a_i - a_{i-1}}\vartheta(a_{i-1}) + \frac{a - a_{i-1}}{a_i - a_{i-1}}\vartheta(a_i)$$

and $\tilde{\varphi}^i_a = \varphi_a|_{[a_{i-1},a_i]}$. Notice that $(\varphi^i_a)_a$ is in HLS_{ϑ_i}. Hence

$$\tilde{W}^2_{\vartheta_i}(\mu_{a_{i-1}},\mu_{a_i}) \le 2\left\{\int \varphi_{a_i}d\mu_{a_i} - \int \varphi_{a_{i-1}}d\mu_{a_{i-1}}\right\}.$$

Then summing over the partitions and taking the scalings into account we end up with

$$\sum_{i=1}^{n}(a_i - a_{i-1})^{-1} W^2_{\vartheta(a_{i-1})}(\mu_{a_{i-1}}, \mu_{a_i}) \le e^{2Lh|s-t|} \sum_{i=1}^{n} \tilde{W}^2_{\vartheta_i}(\mu_{a_{i-1}}, \mu_{a_i})$$
$$\le 2e^{2Lh|s-t|} \sum_{i=1}^{n} \left\{ \int \varphi_{a_i} d\mu_{a_i} - \int \varphi_{a_{i-1}} d\mu_{a_{i-1}} \right\}$$
$$= 2e^{2Lh|s-t|} \left\{ \int \varphi_1 d\mu_1 - \int \varphi_0 d\mu_0 \right\},$$

where we made use of Lemma 2.59(ii) in the first inequality. Taking the supremum over all $(\varphi_a) \in HLS_\vartheta$ we deduce

$$\sum_{i=1}^{n}(a_i - a_{i-1})^{-1} W^2_{\vartheta(a_{i-1})}(\mu_{a_{i-1}}, \mu_{a_i}) \le e^{2Lh|s-t|} \tilde{W}^2_{\vartheta}(\mu_0, \mu_1), \tag{86}$$

We conclude

$$W^2_{s,t}(\mu_0, \mu_1) \le \tilde{W}^2_{\vartheta}(\mu_0, \mu_1),$$

from taking the supremum in (86) over the partition $0 = a_0 < a_1 < \cdots < a_n = 1$ with $a_i - a_{i-1} < h$ and subsequently letting $h \searrow 0$.

□

2.9.2 Action Estimates

Let us recall the following estimate about the oscillation of $a \mapsto \int \varphi d\rho^a$ from [7, Lemma 4.12]. For fixed $t > 0$, let $(\rho^a)_a$ be a 2-absolutely continuous curve in $\mathcal{P}$ with $\rho^a = u^a m_t$ and $u \in \mathcal{C}^1((0,1), L^1(X, m_t))$. Then for any Lipschitz function φ we have

$$\left| \int \dot{u}^a \varphi dm_t \right| \le \frac{1}{2} |\dot{\rho}^a|^2_t + \frac{1}{2} \int \Gamma_t(\varphi) d\rho^a. \tag{87}$$

Actually, we have inequality (87) for each $\varphi \in Dom(\mathcal{E})$ since we assume that each (X, d_t, m_t) is a static RCD(K, ∞) which implies that Lipschitz functions are dense in the domain of the quadratic form $\mathcal{E}$ with respect to the norm $\sqrt{||\varphi||^2 + \mathcal{E}(\varphi)}$ (Proposition 4.10 in [6]).

Moreover we will use the following result about difference quotients and concatenations of functions in $\mathcal{F}_{(s,t)}$.

Lemma 2.62. *Let* $0 < s < T$.

1. *Let* $u \in \mathcal{F}_{(s,t)}$. *Then for almost every* $a \in (s,t)$

$$\frac{1}{h}(u_{a+h} - u_a) \to \partial_a u_a \text{ weakly}^* \text{ in } \mathcal{F}^*,$$

i.e. for every $v \in \mathcal{F}$ *and for almost every* $a \in (s,t)$

$$\int \frac{1}{h}(u_{a+h} - u_a) v dm_\diamond \to \langle \partial_a u_a, v \rangle.$$

2. *For $u \in \mathcal{F}_{(s,t)}$ and $\vartheta \in \mathcal{C}^1([0,1])$ the linear interpolation from s to t, we have that $(u \circ \vartheta) \in \mathcal{F}_{(0,1)}$ with distributional derivative*

$$\partial_a (u \circ \vartheta)(a) = (t-s)\partial_a u_{\vartheta(a)}.$$

Proof. From Corollary 5.6. in [41] it follows for $u \in \mathcal{F}_{(s,t)}$ and $v \in \mathcal{F}$

$$\int u_{a+h} v dm_\diamond - \int u_a v dm_\diamond = \int_a^{a+h} \langle \partial_b u_b, v\rangle db.$$

Since $b \mapsto \langle \partial_b u_b, v\rangle$ is in $L^1(s,t)$ we apply the Lebesgue differentiation theorem and obtain that for almost every $a \in (s,t)$

$$\lim_{h\to 0} \frac{1}{h} \int u_{a+h} v dm_\diamond - \int u_a v dm_\diamond = \lim_{h\to 0} \frac{1}{h} \int_a^{a+h} \langle \partial_b u_b, v\rangle db = \langle \partial_a u_a, v\rangle.$$

This proves the first assertion. To show the second recall that we can approximate each $u \in \mathcal{F}_{(s,t)}$ by smooth functions $(u^n) \subset \mathcal{C}^\infty([s,t] \to \mathcal{F})$ by virtue of [41, Lemma 5.3]. So for each $n \in \mathbb{N}$ and for each smooth compactly supported test function $\psi\colon (0,1) \to \mathcal{F}$ we have that

$$\int_0^1 \int (u^n \circ \vartheta)(a) \partial_a \psi_a dm_\diamond da = -\int_0^1 \int \dot\vartheta(a) \partial_a u^n_{\vartheta(a)} \psi_a dm_\diamond da.$$

Note that the term on the left-hand side converges to $\int_0^1 \int (u \circ \vartheta)(a)\partial_a \psi_a dm_\diamond da$ as $n \to \infty$ since

$$\left| \int_0^1 \int (u^n \circ \vartheta - u \circ \vartheta) \partial_a \psi_a dm_\diamond da \right| \leq (t-s)^{-1} \int_s^t ||u^n_a - u_a||_{\mathcal{F}} ||\partial_a \psi_{\vartheta^{-1}(a)}||_{\mathcal{F}} da,$$

where we applied integration by substitution. Similarly for the right-hand side

$$\left| \int_0^1 \dot\vartheta(a) \langle \partial_a u^n_{\vartheta(a)} - \partial_a u_{\vartheta(a)}, \psi_a \rangle dm_\diamond da \right| \leq \int_s^t ||\partial_a u^n_a - \partial_a u_a||_{\mathcal{F}^*} ||\psi_{\vartheta^{-1}(a)}||_{\mathcal{F}} da,$$

and consequently as $n \to \infty$

$$\int_0^1 \int (u \circ \vartheta)(a) \partial_a \psi_a dm_\diamond da = -\int_0^1 (t-s) \langle \partial_a u_{\vartheta(a)}, \psi_a\rangle da,$$

which is the assertion. ☐

For the following lemmas let $(\rho_a)_{a\in[0,1]}$ be a regular curve and let $\vartheta\colon [0,1] \to [0,\infty)$

$$\vartheta(a) := (1-a)s + at, \text{ where } s < t.$$

Set $\rho_{a,\vartheta} := \hat P_{t,\vartheta(a)}(\rho_a) = u_{a,\vartheta} m_{\vartheta(a)}$.

Lemma 2.63. *The curve $(u_{a,\vartheta})_{a\in[0,1]}$ belongs to* $\mathrm{Lip}([0,1], \mathcal{F}^*)$ *with $u_{a,\vartheta} \in L^2([0,1] \to \mathcal{F})$ and distributional derivative $\partial_a u_{a,\vartheta} \in L^\infty([0,1] \to \mathcal{F}^*)$ satisfying*

$$\partial_a u_{a,\vartheta} = -(t-s)\Delta_{\vartheta(a)} u_{a,\vartheta} + \partial_a f_{\vartheta(a)} u_{a,\vartheta} - P^*_{t,\vartheta(a)}(\dot u_a).$$

Proof. First we show that $(u_{a,\vartheta})$ is in $L^2([0,1]\to\mathcal{F})$. For this recall that, since (ρ_a) is regular, $u_a \le R$ and $\mathcal{E}_t(\sqrt{u_a}) \le E$ for all $a\in[0,1]$ and hence by Lemma 2.23 we get

$$\int_0^1 ||u_{a,\vartheta}||^2_{L^2(m_{\vartheta(a)})}da \le e^{L(t-s)}\int_0^1 ||u_a||^2_{L^2(m_t)}da$$
$$\le Re^{L(t-s)}\int_0^1 ||u_a||_{L^1(m_t)}da = Re^{L(t-s)},$$

and by Theorem 2.20

$$\int_0^1 \mathcal{E}_{\vartheta(a)}(u_{a,\vartheta})da \le e^{3L(t-s)}\int [\mathcal{E}_t(u_a)+||u_a||^2_{L^2(m_t)}]da$$
$$\le e^{3L(t-s)}\sqrt{R}[\int_0^1 2\mathcal{E}_t(\sqrt{u_a})da + R] \le e^{3L(t-s)}\sqrt{R}(2E+R).$$

This shows that $(u_{a,\vartheta})$ is in $L^2([0,1]\to\mathcal{F})$.

Next we show that $(u_{a,\vartheta})$ is contained in $\mathrm{Lip}([0,1],\mathcal{F}^*)$. For this let $\psi\in\mathcal{F}$. Then, for almost every $a_0,a_1\in(0,1)$, we obtain with Lemma 2.62, since $P^*_{t,\vartheta(a)}u_{a_0}\in\mathcal{F}_{(0,1)}$,

$$\begin{aligned}
&\int \psi u_{a_1,\vartheta}dm_\diamond - \int \psi u_{a_0,\vartheta}dm_\diamond\\
=&\int \psi(P^*_{t,\vartheta(a_1)}u_{a_0} - P^*_{t,\vartheta(a_0)}u_{a_0})dm_\diamond + \int \psi P^*_{t,\vartheta(a_1)}(u_{a_1}-u_{a_0})dm_\diamond\\
=&(t-s)\int_{a_0}^{a_1}\mathcal{E}^\diamond_{\vartheta(a)}(P^*_{t,\vartheta(a)}u_{a_0},\psi)da + (t-s)\int_{a_0}^{a_1}\int \dot f_{\vartheta(a)}P^*_{t,\vartheta(a)}u_{a_0}\psi dm_\diamond da\\
&+\int P_{t,\vartheta(a_1)}(\psi e^{f_{\vartheta(a_1)}})(u_{a_1}-u_{a_0})dm_t\\
\le&(t-s)\int_{a_0}^{a_1}\mathcal{E}_{\vartheta(a)}(P^*_{t,\vartheta(a)}u_{a_0})^{1/2}\mathcal{E}_{\vartheta(a)}(\psi e^{f_{\vartheta(a)}})^{1/2}da\\
&+(t-s)\int_{a_0}^{a_1}||\dot f_{\vartheta(a)}||_\infty ||P^*_{t,\vartheta(a)}u_{a_0}||_{L^2(m_{\vartheta(a)})}||\psi e^{f_{\vartheta(a)}}||_{L^2(m_\diamond)}da\\
&+||e^{-f_t}||_\infty \mathcal{E}_\diamond(P_{t,\vartheta(a_1)}(\psi e^{f_{\vartheta(a_1)}}))^{1/2}\sup_a ||\dot u_a||_{\mathcal{F}^*}(a_1-a_0)\\
\le&(t-s)\mathcal{E}_{\vartheta(a)}(\psi)^{1/2}\int_{a_0}^{a_1}\mathrm{Lip}(f_{\vartheta(a)})\mathcal{E}_{\vartheta(a)}(P^*_{t,\vartheta(a)}u_{a_0})^{1/2}da\\
&+(t-s)\int_{a_0}^{a_1}||\dot f_{\vartheta(a)}||_\infty ||P^*_{t,\vartheta(a)}u_{a_0}||_{L^2(m_{\vartheta(a)})}||\psi e^{f_{\vartheta(a)}}||_{L^2(m_\diamond)}da\\
&+||e^{-f_t}||_\infty \mathcal{E}_\diamond(P_{t,\vartheta(a_1)}(\psi e^{f_{\vartheta(a_1)}}))^{1/2}\sup_a ||\dot u_a||_{\mathcal{F}^*}(a_1-a_0).
\end{aligned}$$

Due to our assumptions on f we have that

$$\mathrm{Lip}(f_{\vartheta(a)})\le C,\ ||\dot f_{\vartheta(a)}||_\infty \le L,\ ||f_t||_\infty \le C,$$

while the energy estimate Theorem 2.20 and Corollary 2.23 yields

$$\mathcal{E}_{\vartheta(a)}(P^*_{t,\vartheta(a)}u_{a_0}) \le e^{3L(t-s)}[\mathcal{E}_t(u_{a_0})+||u_{a_0}||^2_{L^2(m_t)}],$$
$$||P^*_{t,\vartheta(a)}u_{a_0}||_{L^2(m_{\vartheta(a)})} \le e^{L(t-s)/2}||u_{a_0}||_{L^2(m_t)}.$$

Note that the last two expressions are bounded since u is a regular curve. Moreover from (23), the gradient estimate (81) and Corollary 2.23 we find

$$\mathcal{E}_\diamond(P_{t,\vartheta(a_1)}(\psi e^{f_{\vartheta(a_1)}})) \leq Ce^{L(t-s)}\mathrm{Lip}(e^{f_{\vartheta(a_1)}})^2\mathcal{E}_{\vartheta(a_1)}(\psi)$$

Applying (23) once more we find that there exists a constant λ such that

$$\int \psi u_{a_1,\vartheta} dm_\diamond - \int \psi u_{a_0,\vartheta} dm_\diamond \leq (a_1 - a_0)\lambda||\psi||_{\mathcal{F}}, \tag{88}$$

and thus

$$||u_{a_1} - u_{a_0}||_{\mathcal{F}^*} \leq \lambda.$$

Note also that (88) holds for every a_0, a_1 by approximating with Lebesgue points. This implies the existence of $\partial_a u_{a,\vartheta} \in L^\infty([0,1], \mathcal{F}^*)$ such that

$$\int \psi u_{a_1,\vartheta} dm_\diamond - \int \psi u_{a_0,\vartheta} dm_\diamond = \int_{a_0}^{a_1} \langle \partial_a u_{a,\vartheta}, \psi\rangle_{\mathcal{F}^*,\mathcal{F}} da.$$

Fix $\psi \in \mathrm{Lip}_b(X)$. By a similar calculation as above it ultimately follows that

$$\begin{aligned}
&\lim_{h\to 0} \frac{1}{h}\Big(\int \psi u_{a+h,\vartheta} dm_\diamond - \int \psi u_{a,\vartheta} dm_\diamond\Big) \\
&= (t-s)\mathcal{E}^\diamond_{\vartheta(a)}(P^*_{t,\vartheta(a)}u_a, \psi) + (t-s)\int \dot{f}_{\vartheta(a)} P^*_{t,\vartheta(a)} u_a \psi dm_\diamond \\
&+ \lim_{h\to 0}\int P_{t,\vartheta(a+h)}(\psi e^{f_{\vartheta(a+h)}})\frac{(u_{a+h} - u_a)}{h} dm_t
\end{aligned}$$

almost everywhere. To determine the last integral recall that $u \in \mathcal{C}^1([0,1], L^1(X))$. Then since $\psi \in \mathrm{Lip}_b(X)$

$$\begin{aligned}
\lim_{h\to 0}\int P_{t,\vartheta(a+h)}(\psi e^{f_{\vartheta(a+h)}})\frac{(u_{a+h} - u_a)}{h} dm_t &= \int P_{t,\vartheta(a)}(\psi e^{f_{\vartheta(a)}})\dot{u}_a dm_t \\
&= \int (\psi e^{f_{\vartheta(a)}})P^*_{t,\vartheta(a)}\dot{u}_a dm_{\vartheta(a)} = \langle P^*_{t,\vartheta(a)}\dot{u}_a, \psi\rangle_{\mathcal{F}^*,\mathcal{F}}.
\end{aligned}$$

From the Lipschitz continuity of $(u_{a,\vartheta})$ we deduce that for almost every $a \in [0,1]$

$$\langle \partial_a u_{a,\vartheta}, \psi\rangle_{\mathcal{F}^*,\mathcal{F}} = \langle -(t-s)\Delta_{\vartheta(a)} u_{a,\vartheta} + \partial_a f_{\vartheta(a)} u_{a,\vartheta} - P^*_{t,\vartheta(a)}(\dot{u}_a), \psi\rangle_{\mathcal{F}^*,\mathcal{F}}.$$

We conclude the proof by approximating $\psi \in \mathcal{F}$ with bounded Lipschitz functions. $\square$

Lemma 2.64. *For any map $\varphi \in HLS_\vartheta$ the map $a \mapsto \int \varphi_a d\rho_{a,\vartheta}$ is absolutely continuous and*

$$\begin{aligned}
\int \varphi_1 d\rho_{1,\vartheta} - \int \varphi_0 d\rho_{0,\vartheta} \leq \int_0^1 \Big[&-\frac{1}{2}\int \Gamma_{\vartheta(a)}(\varphi_a) d\rho_{a,\vartheta} + \int P_{t,\vartheta(a)}(\varphi_a)\, \partial_a u_a \, dm_t \\
&+ (t-s)\int \Gamma_{\vartheta(a)}(\varphi_a, u_{a,\vartheta}) dm_{\vartheta(a)}\Big] da.
\end{aligned}$$

Proof. Let us begin by showing that $a \mapsto \rho_{a,\vartheta}$ is 2-absolutely continuous. Indeed, let $a_0 < a_1$, we have with the equivalence of the gradient estimate (81) and the Wasserstein contraction (68)

$$\begin{aligned}
&W_{\vartheta(a_0)}(\rho_{a_0,\vartheta}, \rho_{a_1,\vartheta}) \\
&\leq W_{\vartheta(a_0)}(\hat{P}_{t,\vartheta(a_0)}\rho_{a_0}, \hat{P}_{t,\vartheta(a_0)}\rho_{a_1}) + W_{\vartheta(a_0)}(\hat{P}_{t,\vartheta(a_0)}\rho_{a_1}, \hat{P}_{t,\vartheta(a_1)}\rho_{a_1}) \\
&\leq W_t(\rho_{a_0}, \rho_{a_1}) + W_{\vartheta(a_0)}(\hat{P}_{t,\vartheta(a_0)}\rho_{a_1}, \hat{P}_{t,\vartheta(a_1)}\rho_{a_1}).
\end{aligned}$$

By virtue of Lemma 2.33(iv) we have that $\tilde{\rho}_a = \hat{P}_{t,\vartheta(a)}\rho_{a_1} = \tilde{u}_a m_{\vartheta(a)}$ is in $AC^2([0,1], \mathcal{P}(X))$. This proves that $a \mapsto \rho_{a,\vartheta}$ is 2-absolutely continuous.

To conclude that $a \mapsto \int \varphi_a d\rho_{a,\vartheta}$ is absolutely continuous we write

$$\begin{aligned}
&\int \varphi_{a_1} d\rho_{a_1,\vartheta} - \int \varphi_{a_0} d\rho_{a_0,\vartheta} \\
&= \int (\varphi_{a_1} - \varphi_{a_0}) d\rho_{a_1,\vartheta} + \int \varphi_{a_0} d\rho_{a_1,\vartheta} - \int \varphi_{a_0} d\rho_{a_0,\vartheta} \\
&\leq ||\varphi_{a_1} - \varphi_{a_0}||_\infty + \mathrm{Lip}(\varphi_{a_0}) W(\rho_{a_1,\vartheta}, \rho_{a_0,\vartheta}).
\end{aligned}$$

To compute its derivative we consider difference quotients. Since $\varphi \in \mathrm{Lip}([0,1], L^\infty(X))$ is in HLS_ϑ and $u_{a+h,\vartheta} \to u_{a,\vartheta}$ in $L^1(X)$ we have

$$\lim_{h\to 0} h^{-1} \int (\varphi_{a+h} - \varphi_a) d\rho_{a+h,\vartheta} \leq -\frac{1}{2} \int |\nabla_{\vartheta(a)} \varphi_a|^2 d\rho_{a,\vartheta}. \tag{89}$$

Now we need to determine

$$\lim_{h\to 0} \frac{1}{h} (\int \varphi_a e^{-f_{\vartheta(a)}} (u_{a+h,\vartheta} - u_{a,\vartheta}) dm_\diamond + \int \varphi_a u_{a+h,\vartheta} d(m_{\vartheta(a+h)} - m_{\vartheta(a)})).$$

The expression on the right hand side clearly converges to

$$-\dot{\vartheta}(a) \int \varphi_a \dot{f}_{\vartheta(a)} u_{a,\vartheta} dm_{\vartheta(a)}, \tag{90}$$

while from Lemma 2.63 we deduce

$$\lim_{h\to 0} \int e^{-f_{\vartheta(a)}} \varphi_a \frac{1}{h} (u_{a+h,\vartheta} - u_{a,\vartheta}) dm_\diamond = \langle \partial_a u_{a,\vartheta}, e^{-f_{\vartheta(a)}} \varphi_a \rangle_{\mathcal{F},\mathcal{F}^*},$$

and after inserting

$$\langle \partial_a u_a, e^{-f_{\vartheta(a)}} \varphi_a \rangle_{\mathcal{F},\mathcal{F}^*} \tag{91}$$

$$= (t-s)\Big(\int \dot{f}_{\vartheta(a)} u_{a,\vartheta} \varphi_a e^{-f_{\vartheta(a)}} dm_\diamond + \mathcal{E}^\diamond_{\vartheta(a)}(u_{a,\vartheta}, \varphi_a e^{-f_{\vartheta(a)}}) \Big) \tag{92}$$

$$= (t-s)\Big(\int \dot{f}_{\vartheta(a)} u_{a,\vartheta} \varphi_a dm_{\vartheta(a)} + \int \Gamma_{\vartheta(a)}(u_{a,\vartheta}, \varphi_a) dm_{\vartheta(a)} \Big). \tag{93}$$

Then from the absolute continuity of $a \mapsto \int \varphi_a d\rho_{a,\vartheta}$ together with (89), (90) and (93), we obtain

$$\begin{aligned}
&\int \varphi_1 d\rho_{1,\vartheta} - \int \varphi_0 d\rho_{0,\vartheta} = \int_0^1 \partial_a \int \varphi_a d\rho_{a,\vartheta} da \\
&\leq \int_0^1 \Big[-\frac{1}{2} \int |\nabla_{\vartheta(a)} \varphi_a|^2 d\rho_{a,\vartheta} + \int P_{t,\vartheta(a)} \varphi_a \dot{u}_a dm_t \\
&\qquad - (t-s) \int \varphi_a \dot{f}_{\vartheta(a)} u_{a,\vartheta} dm_{\vartheta(a)} + (t-s) \int \dot{f}_{\vartheta(a)} u_{a,\vartheta} \varphi_a dm_{\vartheta(a)} \\
&\qquad + (t-s) \int \Gamma_{\vartheta(a)}(u_{a,\vartheta}, \varphi_a) dm_{\vartheta(a)} \Big] da \\
&\leq \int_0^1 \Big[-\frac{1}{2} \int |\nabla_{\vartheta(a)} \varphi_a|^2 d\rho_{a,\vartheta} + \int P_{t,\vartheta(a)} \varphi_a \dot{u}_a dm_t \\
&\qquad + (t-s) \int \Gamma_{\vartheta(a)}(u_{a,\vartheta}, \varphi_a) dm_{\vartheta(a)} \Big] da.
\end{aligned}$$

□

We regularize the entropy functional by truncating the singularities of the logarithm. Define $e_\varepsilon \colon [0,\infty)$ by setting $e_\varepsilon'(r) = \log(\varepsilon + r) + 1$ and $e_\varepsilon(0) = 0$. Then e_ε is still a convex function and $e_\varepsilon' \in \mathrm{Lip}_b([0,R])$. For any t and $\rho = um_t \in \mathcal{P}(X)$ we define

$$S_t^\varepsilon(\rho) = \int e_\varepsilon(u) dm_t.$$

Note that for any $\rho \in Dom(S)$ we clearly have $S^\varepsilon(\rho) \to S(\rho)$as $\varepsilon \to 0$.

As in [7] we introduce

$$p_\varepsilon(r) := e_\varepsilon'(r^2) - \log \varepsilon.$$

Lemma 2.65. *With the same notation as in Lemma 2.64 we find for any $\varepsilon > 0$*

$$\begin{aligned}
&S_t^\varepsilon(\rho_{1,\vartheta}) - S_s^\varepsilon(\rho_{0,\vartheta}) \\
&\geq \int_0^1 \int \dot{u}_a P_{t,\vartheta(a)}(e_\varepsilon'(u_{a,\vartheta})) dm_{\vartheta(a)} + 4(t-s) \int e_\varepsilon''(u_{a,\vartheta}) \Gamma_{\vartheta(a)}(\sqrt{u_{a,\vartheta}}) d\rho_{a,\vartheta} \\
&\quad + (t-s) \int \dot{f}_{\vartheta(a)}(u_{a,\vartheta} e_\varepsilon'(u_{a,\vartheta}) - e_\varepsilon'(u_{a,\vartheta})) dm_{\vartheta(a)} da.
\end{aligned}$$

Proof. From the convexity of e_ε we get for every $a_0, a_1 \in [0,1]$ by virtue of

Lemma 2.63

$$
\begin{aligned}
&S^{\varepsilon}_{\vartheta(a_1)}(\rho_{a_1,\vartheta}) - S^{\varepsilon}_{\vartheta(a_0)}(\rho_{a_0,\vartheta}) \\
=&\int e_\varepsilon(u_{a_1,\vartheta}) - e_\varepsilon(u_{a_0,\vartheta}) e^{-f_{\vartheta(a_0)}} dm_\diamond + \int e_\varepsilon(u_{a_1,\vartheta})(e^{-f_{\vartheta(a_1)}} - e^{-f_{\vartheta(a_0)}}) dm_\diamond \\
\geq &\int e'_\varepsilon(u_{a_0,\vartheta})(u_{a_1,\vartheta} - u_{a_0,\vartheta}) e^{-f_{\vartheta(a_0)}} dm_\diamond + \int e_\varepsilon(u_{a_1,\vartheta})(e^{-f_{\vartheta(a_1)}} - e^{-f_{\vartheta(a_0)}}) dm_\diamond \\
=&\int_{a_0}^{a_1} (\langle \partial_a u_{a,\vartheta}, e^{-f_{\vartheta(a_0)}} e'_\varepsilon(u_{a_0,\vartheta})\rangle - \int e_\varepsilon(u_{a_1,\vartheta})\dot\vartheta(a)\dot f_{\vartheta(a)} e^{-f_{\vartheta(a)}} dm_\diamond) da \\
=&\int_{a_0}^{a_1} (\langle -\dot\vartheta(a)\Delta_{\vartheta(a)} u_{a,\vartheta} + \dot\vartheta(a)\dot f_{\vartheta(a)} u_{a,\vartheta} + P^*_{t,\vartheta(a)}(\dot u_a), e^{-f_{\vartheta(a_0)}} e'_\varepsilon(u_{a_0,\vartheta})\rangle \\
&- \int e_\varepsilon(u_{a_1,\vartheta})\dot\vartheta(a)\dot f_{\vartheta(a)} e^{-f_{\vartheta(a)}} dm_\diamond) da \\
=&\int_{a_0}^{a_1} (-\dot\vartheta(a)\langle \Delta_{\vartheta(a)} u_{a,\vartheta}, e^{-f_{\vartheta(a_0)}} e'_\varepsilon(u_{a_0,\vartheta})\rangle + \int \dot\vartheta(a)\dot f_{\vartheta(a)} u_{a,\vartheta} e^{-f_{\vartheta(a_0)}} e'_\varepsilon(u_{a_0,\vartheta}) dm_\diamond \\
&+ \int P^*_{t,\vartheta(a)}(\dot u_a) e^{-f_{\vartheta(a_0)}} e'_\varepsilon(u_{a_0,\vartheta}) dm_\diamond - \int e_\varepsilon(u_{a_1,\vartheta})\dot\vartheta(a)\dot f_{\vartheta(a)} e^{-f_{\vartheta(a)}} dm_\diamond) da.
\end{aligned}
$$

Now fix $h > 0$ and choose a partition of $[0, 1]$ consisting of Lebesgue points $\{a_i\}_{i=0}^n$ such that $0 \leq a_{i+1} - a_i \leq h$. Then

$$
\begin{aligned}
&S^{\varepsilon}_t(\rho_{1,\vartheta}) - S^{\varepsilon}_s(\rho_{0,\vartheta}) = \sum_{i=1}^n (S^{\varepsilon}_{\vartheta(a_i)}(\rho_{a_i,\vartheta}) - S^{\varepsilon}_{\vartheta(a_{i-1})}(\rho_{a_{i-1},\vartheta})) \\
\geq &\sum_{i=1}^n \int_{a_{i-1}}^{a_i} (-\dot\vartheta(a)\langle \Delta_{\vartheta(a)} u_{a,\vartheta}, e^{-f_{\vartheta(a_{i-1})}} e'_\varepsilon(u_{a_{i-1},\vartheta})\rangle \\
&+ \int \dot\vartheta(a)\dot f_{\vartheta(a)} u_{a,\vartheta} e^{-f_{\vartheta(a_{i-1})}} e'_\varepsilon(u_{a_{i-1},\vartheta}) dm_\diamond \\
&+ \int P^*_{t,\vartheta(a)}(\dot u_a) e^{-f_{\vartheta(a_{i-1})}} e'_\varepsilon(u_{a_{i-1},\vartheta}) dm_\diamond - \int e_\varepsilon(u_{a_i,\vartheta})\dot\vartheta(a)\dot f_{\vartheta(a)} e^{-f_{\vartheta(a)}} dm_\diamond) da \\
=&\int_0^1 (-\dot\vartheta(a)\langle \Delta_{\vartheta(a)} u_{a,\vartheta}, \varsigma^h_a\rangle + \int \dot\vartheta(a)\dot f_{\vartheta(a)} u_{a,\vartheta}\varsigma^h_a dm_\diamond \\
&+ \int P^*_{t,\vartheta(a)}(\dot u_a)\varsigma^h_a dm_\diamond - \int \omega^h_a \dot\vartheta(a)\dot f_{\vartheta(a)} e^{-f_{\vartheta(a)}} dm_\diamond) da,
\end{aligned}
$$

where

$$
\begin{aligned}
\varsigma^h_a &= e^{-f_{\vartheta(a_{i-1})}} e'_\varepsilon(u_{a_{i-1},\vartheta}), \text{ for } a \in (a_{i-1}, a_i] \\
\omega^h_a &= e_\varepsilon(u_{a_i,\vartheta}), \text{ for } a \in (a_{i-1}, a_i].
\end{aligned}
$$

Letting $h \to 0$ we obtain

$$
\begin{aligned}
\varsigma^h_a &\to e^{-f_{\vartheta(a)}} e'_\varepsilon(u_{a,\vartheta}), \text{ in } L^1(X) \text{ for a.e. } a \in (0, 1) \\
\omega^h_a &\to e_\varepsilon(u_{a,\vartheta}), \text{ in } L^1(X) \text{ for a.e. } a \in (0, 1),
\end{aligned}
$$

and thus from dominated convergence

$$
\begin{aligned}
&S_t^\varepsilon(\rho_{1,\vartheta}) - S_s^\varepsilon(\rho_{0,\vartheta}) \\
&\geq \limsup_{h\to 0}\Big[\int_0^1 (-\dot\vartheta(a)\langle \Delta_{\vartheta(a)} u_{a,\vartheta}, \varsigma_a^h\rangle + \int \dot\vartheta(a)\dot f_{\vartheta(a)} u_{a,\vartheta}\varsigma_a^h dm_\diamond \\
&\quad + \int P^*_{t,\vartheta(a)}(\dot u_a)\varsigma_a^h dm_\diamond - \int \omega_a^h \dot\vartheta(a)\dot f_{\vartheta(a)} e^{-f_{\vartheta(a)}} dm_\diamond)da\Big] \\
&\geq \limsup_{h\to 0}\Big[\int_0^1 (-\dot\vartheta(a)\langle \Delta_{\vartheta(a)} u_{a,\vartheta}, \varsigma_a^h\rangle da\Big] \\
&\quad + \int_0^1 (\int \dot\vartheta(a)\dot f_{\vartheta(a)} u_{a,\vartheta} e^{-f_{\vartheta(a)}} e'_\varepsilon(u_{a,\vartheta}) dm_\diamond \\
&\quad + \int P^*_{t,\vartheta(a)}(\dot u_a) e^{-f_{\vartheta(a)}} e'_\varepsilon(u_{a,\vartheta}) dm_\diamond - \int e_\varepsilon(u_{a,\vartheta})\dot\vartheta(a)\dot f_{\vartheta(a)} e^{-f_{\vartheta(a)}} dm_\diamond)da.
\end{aligned}
$$

To see that $\langle \Delta_{\vartheta(a)} u_{a,\vartheta}, \varsigma_a^h\rangle \to \langle \Delta_{\vartheta(a)} u_{a,\vartheta}, e^{-f_{\vartheta(a)}} e'_\varepsilon(u_{a,\vartheta})\rangle$, recall that from Theorem 2.20 it suffices to show that

$$\varsigma_a^h \to e^{-f_{\vartheta(a)}} e'_\varepsilon(u_{a,\vartheta}) \text{ in } L^2(X).$$

This is a consequence of the boundedness of $u_{a,\vartheta}$ and $f_{\vartheta(a)}$. Then again by dominated convergence we have

$$
\begin{aligned}
&S_t^\varepsilon(\rho_{1,\vartheta}) - S_s^\varepsilon(\rho_{0,\vartheta}) \\
&\geq \int_0^1 [\dot\vartheta(a)\mathcal{E}^\diamond_{\vartheta(a)}(u_{a,\vartheta}, e^{-f_{\vartheta(a)}} e'_\varepsilon(u_{a,\vartheta})) + \int \dot\vartheta(a)\dot f_{\vartheta(a)} u_{a,\vartheta} e^{-f_{\vartheta(a)}} e'_\varepsilon(u_{a,\vartheta}) dm_\diamond \\
&\quad + \int P^*_{t,\vartheta(a)}(\dot u_a) e^{-f_{\vartheta(a)}} e'_\varepsilon(u_{a,\vartheta}) dm_\diamond - \int e_\varepsilon(u_{a,\vartheta})\dot\vartheta(a)\dot f_{\vartheta(a)} e^{-f_{\vartheta(a)}} dm_\diamond] da \\
&= \int_0^1 [\dot\vartheta(a)\mathcal{E}_{\vartheta(a)}(u_{a,\vartheta}, e'_\varepsilon(u_{a,\vartheta})) + \int \dot\vartheta(a)\dot f_{\vartheta(a)} u_{a,\vartheta} e'_\varepsilon(u_{a,\vartheta}) dm_{\vartheta(a)} \\
&\quad + \int P^*_{t,\vartheta(a)}(\dot u_a) e'_\varepsilon(u_{a,\vartheta}) dm_{\vartheta(a)} - \int e_\varepsilon(u_{a,\vartheta})\dot\vartheta(a)\dot f_{\vartheta(a)} dm_{\vartheta(a)}] da.
\end{aligned}
$$

□

2.9.3 The Dynamic EVI$^-$-Property

Proposition 2.66. *Let $\rho^a = u^a m_t$ be a regular curve. Then setting $\rho^a_\vartheta = \hat P_{t,\vartheta(a)}\rho^a$, it holds*

$$
\begin{aligned}
&\frac{1}{2}\tilde W^2_\vartheta(\rho_{1,\vartheta}, \rho_{0,\vartheta}) - (t-s)(S_t(\rho_{1,\vartheta}) - S_s(\rho_{0,\vartheta})) \\
&\leq \frac{1}{2}\int_0^1 |\dot\rho_a|_t^2 da - (t-s)^2 \int_0^1 \int \dot f_{\vartheta(a)} d\rho_{a,\vartheta} da.
\end{aligned}
\tag{94}
$$

Proof. Applying Lemma 2.64 and Lemma 2.65, we find

$$\begin{aligned}
&\int \varphi_1 d\rho_{1,\vartheta} - \int \varphi_0 d\rho_{0,\vartheta} - (t-s)(S_t^\varepsilon(\rho_{1,\vartheta}) - S_s^\varepsilon(\rho_{0,\vartheta})) \\
&\le \int_0^1 \Big[\int \dot{u}_a P_{t,\vartheta(a)}(\varphi_a - (t-s)e'_\varepsilon(u_{a,\vartheta}))dm_t - \frac{1}{2}\int \Gamma_{\vartheta(a)}(\varphi_a)d\rho_{a,\vartheta} \\
&+ (t-s)\int \Gamma_{\vartheta(a)}(\varphi_a, u_{a,\vartheta})dm_{\vartheta(a)} - 4(t-s)^2 \int e''_\varepsilon(u_{a,\vartheta})\Gamma_{\vartheta(a)}(\sqrt{u_{a,\vartheta}})d\rho_{a,\vartheta} \\
&- (t-s)^2 \int (e_\varepsilon(u_{a,\vartheta}) - e'_\varepsilon(u_{a,\vartheta})u_{a,\vartheta})\dot{f}_{\vartheta(a)}dm_{\vartheta(a)}\Big] da.
\end{aligned} \tag{95}$$

Then since

$$4re''_\varepsilon(r) \ge 4r^2(e''_\varepsilon(r))^2 = r(p'_\varepsilon(\sqrt{r}))^2,$$

we can estimate

$$\begin{aligned}
-4u_{a,\vartheta}e''_\varepsilon(u_{a,\vartheta})\Gamma_{\vartheta(a)}(\sqrt{u_{a,\vartheta}}) &\le -u_{a,\vartheta}(p'_\varepsilon(\sqrt{u_{a,\vartheta}}))^2\Gamma_{\vartheta(a)}(\sqrt{u_{a,\vartheta}}) \\
&= -u_{a,\vartheta}\Gamma_{\vartheta(a)}(p_\varepsilon(\sqrt{u_{a,\vartheta}})),
\end{aligned}$$

and while, with $q_\varepsilon(r) := \sqrt{r}(2 - \sqrt{r}p'_\varepsilon(\sqrt{r}))$,

$$\begin{aligned}
\Gamma_{\vartheta(a)}(u_{a,\vartheta}, \varphi_a) &= 2\sqrt{u_{a,\vartheta}}\Gamma_{\vartheta(a)}(\sqrt{u_{a,\vartheta}}, \varphi_a) \\
&= u_{a,\vartheta}\Gamma_{\vartheta(a)}(p_\varepsilon(\sqrt{u_{a,\vartheta}}), \varphi_a) + q_\varepsilon(u_{a,\vartheta})\Gamma_{\vartheta(a)}(\sqrt{u_{a,\vartheta}}, \varphi_a)
\end{aligned}$$

we find

$$\begin{aligned}
&\int \varphi_1 d\rho_{1,\vartheta} - \int \varphi_0 d\rho_{0,\vartheta} - (t-s)(S_t^\varepsilon(\rho_{1,\vartheta}) - S_s^\varepsilon(\rho_{0,\vartheta})) \\
&\le \int_0^1 \Big[\int \dot{u}_a P_{t,\vartheta(a)}(\varphi_a - (t-s)e'_\varepsilon(u_{a,\vartheta}))dm_t - \frac{1}{2}\int \Gamma_{\vartheta(a)}(\varphi_a)d\rho_{a,\vartheta} \\
&+ (t-s)\int \Gamma_{\vartheta(a)}(\varphi_a, p_\varepsilon(\sqrt{u_{a,\vartheta}}))d\rho_{a,\vartheta} - (t-s)^2 \int \Gamma_{\vartheta(a)}(p_\varepsilon(\sqrt{u_{a,\vartheta}}))d\rho_{a,\vartheta} \\
&+ (t-s)\int q_\varepsilon(u_{a,\vartheta})\Gamma_{\vartheta(a)}(\sqrt{u_{a,\vartheta}}, \varphi_a)dm_{\vartheta(a)} \\
&- (t-s)^2 \int (e_\varepsilon(u_{a,\vartheta}) - e'_\varepsilon(u_{a,\vartheta})u_{a,\vartheta})\dot{f}_{\vartheta(a)}dm_{\vartheta(a)}\Big] da.
\end{aligned} \tag{96}$$

Hence, by means of (87), the gradient estimate (81), and Young inequality

$2xy \le \delta x^2 + y^2/\delta$ this yields

$$\begin{aligned}
&\int \varphi_1 d\rho_{1,\vartheta} - \int \varphi_0 d\rho_{0,\vartheta} - (t-s)(S_t^\varepsilon(\rho_{1,\vartheta}) - S_s^\varepsilon(\rho_{0,\vartheta}))\\
&\le \int_0^1 \Big[\frac{1}{2}|\dot\rho_a|_t^2 + \frac{1}{2}\int \Gamma_t(P_{t,\vartheta(a)}(\varphi_a - (t-s)e_\varepsilon'(u_{a,\vartheta}))d\rho_a\\
&\qquad - \frac{1}{2}\int P_{t,\vartheta(a)}\Gamma_{\vartheta(a)}(\varphi_a - (t-s)p_\varepsilon(\sqrt{u_{a,\vartheta}}))d\rho_a\\
&\qquad + (t-s)\int q_\varepsilon(u_{a,\vartheta})\Gamma_{\vartheta(a)}(\sqrt{u_{a,\vartheta}},\varphi_a)dm_{\vartheta(a)}\\
&\qquad - (t-s)^2\int (e_\varepsilon(u_{a,\vartheta}) - e_\varepsilon'(u_{a,\vartheta})u_{a,\vartheta})\dot f_{\vartheta(a)}dm_{\vartheta(a)}\Big]da\\
&\le \int_0^1 \Big[\frac{1}{2}|\dot\rho_a|_t^2 + +(t-s)\int |q_\varepsilon(u_{a,\vartheta})||\Gamma_{\vartheta(a)}(\sqrt{u_{a,\vartheta}},\varphi_a)|dm_{\vartheta(a)}\\
&\qquad - (t-s)^2\int (e_\varepsilon(u_{a,\vartheta}) - e_\varepsilon'(u_{a,\vartheta})u_{a,\vartheta})\dot f_{\vartheta(a)}dm_{\vartheta(a)}\Big]da\\
&\le \int_0^1 \Big[\frac{1}{2}|\dot\rho_a|_t^2 + \frac{(t-s)}{2\delta}\int (q_\varepsilon(u_{a,\vartheta}))^2\Gamma_{\vartheta(a)}(\varphi_a)dm_{\vartheta(a)}\\
&\qquad + \frac{(t-s)\delta}{2}\int \Gamma_{\vartheta(a)}(\sqrt{u_{a,\vartheta}})dm_{\vartheta(a)}\\
&\qquad - (t-s)^2\int (e_\varepsilon(u_{a,\vartheta}) - e_\varepsilon'(u_{a,\vartheta})u_{a,\vartheta})\dot f_{\vartheta(a)}dm_{\vartheta(a)}\Big]da.
\end{aligned}$$

We first pass to the limit $\varepsilon \to 0$,

$$\lim_{\varepsilon\to 0} q_\varepsilon^2(r) = 0, \quad q_\varepsilon^2(r) = 4r(1 - \frac{r}{\varepsilon + r})^2 \le 4r,$$

$\lim_{\varepsilon\to 0}(e_\varepsilon(r) - re_\varepsilon'(r)) = -r,$

$$|e_\varepsilon(r) - re_\varepsilon'(r)| \le 2(\varepsilon + r)|\log(\varepsilon + r)| + r + \varepsilon\log\varepsilon \le 2\sqrt{\varepsilon + r} + r + \varepsilon\log\varepsilon,$$

and then, $\delta \to 0$,

$$\begin{aligned}
&\int \varphi_1 d\rho_{1,\vartheta} - \int \varphi_0 d\rho_{0,\vartheta} - (t-s)(S_t(\rho_{1,\vartheta}) - S_s(\rho_{0,\vartheta}))\\
&\le \int_0^1 \Big[\frac{1}{2}|\dot\rho_a|_t^2 + (t-s)^2\int \dot f_{\vartheta(a)}d\rho_{a,\vartheta}\Big]da.
\end{aligned}$$

Taking the supremum over φ we obtain the desired estimate (94). $\square$

Theorem 2.67. *Assume that the gradient estimate holds true for the time-dependent metric measure space $(X, d_t, m_t)_{t\in(0,T)}$. Then for every $\mu \in Dom(S)$ and every $\tau \in (0,T]$ the dual heat flow $\mu_t := \hat P_{t,\tau}\mu$ emanating in μ we have*

$$S_s(\mu_s) - S_t(\sigma) \le \frac{1}{2(t-s)}(W_t^2(\mu_t,\sigma) - W_{s,t}^2(\mu_s,\sigma)) - (t-s)\int_0^1\int \dot f_{\vartheta(a)}d\rho_{a,\vartheta}da \tag{97}$$

for all $s \in (0,\tau)$ and all $\sigma, \mu \in Dom(S)$. Here $(\rho_a)_{a\in[0,1]}$ denotes the W_t-geodesic connecting $\rho_0 = \mu_t$, $\rho_1 = \sigma$ and $\rho_{a,\vartheta} = \hat P_{t,\vartheta(a)}(\rho_a)$.

In particular μ_t is a dynamic upward EVI$^-$-gradient flow, i.e. for every $t \in (0,\tau)$ and every $\sigma \in Dom(S)$ we have

$$\frac{1}{2}\partial_s^- W_{s,t}^2(\mu_s,\sigma)_{|s=t-} \geq S_t(\mu_t) - S_t(\sigma).$$

Proof. Let $(\rho_a)_{a\in[0,1]}$ be a W_t-geodesic connecting μ_t and σ, which exists and is unique. We approximate the geodesic $(\rho_a)_{a\in[0,1]}$ by regular curves $(\rho_a^n)_{a\in[0,1]}$. Proposition 2.66 states that for each $(\rho_a^n)_{a\in[0,1]}$

$$\begin{aligned}
&\frac{1}{2}\tilde{W}_\vartheta^2(\rho_{1,\vartheta}^n,\rho_{0,\vartheta}^n) - (t-s)(S_t(\rho_{1,\vartheta}^n) - S_s(\rho_{0,\vartheta}^n)) \\
&\leq \frac{1}{2}\int_0^1 |\dot{\rho}_a^n|_t^2 da - (t-s)^2 \int_0^1 \int \dot{f}_{\vartheta(a)} d\rho_{a,\vartheta}^n da.
\end{aligned} \tag{98}$$

Since for every $a\in[0,1]$ ρ_a^n converges to ρ_a in duality with bounded continuous functions, $\rho_{a,\vartheta}^n$ converges to $\rho_{a,\vartheta}$ in duality with bounded continuous functions as well. By virtue of Lemma 2.59 we obtain

$$\liminf_{n\to\infty} \tilde{W}_\vartheta^2(\rho_{1,\vartheta}^n,\rho_{0,\vartheta}^n) \geq \tilde{W}_\vartheta^2(\rho_{1,\vartheta},\rho_{0,\vartheta}).$$

Note that (ρ_a^n) also converges to ρ_a in duality with L^∞ functions, since Lemma 2.28 provides $\sup_n S_t(\rho_a^n) < \infty$. The same argument applies then to $\rho_{a\vartheta}^n$. Hence

$$\lim_{n\to\infty} \int \dot{f}_{\vartheta(a)} d\rho_{a,\vartheta}^n = \int \dot{f}_{\vartheta(a)} d\rho_{a,\vartheta}.$$

Then we end up with

$$\begin{aligned}
&\frac{1}{2}\tilde{W}_\vartheta^2(\mu_s,\sigma) - (t-s)(S_t(\sigma) - S_s(\mu_s)) \\
&\leq \frac{1}{2}W_t^2(\mu_t,\sigma) - (t-s)^2 \int_0^1 \int \dot{f}_{\vartheta(a)} d\rho_{a,\vartheta} da.
\end{aligned} \tag{99}$$

Applying Corollary 2.61 we obtain

$$\begin{aligned}
&(t-s)(S_s(\mu_s) - S_t(\sigma)) \\
&\leq \frac{1}{2}W_t^2(\mu_t,\sigma) - \frac{1}{2}W_{s,t}^2(\mu_s,\sigma) - (t-s)^2 \int_0^1 \int \dot{f}_{\vartheta(a)} d\rho_{a,\vartheta} da.
\end{aligned}$$

Dividing by $t-s$ and letting $s \nearrow t$ we find

$$\begin{aligned}
S_t(\mu_t) - S_t(\sigma) &\leq \liminf_{s\nearrow t} \frac{1}{2(t-s)}\left(W_t^2(\mu_t,\sigma) - W_{s,t}^2(\mu_s,\sigma)\right) \\
&= \frac{1}{2}\partial_s^- W_{s,t}^2(\mu_s,\sigma)_{|s=t-}.
\end{aligned}$$

□

2.9.4 Summarizing

The precise integrated version (97) of the EVI$^-$-property indeed also implies a relaxed version of the EVI$^+$-property which then in turn allows to prove uniqueness of dynamic EVI-flows for the entropy.

Corollary 2.68. *The gradient estimate* **(III)** *implies the* $\mathbf{EVI}^+(-2L,\infty)$*-property. More precisely, for every* $\mu\in Dom(S)$ *and every* $\tau\le T$ *the dual heat flow* $\mu_t:=\hat P_{t,\tau}\mu$ *emanating in* μ *satisfies*

$$\frac12\partial_s^- W_{s,t}^2(\mu_s,\sigma)_{|s=t}\ge S_t(\mu_t)-S_t(\sigma)-L\,W_t^2(\mu_t,\sigma)$$

for all $t<\tau$ *and all* $\sigma\in\mathcal P(X)$.

Proof. Given $\mu_t:=\hat P_{t,\tau}\mu$ for $t\tau$, consider (97) for fixed $s<\tau$ and with $s\searrow t$. Then

$$\begin{aligned}
S_s(\mu_s)-S_s(\sigma) &= \lim_{s\searrow t} S_s(\mu_s)-S_t(\sigma)\\
&\le \lim_{s\searrow t}\frac{1}{2(t-s)}\Big[W_t^2(\mu_t,\sigma)-W_{s,t}^2(\mu_s,\sigma)\Big]\\
&\le \Big(\lim_{s\searrow t}\frac{1}{2(t-s)}\Big[W_{t,s}^2(\mu_t,\sigma)-W_s^2(\mu_s,\sigma)\Big]\\
&\qquad\qquad +\frac{L}{2}\Big[W_t^2(\mu_t,\sigma)+W_s^2(\mu_s,\sigma)\Big]\Big)\\
&= \frac12\partial_t^- W_{t,s}^2(\mu_t,\sigma)_{t=s+}+L\,W_s^2(\mu_s,\sigma)
\end{aligned}$$

where the last estimate follows from (102). ☐

Corollary 2.69. *Assume that* **(III)** *holds true and that* $(\mu_t)_{t\in(\sigma,\tau)}$ *is a dynamic upward EVI*$^-$*- or EVI*$^+$*gradient flow for* S *emanating in some* $\mu\in\mathcal P$*. Then*

$$\mu_t=\hat P_{t,\tau}\mu$$

for all $t\in(\sigma,\tau)$*. That is, the dual heat flow is the unique dynamic backward EVI*$^-$*-flow for the Boltzmann entropy.*

Proof. Corollary 2.78 together with Corollary 2.68 and Theorem 2.67. ☐

Theorem 2.70. *The gradient estimate* $(\mathbf{III}_N)$ *implies the dynamic* N*-convexity of the Boltzmann entropy* $(\mathbf{I}_N)$*.*

Proof. According to Theorem 2.41 and Theorem 2.67 the gradient estimate $(\mathbf{III}_N)$ implies both

- the transport estimate $(\mathbf{II}_N)$ and
- the $\mathbf{EVI}^-(0,\infty)$-property

According to Theorem 2.80, both properties together imply dynamic N-convexity. ☐

2.10 EVI, Contraction Estimates and Dynamic Convexity

2.10.1 Time-dependent Geodesic Spaces

For this chapter, our basic setting will be a space X equipped with a 1-parameter family of complete geodesic metrics $(d_t)_{t\in I}$ where $I \subset \mathbb{R}$ is a bounded open interval, say for convenience $I = (0,T)$. (More generally, one might allow d_t to be pseudo metrics where the existence of connecting geodesics is only requested for pairs $x,y \in X$ with $d_t(x,y) < \infty$.) We always request that there exists a constant $L \in \mathbb{R}$ ('log-Lipschitz bound') such that

$$\left|\log \frac{d_t(x,y)}{d_s(x,y)}\right| \le L \cdot |t-s| \tag{100}$$

for all s,t and all x,y ('log Lipschitz continuity in t');

Let us first introduce a natural 'distance' on $I \times X$.

Definition 2.71. *Given $s,t \in I$ and $x,y \in X$ we put*

$$d_{s,t}(x,y) := \inf \left\{ \int_0^1 |\dot\gamma^a|^2_{s+a(t-s)} da \right\}^{1/2} \tag{101}$$

where the infimum runs over all absolutely continuous curves $(\gamma^a)_{a\in[0,1]}$ in X connecting x and y.

Proposition 2.72. *(i) The infimum in the above formula is attained. Each minimizer $(\gamma^a)_{a\in[0,1]}$ is a curve of constant speed, i.e. $|\dot\gamma^a|_{s+a(t-s)} = d_{s,t}(x,y)$ for all $a \in [0,1]$.*

(ii) A point $z \in X$ lies on some minimizing curve γ with $z = \gamma^a$ if and only if

$$d_{s,t}(x,y) = d_{s,r}(x,z) + d_{r,t}(z,y)$$

with $r = s + a(t-s)$.

(iii) For all $s,t \in I$ and $x,y \in X$

$$\frac{1-e^{-L|t-s|}}{L|t-s|} \le \frac{d_{s,t}(x,y)}{d_s(x,y)} \le \frac{e^{L|t-s|}-1}{L|t-s|}.$$

Thus in particular,

$$\left|\partial_t d_{s,t}(x,y)\big|_{t=s}\right| \le \frac{L}{2} d_s(x,y). \tag{102}$$

(iv) For all $s < t \in I$ and $x,y \in X$

$$d_{s,t}(x,y) = \lim_{\delta\to 0} \inf_{(t_i,x_i)_i} \left\{ \sum_{i=1}^k \frac{t-s}{t_i - t_{i-1}} \, d^2_{t_i}\big(x_i, x_{i-1}\big) \right\}^{1/2} \tag{103}$$

where the infimum runs over all $k \in \mathbb{N}$. all partitions $(t_i)_{i=0,\dots,k}$ of $[s,t]$ with $t_0 = s, t_k = 1$ and $|t_i - t_{i-1}| \le \delta$ as well as over all $x_i \in X$ with $x_0 = x, x_k = y$.

Proof. (i) For each absolutely continuous curve $(\gamma^a)_{a\in[0,1]}$

$$\left(\int_0^1 |\dot\gamma^a|^2_{s+a(t-s)} da\right)^{1/2} \ge \int_0^1 |\dot\gamma^a|_{s+a(t-s)} da$$

with equality if and only if the curve has constant speed.

(ii) Restricting the minimizing curve for $d_{s,t}$ to parameter intervals $[0,a]$ and $[a,1]$ provides upper estimates for $d_{s,r}(x,z)$ and $d_{r,t}(z,y)$, , resp., and thus yields the "$\geq$"-inequality. Conversely, given any pair of minimizers for $d_{s,r}(x,z)$ and $d_{r,t}(z,y)$ by concatenation a curve connecting x and y can be constructed with action bounded by the scaled action of the two ingredients. This proves the "$\leq$"-inequality.

(iii) The log-Lipschitz continuity of the distance implies that for each absolutely continuous curve

$$e^{-La|t-s|}\int_0^1 |\dot\gamma^a|_s da \leq \int_0^1 |\dot\gamma^a|_{s+a(t-s)} da \leq e^{La|t-s|}\int_0^1 |\dot\gamma^a|_s da.$$

(iv) see section 2.9.1 for the argument in the case of $W_{s,t}$. □

2.10.2 EVI Formulation of Gradient Flows

For the subsequent discussion, a lower semi-bounded function $V : I \times X \to (-\infty,\infty]$ will be given with $V_s(x) \leq C_0 \cdot V_t(x) + C_1$ for all $s,t \in I$ and $x \in X$ (thus, in particular, $Dom(V) = \{x \in X : \ V_t(x) < \infty\}$ is independent of x) and such that for each $t \in I$ the function $x \mapsto V_t(x)$ is κ-convex along each d_t-geodesic (for some $\kappa \in \mathbb{R}$). We also assume that minimizing d_t-geodesics between pairs of points in $Dom(V)$ are unique.

In previous chapters, the following results will be applied

- to the Boltzmann entropy S_t on the time-dependent geodesic space $(\mathcal{P}, W_t)_{t\in I}$ as well as
- to the Dirichlet energy $\mathcal{E}_t$ on the time-dependent geodesic space $L^2(X,m_t)_{t\in I}$

in the place of the function V_t on the time-dependent geodesic space $(X,d_t)_{t\in I}$.

Definition 2.73. *Given a left-open interval $J \subset I$, an absolutely continuous curve $(x_t)_{t\in J}$ will be called* dynamic backward EVI$^-$-gradient flow *for V if for all $t \in J$ and all $z \in Dom(V_t)$*

$$\frac{1}{2}\partial_s^- d^2_{s,t}(x_s,z)\Big|_{s=t-} \geq V_t(x_t) - V_t(z) \tag{104}$$

where $d_{s,t}$ is defined in Definition 2.71.

A curve $(x_t)_{t\in J}$ with a right-open interval $J \subset I$ will be called dynamic backward EVI$^+$-gradient flow *for V if instead*

$$\frac{1}{2}\partial_s^- d^2_{s,t}(x_s,z)\Big|_{s=t+} \geq V_t(x_t) - V_t(z)$$

for all $t \in J$.

It is called dynamic backward EVI-gradient flow *if it is both, a* dynamic backward EVI$^+$-gradient flow *and a* dynamic backward EVI$^-$-gradient flow.

We say that the backward gradient flow $(x_t)_{t\in J}$ emanates *in $x' \in X$ if $\lim_{t\nearrow \sup J} x_t = x'$.*

Being a dynamic backward EVI$^\pm$-gradient flow for V obviously implies that $x_t \in Dom(V_t)$ for all $t < \tau$.

Remark. *Note that these definitions are slightly different from a previous one presented in [59]. If d_s depends smoothly on s then*

$$\partial_s^- d_{s,t}^2(x_s, z)\big|_{s=t-} = \partial_s^- d_t^2(x_s, z)\big|_{s=t-} + \partial_s^- d_{s,t}^2(x_t, z)\big|_{s=t-}$$

and always $\partial_s^- d_{s,t}^2(x_t, z)\big|_{s=t-} \geq \mathfrak{b}_t^0(\gamma)$ for any $d_t$geodesic γ connecting x_t and z.

Often, we ask for an improved notion of dynamic backward EVI-gradient flows, involving parameters $N \in [1, \infty]$ (regarded as an upper bound for the 'dimension') and/or $K \in \mathbb{R}$ (regarded as a lower bound for the 'curvature'). The choices $N = \infty$ and $K = 0$ will yield the previous concept.

Definition 2.74. *We say that an absolutely continuous curve $(x_t)_{t\in(\sigma,\tau)}$ is a* dynamic backward EVI(K, N)-gradient flow *for V if for all $z \in Dom(V_t)$ and all $t \in (\sigma, \tau)$*

$$\begin{aligned} &\frac{1}{2}\partial_s^- d_{s,t}^2(x_s, z)\Big|_{s=t} - \frac{K}{2} \cdot d_t^2(x_t, z) \\ &\geq V_t(x_t) - V_t(z) + \frac{1}{N}\int_0^1 \Big(\partial_a V_t(\gamma^a)\Big)^2 (1-a)da \end{aligned} \tag{105}$$

where γ denotes the d_t-geodesic connecting x_t and z.

Analogously, we define dynamic backward EVI$^{\pm}(K, N)$-gradient flows *for V.*

In the case, $K = 0$, dynamic backward EVI(K, N)-gradient flows will be simply called dynamic backward EVI$_N$-gradient flows.

The concept of 'backward' gradient flows is tailor-made for our later application to the dual heat flow. This flow is running backward in time and on its way it tries to minimize the Boltzmann entropy. Regarded in positive time direction, it follows the 'upward gradient' of the entropy.

On the other hand, in calculus of variations mostly the 'downward' gradient flow will be considered where a curve tries to follow the negative gradient of a given functional.

Definition 2.75. *We say that an absolutely continuous curve $(x_t)_{t\in(\sigma,\tau)}$ is a* dynamic forward EVI(K, N)-gradient flow *for V if for all $z \in Dom(V_t)$ and all $t \in (\sigma, \tau)$*

$$\begin{aligned} &-\frac{1}{2}\partial_s^+ d_{s,t}^2(x_s, z)\Big|_{s=t} - \frac{K}{2} \cdot d_t^2(x_t, z) \\ &\geq V_t(x_t) - V_t(z) + \frac{1}{N}\int_0^1 \Big(\partial_a V_t(\gamma^a)\Big)^2 (1-a)da \end{aligned} \tag{106}$$

where γ denotes the d_t-geodesic connecting x_t and z.

We say that a forward gradient flow emanates in a given point $x' \in X$ if $\lim_{t\searrow\sigma} x_t = x'$.

We will formulate all our results for 'backward' gradient flows and leave it to the reader to carry them over to the case of 'forward' gradient flows.

Lemma 2.76. *For each dynamic backward EVI$^{\pm}(K, \infty)$-gradient flow $(x_t)_{t\in(\sigma,\tau)}$ for V*

$$\int_\sigma^\tau V_t(x_t)dt < \infty.$$

Proof. Choose $z \in Dom(V)$, apply the EVI(K,∞)-property at time t, and then integrate w.r.t. time t

$$\begin{aligned}
\int_\sigma^\tau V_t(x_t)dt &\le \int_\sigma^\tau \Big[V_t(z) + \frac{1}{2}\partial_s d^2_{s,t}(x_s,z)\big|_{s=t} - \frac{K}{2}d_t^2(x_t,z)\Big]dt \\
\le\ & (C_0\,V_\tau(z)+C_1)(\tau-\sigma) + \frac{1}{2}\int_\sigma^\tau \Big[\partial_t d_t^2(x_t,z) + (L-K)\,d_t^2(x_t,z)\Big]dt \\
=\ & (C_0\,V_\tau(z)+C_1)(\tau-\sigma) + \frac{1}{2}d_\tau^2(x_\tau,z) - \frac{1}{2}d_\sigma^2(x_\sigma,z) + \frac{L-K}{2}\int_\sigma^\tau d_t^2(x_t,z)dt.
\end{aligned}$$

Obviously, the right hand side is finite which thus proves the claim. $\square$

2.10.3 Contraction Estimates

Theorem 2.77. *Given two curves $(x_t)_{t\in(\sigma,\tau)}$ and $(y_t)_{t\in(\sigma,\tau)}$, one of which is an is a dynamic backward EVI$^-(K,N)$-gradient flow for V and the other is a dynamic backward EVI$^+(K,N)$-gradient flow for V, then for all $\sigma < s < t < \tau$*

$$d_s^2(x_s,y_s) \le e^{-2K(t-s)}\cdot d_t^2(x_t,y_t) - \frac{2}{N}\int_s^t e^{-2K(r-s)}\cdot\Big|V_r(x_r)-V_r(y_r)\Big|^2 dr. \tag{107}$$

Proof. Assume that the curve $(x_t)_{t\in(\sigma,\tau]}$ is a dynamic backward EVI$^-$-gradient flow for V and $(y_t)_{t\in(\sigma,\tau]}$ is a dynamic backward EVI$^+$-gradient flow for V. It implies that $r\mapsto d_r(x_r,y_r)$ is absolutely continuous since

$$|d_t(x_t,y_t) - d_s(x_s,y_s)| \le d_s(x_s,x_t) + d_s(y_s,y_t) + L(t-s)d_t(x_t,y_t).$$

Thus by the very definition of EVI flows

$$\begin{aligned}
& d_t^2(x_t,y_t) - d_s^2(x_s,y_s) = \limsup_{\delta\searrow 0}\Big[\frac{1}{\delta}\int_{t-\delta}^t d_r^2(x_r,y_r)\,dr - \frac{1}{\delta}\int_s^{s+\delta} d_r^2(x_r,y_r)\,dr\Big] \\
& = \limsup_{\delta\searrow 0}\frac{1}{\delta}\int_{s+\delta}^t \Big[d_r^2(x_r,y_r) - d^2_{r-\delta}(x_{r-\delta},y_{r-\delta})\Big]\,dr \\
& \ge \liminf_{\delta\searrow 0}\frac{1}{\delta}\int_{s+\delta}^t \big[d_r^2(x_r,y_r) - d^2_{r,r-\delta}(x_r,y_{r-\delta})\big]\,dr \\
& + \liminf_{\delta\searrow 0}\frac{1}{\delta}\int_{s+\delta}^t \big[d^2_{r,r-\delta}(x_r,y_{r-\delta}) - d^2_{r-\delta}(x_{r-\delta},y_{r-\delta})\big]\,dr \\
& = \liminf_{\delta\searrow 0}\frac{1}{\delta}\int_{s+\delta}^t \big[d_r^2(x_r,y_r) - d^2_{r,r-\delta}(x_r,y_{r-\delta})\big]\,dr \\
& + \liminf_{\delta\searrow 0}\frac{1}{\delta}\int_s^{t-\delta} \big[d^2_{r+\delta,r}(x_{r+\delta},y_r) - d_r^2(x_r,y_r)\big]\,dr.
\end{aligned}$$

If we could interchange the liminf with integration we calculate further

$$
\begin{aligned}
&\overset{(*)}{\geq} \int_s^t \liminf_{\delta\searrow 0} \frac{1}{\delta}\big[d_r^2(x_r,y_r) - d^2_{r,r-\delta}(x_r,y_{r-\delta})\big]\,dr \\
&+ \int_s^t \liminf_{\delta\searrow 0} \frac{1}{\delta}\big[d^2_{r+\delta,r}(x_{r+\delta},y_r) - d_r^2(x_r,y_r)\big]\,dr \\
&\geq 2\int_s^t \Big[\frac{K}{2} d_r^2(x_r,y_r) + V_r(y_r) - V_r(x_r) + \frac{1}{N}\int_0^1 \Big(\partial_a V_r(\gamma_r^a)\Big)^2 a\,da\Big]dr \\
&+ 2\int_s^t \Big[\frac{K}{2} d_r^2(x_r,y_r) + V_r(x_r) - V_r(y_r) + \frac{1}{N}\int_0^1 \Big(\partial_a V_r(\gamma_r^a)\Big)^2 (1-a)\,da\Big]dr \\
&= 2K\int_s^t d_r^2(x_r,y_r)\,dr + \frac{2}{N}\int_s^t\int_0^1 \Big(\partial_a V_r(\gamma_r^a)\Big)^2 da\,dr \\
&\geq 2K\int_s^t d_r^2(x_r,y_r)\,dr + \frac{2}{N}\int_s^t \Big|V_r(x_r) - V_r(y_r)\Big|^2 dr.
\end{aligned}
$$

Dividing by $t-s$ and passing to the limit $t-s\searrow 0$ yields

$$
\partial_t d_t^2(x_t,y_t) \geq 2K d_t^2(x_t,y_t) + \frac{2}{N}\Big|V_t(x_t) - V_t(y_t)\Big|^2
$$

for a.e. t. The claim now follows via 'variation of constants'.

It remains to justify the interchange of $\liminf_{\delta\searrow 0}$ and $\int \ldots dr$ in $(*)$ which requires quite some effort. Recall from Proposition 2.72 that $|\frac{d^2_{s,t}(x,y)}{d^2_s(x,y)} - 1| \leq 2L\cdot|t-s|$ for all x,y,s,t with $|t-s|\leq\frac{1}{L}$. Thus we can estimate

$$
\begin{aligned}
&-\frac{1}{\delta}\Big[d_r^2(x_r,y_r) - d^2_{r,r-\delta}(x_r,y_{r-\delta})\Big] \\
&\quad\leq\quad -\frac{1}{\delta}\Big[d_r^2(x_r,y_r) - d^2_{r-\delta}(x_r,y_{r-\delta})\Big] + o_1 \\
&\quad=\quad -\frac{1}{\delta}\int_{r-\delta}^r \partial_s d_s^2(x_r,y_s)\,ds + o_1 \\
&\quad\leq\quad -\frac{1}{\delta}\int_{r-\delta}^r \partial_t d^2_{s,t}(x_r,y_t)\Big|_{t=s} ds + o_1 + o_2 \\
&\quad\leq\quad \frac{2}{\delta}\int_{r-\delta}^r \Big[V_s(x_r) - V_s(y_s)\Big]ds + o_1 + o_2 + o_3 \\
&\quad\leq\quad 2C_0\cdot V_r(x_r) + 2C_1 + C + o_1 + o_2 + o_3
\end{aligned}
$$

where for the last inequality we used the growth estimate of $s\mapsto V_s(x)$ and the lower boundedness of V and where we put with $o_1(r,\delta) = 2L\,d_r^2(x_r,y_{r-\delta})$, $o_2(r,\delta) = 2L\,\frac{1}{\delta}\int_{r-\delta}^r d_r^2(x_r,y_\sigma)\,d\sigma$, $o_3(r) = K\,d_r^2(x_r,y_r)$. Continuity of $r\mapsto d_r$ and of $r\mapsto x_r$ as well as of $r\mapsto y_r$ imply that for any fixed $z\in X$ the function $r\mapsto d_r^2(x_r,z)$ is bounded as well as $r\mapsto d_r^2(y_{r-\delta},z)$ for $r\in(s,t)$, uniformly in $\delta\in(0,1)$. Thus $o_1(r,\delta)+o_2(r,\delta)+o_3(r,\delta)\leq C'$ which finally justifies the interchange of limit and integral.

Similarly, we can estimate

$$\begin{aligned}
& -\frac{1}{\delta}\Big[d^2_{r+\delta,r}(x_{r+\delta},y_r)-d^2_r(x_r,y_r)\Big] \\
\le\ & -\frac{1}{\delta}\int_r^{r+\delta}\partial_s d_s^2(x_s,y_r)\,ds+o_1' \\
\le\ & 2C_0\cdot V_r(y_r)+2C_0+C+o_1'+o_2'+o_3'.
\end{aligned}$$

In both cases, the final expression is integrable w.r.t. $r\in[s,t]$ according to Lemma 2.76 since by assumption $V_t(x_t)<\infty$ as well as $V_t(y_t)<\infty$. $\square$

Corollary 2.78. *Assume that $(x_t)_{t\in(\sigma,\tau)}$ is a dynamic backward EVI(K,N)-gradient flow for V and that $(y_t)_{t\in(\sigma,\tau)}$ is a dynamic backward EVI$^-(K,N)$- or EVI$^+(K,N)$-gradient flow for V emanating in the same point $x_\tau=y_\tau$. Then*

$$x_t=y_t$$

for all $t\le\tau$.

Corollary 2.79. *Assume that for given τ, a dynamic upward EVI(K,∞)-gradient flow terminating in x' exists for each x' in a dense subset $D\subset X$. Then this flow can be extended to a flow terminating in any $x'\in X$ and satisfying*

$$d_s(x_s,y_s)\le e^{-K(t-s)}\cdot d_t(x_t,y_t) \tag{108}$$

for any $s<t\le\tau$.

2.10.4 Dynamic Convexity

Let us recall the notion of dynamic convexity as introduced in [59].

Definition 2.80. *We say that the function $V:I\times X\to(-\infty,\infty]$ is* strongly dynamically (K,N)-convex *if for a.e. $t\in I$ and for every d_t-geodesic $(\gamma^a)_{a\in[0,1]}$ with $\gamma^0,\gamma^1\in Dom(V_t)$*

$$\partial_a^+V_t(\gamma_t^{1-})-\partial_a^-V_t(\gamma_t^{0+})\ge-\frac12\partial_t^-d^2_{t-}(\gamma^0,\gamma^1)+\frac{K}{2}d_t^2(\gamma^0,\gamma^1)+\frac1N\left|V_t(\gamma^0)-V_t(\gamma^1)\right|^2. \tag{109}$$

Theorem 2.81. *Assume that for each $t\in I$ and each $x'\in Dom(V_t)$ there exists a dynamic backward EVI(K,N)-gradient flow $(x_s)_{s\in(\sigma,t]}$ for V emanating in x' and such that $\lim_{s\nearrow t}V_s(x_s)=V_t(x_t)$. Then V is strongly dynamically (K,N)-convex.*

To be more precise, we request the inequality (104) at the point t and the inequality (105) at all times before t.

Proof. Fix $t\in I$ and a d_t-geodesic $(\gamma^a)_{a\in[0,1]}$ with $\gamma^0,\gamma^1\in Dom(V_t)$. The a priori assumption of κ-convexity implies $\gamma^a\in Dom(V_t)$ for all $a\in[0,1]$. For each a, let $(\gamma_s^a)_{s\le t}$ denote the EVI$_N$-gradient flow for V emanating in $\gamma^a=\gamma_t^a$. Then for all $a\in(0,\frac12)$

$$\begin{aligned}
V_t(\gamma^a)-V_t(\gamma^0)\ &\le\ \frac12\partial_s^-d^2_{s,t}(\gamma_s^a,\gamma^0)\Big|_{s=t-}\\
&\le\ \frac12\partial_s^-d^2_s(\gamma_s^a,\gamma^0)\Big|_{s=t-}+a^2L\,d_t^2(\gamma^0,\gamma^1)
\end{aligned}$$

(due to the log-Lipschitz continuity of $s \mapsto d_s$) and

$$\begin{aligned} V_t(\gamma^{1-a}) - V_t(\gamma^1) &\leq \frac{1}{2}\partial_s^- d^2_{s,t}(\gamma_s^{1-a},\gamma^1)\Big|_{s=t-} \\ &\leq \frac{1}{2}\partial_s^- d^2_s(\gamma_s^{1-a},\gamma^1)\Big|_{s=t-} + a^2 L\, d_t^2(\gamma^0,\gamma^1). \end{aligned}$$

Moreover, the previous Theorem 2.77 implies

$$\begin{aligned} &\frac{1}{2}\partial_s^- d_s^2(\gamma_s^a,\gamma_s^{1-a})\Big|_{s=t-} - K\, d_t^2(\gamma^a,\gamma^{1-a}) - \frac{1}{N}\Big|V_t(\gamma^a) - V_t(\gamma^{1-a})\Big|^2 \\ &= \liminf_{s\nearrow t} \frac{1}{t-s}\Big[\frac{1}{2}d_t^2(\gamma^a,\gamma^{1-a}) - \frac{1}{2}d_s^2(\gamma_s^a,\gamma_s^{1-a}) - K\, d_t^2(\gamma^a,\gamma^{1-a}) \\ &\qquad - \frac{1}{N}\int_s^t \Big|V_r(\gamma_r^a) - V_r(\gamma_r^{1-a})\Big|^2 dr \geq 0. \end{aligned}$$

(Here we used the requested continuity $V_r(\gamma_r^a) \to V_t(\gamma^a)$ for $r \nearrow t$.)

Adding up these inequalities (the last one multiplied by $\frac{1}{1-2a}$ and the previous ones by $\frac{1}{a}$) yields

$$\begin{aligned} &\frac{1}{a}\Big[V_t(\gamma^a) - V_t(\gamma^0) + V_t(\gamma^{1-a}) - V_t(\gamma^1)\Big] \\ &\leq \liminf_{s\nearrow t} \frac{1}{2(t-s)}\Big(\big[\frac{1}{a}d_t^2(\gamma^0,\gamma^a) + \frac{1}{1-2a}d_t^2(\gamma^a,\gamma^{1-a}) + \frac{1}{a}d_t^2(\gamma^{1-a},\gamma^1)\big] \\ &\qquad - \big[\frac{1}{a}d_s^2(\gamma^0,\gamma_s^a) + \frac{1}{1-2a}d_s^2(\gamma_s^a,\gamma_s^{1-a}) + \frac{1}{a}d_s^2(\gamma_s^{1-a},\gamma^1)\big]\Big) \\ &\qquad + 2aL\, d_t^2(\gamma^0,\gamma^1) - \frac{K}{1-2a}d_t^2(\gamma^a,\gamma^{1-a}) - \frac{1}{N(1-2a)}\Big|V_t(\gamma^a) - V_t(\gamma^{1-a})\Big|^2 \\ &\leq \liminf_{s\nearrow t} \frac{1}{2(t-s)}\Big(d_t^2(\gamma^0,\gamma^1) - d_s^2(\gamma^0,\gamma^1)\Big) \\ &\qquad - \big[(1-2a)K - 2aL\big]\cdot d_t^2(\gamma^0,\gamma^1) - \frac{1}{N(1-2a)}\Big|V_t(\gamma^a) - V_t(\gamma^{1-a})\Big|^2. \end{aligned}$$

In the limit $a \to 0$ this yields the claim. □

3 Improved Gradient Estimates for the Heat Flow and Couplings of Brownian motions

In this chapter we show that the gradient estimates for the heat flow are self-improving. On the level of the dual heat flow this means that if the transport estimate holds with respect to the L^2-Kantorovich distance it also holds with respect to the L^∞-Kantorovich distance. We use this observation for the construction of couplings of Brownian motions and obtain pathwise contraction for their trajectories.

3.1 Main Results

In Chapter 2 it has been shown that a family of metric measure spaces $(X, d_t, m_t)_{t\in I}$ is a super-Ricci flow if and only if the gradient estimate

$$\Gamma_t(P_{t,s}u) \leq P_{t,s}(\Gamma_s(u))$$

holds for every $u \in Dom(\mathcal{E})$ and every $0 < s < t < T$, or, equivalently by duality [36], if the L^2-Kantorovich transport estimate

$$W_{2,s}(\hat{P}_{t,s}\mu, \hat{P}_{t,s}\nu) \leq W_{2,t}(\mu, \nu) \tag{110}$$

holds for every $\mu, \nu \in \mathcal{P}(X)$ and every $0 \leq s < t \leq T$.

In this paper we improve the gradient estimate (and therefore the transport estimate) in the sense of Savaré [55] and Bakry [12] respectively. For this we aggravate our assumption regarding the time-dependence of the metric. We will restrict ourselves to metrics such that the map $t \mapsto \log d_t(x, y)$ is continuously differentiable (instead of Lipschitz continuous as in Chapter 2) and its derivative $h_t(x, y)$ is continuous as $y \to x$, see (117) and (118). We then obtain that $t \mapsto \Gamma_t(u)$ is differentiable and derive a dynamic version of Bochner's inequality

$$\frac{1}{2}\Delta_t(\Gamma_t(u)) - \Gamma_t(u, \Delta_t u) \geq \frac{1}{2}(\partial_t\Gamma_t)(u).$$

In contrast to Chapter 2, where also a dynamic version of Bochner's inequality has been derived, the function u does not need to arise as a heat flow $P_{t,s}u_s$.

Theorem 3.1. *Let $(X, d_t, m_t)_{t\in I}$ be a one-parameter family of geodesic Polish metric measure spaces satisfying* (113), (114), (117) *and* (118) *such that each (X, d_t, m_t) is a RCD(K, N) space. If the transport estimate* (110) *holds, then the dynamic Bochner inequality* (120) *holds at all $t \in I$.*

Having established a dynamic version of Bochner's inequality we can follow the arguments in [55] and obtain the improved gradient estimate.

Theorem 3.2. *Let $(X, d_t, m_t)_{t\in I}$ be as in Theorem 3.1. Then, if the dynamic Bochner inequality* (120) *and the regularity assumption* (132) *is satisfied, for every $\alpha \in [1/2, 1]$ we have for a.e. $\tau \leq t$ and $\sigma \geq s$*

$$\Gamma_\tau(P_{\tau,\sigma}u)^\alpha \leq P_{\tau,\sigma}(\Gamma_\sigma(u)^\alpha), \tag{111}$$

for every $u \in Dom(\mathcal{E})$ and m-a.e. $x \in X$.

As a consequence we obtain that if (X, d_t, m_t) is a super-Ricci flow in the sense of Chapter 2, then for $\beta \in [1, 2]$

$$|\nabla_t P_{t,s} u|_*^\beta \leq P_{t,s}(|\nabla_s u|_*^\beta)$$

and for $p \in [1, \infty]$

$$W_{p,s}(\hat{P}_{t,s}\mu, \hat{P}_{t,s}\nu) \leq W_{p,t}(\mu, \nu), \tag{112}$$

see Corollary 3.17. Similar as in [58] we will apply these results to Brownian motions and construct a coupling (X_s^1, X_s^2) of Brownian motions such that almost surely

$$d_s(X_s^1, X_s^2) \leq d_t(x, y),$$

see Theorem 3.25.

Example. *A possible example for the setting chosen in this paper is the super-Ricci flow on the spherical cone over the product of the 2-spheres with radius $1/\sqrt{3}$ constructed in Chapter 2. This space is a RCD$^*(4,5)$-space, and the punctured cone is a 5-dimensional (non-complete) Riemannian manifold with constant curvature 4. A possible Ricci flow on the punctured cone is given by distances which shrink to one point homothetically in time. The completion of this flow is a super-Ricci flow which shrinks to a point homothetically in time. Hence, for time points smaller than the collapsing time the metrics satisfy the assumptions* (117) *and* (118)*. The same argumentation can be used to obtain* (113) *for the measures.*

3.2 Proof of the Main Results

In the sequel let $(X, d_t, m_t)_{t \in I}$, where $I = (0, T)$, be a one-parameter family of geodesic Polish metric measure spaces such that the following holds:

1. There exists a finite reference measure m with full topological support such that $m_t = e^{-f_t}$ with Borel functions (f_t) satisfying

$$|f_t(x)| \leq C, \quad |f_t(x) - f_t(y)| \leq C d_t(x, y), \quad |f_t(x) - f_s(x)| \leq L|t - s|, \tag{113}$$

   with constants $C, L > 0$ independent of $x, y \in X$ and $s, t \in I$.

2. the distance is "log-Lipschitz" continuous, i.e.

$$|\log(d_t(x, y)/d_s(x, y))| \leq L|t - s| \tag{114}$$

   for all $x, y \in X$ and all $s, t \in I$,

3. there exist constants $K, N \in \mathbb{R}$ such that for each $t \in I$ the space (X, d_t, m_t) satisfies the Riemannian curvature-dimension bound RCD(K, N) in the sense of [7], [24].

In the sequel let us introduce the time-dependent quantities which we are going to use.

Let $\mathcal{P}(X)$ denote the space of all Borel probability measures. We set for each $p \in [1, \infty)$

$$W_{p,t}(\mu_1, \mu_2) = \min \left\{ \int_{X \times X} d_t^p(x, y)\, d\gamma(x, y) | \gamma \in \Pi(\mu_1, \mu_2) \right\}^{1/p},$$

where $\Pi(\mu_1, \mu_2)$ is the space of all measures in $\mathcal{P}(X \times X)$ whose marginals $(e_i)_\# \mu$ coincide with μ_i. We also set

$$W_{\infty,t}(\mu_1, \mu_2) = \inf \left\{ ||d_t||_{L^\infty(\gamma)} | \gamma \in \Pi(\mu_1, \mu_2) \right\} = \lim_{p \to \infty} W_{p,t}(\mu_1, \mu_2),$$

with essential supremum $||d||_{L^\infty(\gamma)} = \inf\{C \geq 0 | d(x, y) \leq C\, \gamma\text{-a.e. } x, y\}$. For the second equality see e.g. Lemma 3.2 in [36].

We recall that the Cheeger energy Ch_t at time $t \in I$ is defined as the convex and lower-semicontinuous functional in $L^2(X, m_t)$

$$\mathrm{Ch}_t(u) := \inf \left\{ \liminf_{n \to \infty} \frac{1}{2} \int_X \mathrm{lip}_t(u_n)^2\, dm_t \right\}$$

where the infimum is taken over all bounded Lipschitz functions $u_n \in \mathrm{Lip}_b(X)$ such that $u_n \to u$ in $L^2(X, m_t)$ (cf. [5, 59]). Here, $\mathrm{lip}_t u$ denotes the local Lipschitz constant w.r.t. the metric d_t

$$\mathrm{lip}_t u(x) := \limsup_{y \to x} \frac{|u(y) - u(x)|}{d_t(x, y)},$$

and Ch_t admits the local representation formula

$$\mathrm{Ch}_t(u) = \frac{1}{2} \int_X |\nabla_t u|_*^2\, dm_t,$$

where $|\nabla_t u|_*$ is the minimal relaxed gradient [5]. Since (X, d_t, m_t) satisfies a Riemannian curvature bound, (in particular Ch_t is quadratic) $\mathcal{E}_t := 2\mathrm{Ch}_t$ is a strongly local Dirichlet form with Carré du Champ

$$\Gamma_t(u) = |\nabla_t u|_*^2$$

cf. [55, 7, 6], i.e.

$$\mathcal{E}_t(u) = \int_X \Gamma_t(u)\, dm_t. \tag{115}$$

Thanks to (115), $\mathcal{E}(u, v) = \int_X \Gamma_t(u, v)\, dm_t$ where

$$\Gamma_t(u, v) := \frac{1}{4}(\Gamma_t(u + v) - \Gamma_t(u - v)).$$

$\Gamma(\cdot, \cdot)$ satisfies the chain rule and the Leibniz rule

$$\Gamma_t(\theta(u), v) = \theta'(u)\Gamma_t(u, v), \quad \Gamma_t(uv, w) = u\Gamma_t(v, w) + v\Gamma_t(u, w),$$

where $u, v, w \in \mathit{Dom}(\mathcal{E}_t)$ and $\theta \in \mathrm{Lip}(\mathbb{R})$, $\theta(0) = 0$. We call the linear generator Δ_t the Laplacian and

$$-\int_X \Delta_t u\, v\, dm_t = \mathcal{E}_t(u, v) \qquad \forall u \in \mathit{Dom}(\Delta_t), v \in \mathit{Dom}(\mathcal{E}_t),$$

with domain $Dom(\Delta_t) \subset Dom(\mathcal{E}_t)$.

Due to our assumptions (113) and (114), the sets $L^2(X, m_t)$ and $W^{1,2}(X, d_t, m_t) := \mathcal{D}(\mathcal{E}_t)$ do not depend on t and the respective norms for varying t are equivalent to each other. We put $\mathcal{H} = L^2(X, m)$ and $\mathcal{F} = Dom(\mathcal{E}_{t_0})$ for some fixed t_0 as well as

$$\mathcal{F}_{(s,\tau)} = L^2\big((s,\tau) \to \mathcal{F}\big) \cap H^1\big((s,\tau) \to \mathcal{F}^*\big) \subset \mathcal{C}\big([s,\tau] \to \mathcal{H}\big)$$

for each $0 \le s < \tau \le T$.

The Heat Equations

A function u is called *solution to the heat equation*

$$\Delta_t u = \partial_t u \qquad \text{on } (s,\tau) \times X$$

if $u \in \mathcal{F}_{(s,\tau)}$ and if for all $w \in \mathcal{F}_{(s,\tau)}$

$$-\int_s^\tau \mathcal{E}_t(u_t, w_t)dt = \int_s^\tau \langle \partial_t u_t, w_t e^{-f_t} \rangle_{\mathcal{F}^*,\mathcal{F}}\, dt$$

where $\langle \cdot, \cdot \rangle_{\mathcal{F}^*,\mathcal{F}} = \langle \cdot, \cdot \rangle$ denotes the dual pairing. Note that thanks to (113), $w \in L^2\big((s,\tau) \to \mathcal{F}\big)$ if and only if $we^{-f} \in L^2\big((s,\tau) \to \mathcal{F}\big)$.

Further a function v is called *solution to the adjoint heat equation*

$$-\Delta_s v + \partial_s f \cdot v = \partial_s v \qquad \text{on } (\sigma, t) \times X$$

if $v \in \mathcal{F}_{(\sigma,t)}$ and if for all $w \in \mathcal{F}_{(\sigma,t)}$

$$\int_\sigma^t \mathcal{E}_s(v_s, w_s)ds + \int_\sigma^t \int_X v_s \cdot w_s \cdot \partial_s f_s \, dm_s\, ds = \int_\sigma^t \langle \partial_s v_s, w_s e^{-f_s} \rangle_{\mathcal{F}^*,\mathcal{F}}\, ds.$$

We recall the following results from Chapter 2.

Theorem 3.3. *(i) For each $0 \le s < \tau \le T$ and each $h \in \mathcal{H}$ there exists a unique solution $u \in \mathcal{F}_{(s,\tau)}$ to the heat equation $\partial_t u_t = \Delta_t u_t$ on $(s,\tau) \times X$ with $u_s = h$.*

(ii) The heat propagator $P_{t,s} : h \mapsto u_t$ admits a kernel $p_{t,s}(x,y)$ w.r.t. m_s, i.e.

$$P_{t,s}h(x) = \int p_{t,s}(x,y)h(y)\, dm_s(y).$$

If X is bounded, for each $(s', y) \in (s,T) \times X$ the function $(t,x) \mapsto p_{t,s}(x,y)$ is a solution to the heat equation on $(s',T) \times X$.

(iii) All solutions $u : (t,x) \mapsto u_t(x)$ to the heat equation on $(s,\tau) \times X$ are Hölder continuous in t and x. All nonnegative solutions satisfy a scale invariant parabolic Harnack inequality of Moser type.

(iv) The heat kernel $p_{t,s}(x,y)$ is Hölder continuous in all variables, it is Markovian

$$\int p_{t,s}(x,dy) := \int p_{t,s}(x,y)\, dm_s(y) = 1 \qquad (\forall s < t, \forall x)$$

and has the propagator property

$$p_{t,r}(x,z) = \int p_{t,s}(x,y)\, p_{s,r}(y,z)\, dm_s(y) \qquad (\forall r < s < t, \forall s, z).$$

Theorem 3.4. *(i) For each $0 \leq \sigma < t \leq T$ and each $g \in \mathcal{H}$ there exists a unique solution $v \in \mathcal{F}_{(0,t)}$ to the adjoint heat equation $\partial_s v_s = -\Delta_s v_s + (\partial_s f_s) v_s$ on $(\sigma, t) \times X$ with $v_t = g$.*

(ii) This solution is given as $v_s(y) = P^*_{t,s} g(y)$ in term of the adjoint heat propagator

$$P^*_{t,s} g(y) = \int p_{t,s}(x,y) g(x)\, dm_t(x). \tag{116}$$

If X is bounded, for each $(t', x) \in (0,t) \times X$ the function $(s,y) \mapsto p_{t,s}(x,y)$ is a solution to the adjoint heat equation on $(0,t') \times X$.

(iii) All solutions $v : (s,y) \mapsto v_s(y)$ to the adjoint heat equation on $(\sigma,t) \times X$ are Hölder continuous in s and y. All nonnegative solutions satisfy a scale invariant parabolic Harnack inequality of Moser type.

By duality, the propagator $(P_{t,s})_{s \leq t}$ acting on bounded continuous functions induces a *dual propagator* $(\hat{P}_{t,s})_{s \leq t}$ acting on probability measures as follows

$$\int u\, d(\hat{P}_{t,s}\mu) = \int (P_{t,s} u) d\mu \qquad \forall u \in \mathcal{C}_b(X), \forall \mu \in \mathcal{P}(X).$$

The time-dependent function $v_t(x) = P_{t,s} u(x)$ is a solution to the heat equation, whereas the time-dependent measure $\nu_s(dy) = \hat{P}_{t,s}\mu(dy)$ is a solution to the *dual heat equation*

$$-\partial_s \nu = \hat{\Delta}_s \nu.$$

Again $\hat{\Delta}_s$ is defined by duality: $\int u\, d(\hat{\Delta}_s \mu) = \int \Delta_s u\, d\mu \quad \forall u, \forall \mu$.

We recall Theorem 2.20 and Lemma 2.34 from Chapter 2.

Lemma 3.5. *Let $u, g \in \mathcal{F}$ and $t \in I$ with $g m_t \in \mathcal{P}(X)$. Then,*

$$\lim_{h \searrow 0} \frac{1}{h} \left(\int u g dm_t - \int u P^*_{t,t-h} g dm_{t-h} \right) = \int \Gamma_t(u,g) dm_t.$$

Theorem 3.6. *For all $0 < s < \tau < T$ and for all solutions $u \in \mathcal{F}_{(s,T)}$ to the heat equation*

(i) $u_t \in Dom(\Delta_t)$ for a.e. $t \in (s,\tau)$.

(ii) If the initial condition $u_s \in \mathcal{F}$ then

$$u \in L^2\big((s,\tau) \to Dom(A_\cdot)\big) \cap H^1\big((s,\tau) \to \mathcal{H}\big).$$

More precisely,

$$e^{-3L\tau} \mathcal{E}_\tau(u_\tau) + 2 \int_s^\tau e^{-3Lt} \int_X \left|\Delta_t u_t\right|^2 dm_t\, dt \leq e^{-3Ls} \cdot \mathcal{E}_s(u_s).$$

(iii) For all solutions v to the adjoint heat equation on $(\sigma,t) \times X$ and all $s \in (\sigma,t)$

$$\mathcal{E}_s(v_s) + \|v_s\|^2_{L^2(m_s)} \leq e^{3L(t-s)} \cdot \Big[\mathcal{E}_t(v_t) + \|v_t\|^2_{L^2(m_t)} \Big].$$

Moreover, $v_s \in Dom(\Delta_s)$ for a.e. $s \in (\sigma,t)$.

3.2.1 From Transport Estimates to Bochner's Inequality

In this section we aggravate the regularity of the map $r \mapsto \log d_r(x,y)$. We assume that there exists a $\mathcal{C}^0$ map $r \mapsto h_r(x,y)$, uniformly bounded $|h_r(x,y)| \leq C$ such that for each $s,t \in I$ and $x,y \in X$

$$d_t(x,y) = d_s(x,y)e^{\int_s^t h_r(x,y)\,dr}. \tag{117}$$

Consequently, for each $x,y \in X$, $r \mapsto \log d_r(x,y)$ is continuously differentiable with derivative $h_r(x,y) = \frac{d}{dr}\log d_r(x,y)$.
Moreover we assume that

$$\begin{aligned}&\forall x \in X, r \in I \text{ the limit } \lim_{y\to x} h_r(x,y) := H_r(x) \text{ exists, measurable in } x,\\ &\text{and } r \mapsto H_r(x) \text{ is continuous } \forall x \in X.\end{aligned} \tag{118}$$

We obtain the following lemma.

Lemma 3.7. *Let $u \in \mathrm{Lip}(X)$. Then for all $s,t \in I$ and $x \in X$*

$$\mathrm{lip}_t u(x) = \mathrm{lip}_s u(x) e^{-\int_s^t H_r(x)\,dr}.$$

Proof. For $s < t$, we obtain from the very definition of the local slope

$$\begin{aligned}\mathrm{lip}_t u(x) =& \limsup_{y\to x} \frac{|u(y)-u(x)|}{d_t(x,y)} \leq \limsup_{y\to x} \frac{|u(y)-u(x)|}{d_s(x,y)} e^{-\liminf_{y\to x}\int_s^t h_r(x,y)\,dr}\\ =& \mathrm{lip}_s u(x) e^{-\int_s^t H_r(x)\,dr},\end{aligned}$$

where we applied dominated convergence. Changing the roles of s and t yields

$$\begin{aligned}\mathrm{lip}_s u(x) =& \limsup_{y\to x} \frac{|u(y)-u(x)|}{d_s(x,y)} \leq \limsup_{y\to x} \frac{|u(y)-u(x)|}{d_t(x,y)} e^{-\liminf_{y\to x}\int_s^t h_r(x,y)\,dr}\\ =& \mathrm{lip}_t u(x) e^{-\int_s^t H_r(x)\,dr},\end{aligned}$$

which proves the assertion. □

We apply our observation to the minimal relaxed gradient. We say that $G \in L^2(X,m_t)$ is a *t-relaxed gradient* of $u \in L^2(X,m_t)$ if there exists Lipschitz functions $u_n \in L^2(X,m_t)$ such that

$$u_n \to u \text{ in } L^2(X,m_t) \text{ and } \mathrm{lip}_t u_n \rightharpoonup \tilde{G} \text{ in } L^2(X,m_t),\ \tilde{G} \leq G\, m\text{-a.e. in } X.$$

G is the *minimal t-relaxed gradient* $|\nabla_t u|_*$ if its $L^2(X,m_t)$ norm is minimal among all relaxed gradients, see [5, Definition 4.2]. The collection of all t-relaxed gradients is convex and closed in $L^2(X,m_t)$ [5, Lemma 4.3].

Proposition 3.8. *For m-a.e. $x \in X$*

$$|\nabla_t u|_*(x) = |\nabla_s u|_*(x) e^{-\int_s^t H_r(x)\,dr}$$

for each $u \in \mathcal{F}$ and for all $s,t \in I$. In particular for m-a.e. $x \in X$, $t \mapsto |\nabla_t u|_*(x)$ is continuously differentiable.

Proof. Assume $s \leq t$. Let $u_n \in L^2(X, m_s)$ be a sequence of Borel Lipschitz functions such that $u_n \to u$ and $\text{lip}_s u_n \to |\nabla_s u|_*$ in $L^2(X, m_s)$, see Lemma 4.3 in [5]. Then since H is uniformly bounded

$$\text{lip}_s u_n(\cdot) e^{-\int_s^t H_r(\cdot)\, dr} \to |\nabla_s u|_*(\cdot) e^{-\int_s^t H_r(\cdot)\, dr} \text{ in } L^2(X, m_s).$$

This implies that $|\nabla_s u|_*(\cdot) e^{-\int_s^t H_r(\cdot)\, dr}$ is a relaxed gradient of u with respect to the d_t norm, and hence from Lemma 4.4 in [5]

$$|\nabla_t u|_*(\cdot) \leq |\nabla_s u|_*(\cdot) e^{-\int_s^t H_r(\cdot)\, dr} \quad m\text{- a.e. in } X.$$

Changing the roles of s and t yields that

$$|\nabla_t u|_*(\cdot) = |\nabla_s u|_*(\cdot) e^{-\int_s^t H_r(\cdot)\, dr} \quad m\text{- a.e. in } X.$$

Choosing s and t from a dense and countable set D in I the argument from above implies that m-a.e. in X

$$|\nabla_t u|_*(\cdot) = |\nabla_s u|_*(\cdot) e^{-\int_s^t H_r(\cdot)\, dr} \tag{119}$$

for each s and t in D. Since the dependence of the left and the right side of the equality is continuous with respect to s and t, we conclude that for m-a.e. $x \in X$, $|\nabla_t u|_*(\cdot) = |\nabla_s u|_*(\cdot) e^{-\int_s^t H_r(\cdot)\, dr}$ holds for every s and t in I.

Similarly, we choose u in a dense and countable set C in $\mathcal{F}$ ([6, Proposition 4.10]) and obtain that m-a.e. equation (119) holds for every $s, t \in I$ and every $u \in C$. Given $u \in \mathcal{F}$ we approximate u by a sequence $u_n \in C$, i.e. $|\nabla_t u_n| \to |\nabla_t u|$ in $L^2(X, m_t)$. Then there exists a subsequence u_{n_k} such that for m-a.e. $x \in X$, $|\nabla_t u_{n_k}|(x) \to |\nabla_t u|(x)$. Equality (119) implies that for the same subsequence $|\nabla_s u_{n_k}|(x) \to |\nabla_s u|(x)$ for m-a.e. x. Hence we showed that for m-a.e. $x \in X$, (119) holds for every $u \in \mathcal{F}$ and every $s, t \in I$.

The last assertion follows directly from the fact that $r \mapsto H_r(x)$ is supposed to be continuous for all $x \in X$. ☐

We give a refined weak dynamic version of Bochner's inequality, cf. Chapter 2.

Definition 3.9. *We say that the dynamic Bochner inequality holds at time t if for all $u \in Dom(\Delta_t) \cap L^\infty(X, m_t)$ such that $\Gamma_t(u) \in L^\infty(X, m_t)$, and all $g \in Dom(\Delta_t) \cap L^\infty(X, m_t)$ with $g \geq 0$*

$$\frac{1}{2} \int \Gamma_t(u) \Delta_t g \, dm_t + \int (\Delta_t u)^2 g + \Gamma_t(u, g) \Delta_t u \, dm_t \geq \frac{1}{2} \int (\partial_t \Gamma_t)(u) g \, dm_t. \tag{120}$$

This is a "real" Bochner inequality in the sense that on the one hand u and g do not have to arise as a heat flow (see Definition 2.48), and on the other we employ the time-derivative $\partial_t \Gamma_t(u)$ in contrast to the definition in Chapter 2. For the proof of Theorem 3.1 we use the same starting point as in the proof of Theorem 2.51. This argumentation is inspired by [18], where the authors prove the equivalence between Wasserstein contraction estimates and Bochner's inequality in the static setting.

Proof of Theorem 3.1. Define $u = h^t_\varepsilon u_0$, where $u_0 \in L^\infty(X, m_t) \cap L^2(X, m_t)$ and h^t_ε the static semigroup mollification

$$h^t_\varepsilon u_0 := -\frac{1}{\varepsilon^2}\int_0^\infty H^t_r u_0 \kappa(\frac{r}{\varepsilon})\, dr.$$

Here, $(H^t_r)_{r\geq 0}$ denotes the (static) semigroup associated to $\mathcal{E}_t$ and $\kappa \in \mathcal{C}^\infty_c((0,\infty))$ with $\kappa \geq 0$ and $\int_0^\infty \kappa_r\, dr = 1$. Recall that $u, \Delta_t u \in Dom(\Delta_t) \cap \mathrm{Lip}_b(X)$.

Let $g \in \mathcal{F} \cap L^\infty(X, m_t)$ such that $g \geq 0$. Then, the transport estimate (110) together with Lemma 2.52 and Lemma 2.53 in Chapter 2 eventually yields

$$-\frac{1}{2}\int P_{t,s}(\Gamma_s(u))g dm_t + \int \Gamma_t(P_{t,s}u, u) g dm_t \leq \frac{1}{2}\int \Gamma_t(u) g dm_t.$$

We subtract $\frac{1}{2}\int \Gamma_t(u) g dm_t$ on each side and divide by $t-s$ obtaining

$$\begin{aligned}&\frac{1}{2(t-s)}\left[\int \Gamma_t(u) g dm_\tau - \int P_{t,s}(\Gamma_s(u)) g dm_t\right]\\ &+\frac{1}{t-s}\left[\int \Gamma_t(P_{t,s}u, u) g dm_t - \int \Gamma_t(u) g dm_t\right]\\ &\leq 0.\end{aligned} \tag{121}$$

We decompose the first term on the left-hand side into the following two terms

$$\begin{aligned}&\frac{1}{2(t-s)}\left[\int \Gamma_t(u) g dm_t - \int \Gamma_s(u) P^*_{t,s} g dm_s\right]\\ =&\frac{1}{2(t-s)}\left[\int \Gamma_t(u) g dm_t - \int \Gamma_t(u) P^*_{t,s} g dm_s\right] + \frac{1}{2}\int \frac{\Gamma_t(u) - \Gamma_s(u)}{t-s} P^*_{t,s} g dm_s.\end{aligned}$$

Recall that $\Gamma_t(u) \in \mathcal{F}$ [55, Lemma 3.2] and thus we can apply Lemma 3.5, which gives us

$$\lim_{s\nearrow t}\frac{1}{(t-s)}\left[\int \Gamma_t(u) g dm_t - \int \Gamma_t(u) P^*_{t,s} g dm_s\right] = \int \Gamma_t(\Gamma_t(u), g) dm_t, \tag{122}$$

while, since $|\frac{\Gamma_s(u)-\Gamma_t(u)}{(t-s)}| \leq 2L\Gamma_t(u) \in L^\infty(X, m_t)$,

$$\begin{aligned}&\liminf_{s\nearrow t}\int \frac{\Gamma_t(u) - \Gamma_s(u)}{(t-s)}(P^*_{t,s} g) dm_s\\ &\geq \liminf_{s\nearrow t}\int \frac{\Gamma_t(u) - \Gamma_s(u)}{(t-s)} g dm_t + \liminf_{s\nearrow t}\int \frac{\Gamma_t(u) - \Gamma_s(u)}{(t-s)}(P^*_{t,s} g e^{-f_s} - g e^{-f_t}) dm\\ &\geq \int (\partial_t \Gamma_t)(u) g dm_t - \limsup_{s\nearrow t} 2L||\Gamma_t(u)||_{L^\infty(X,m_t)}||P^*_{t,s} g e^{-f_s} - g e^{-f_t}||_{L^1(X,m_t)}\\ &= \int (\partial_t \Gamma_t)(u) g dm_t,\end{aligned} \tag{123}$$

where we used Proposition 3.8 in the last inequality and that $P^*_{t,s} g e^{-f_s} \to g e^{-f_t}$ in $L^1(X, m)$ as $s \to t$.

Regarding the second term on the left-hand side of (121), note that the Leibniz rule and the integration by parts formula is applicable and we get

$$\begin{aligned}\int \Gamma_t(P_{t,s}u,u)gdm_t &= \int \Gamma_t(gP_{t,s}u,u)dm_t - \int \Gamma_t(g,u)P_{t,s}udm_t \\ &= -\int \psi P^*_{t,s}(g\Delta_t u)dm_s - \int P^*_{t,s}(\Gamma_t(g,u))udm_s.\end{aligned} \tag{124}$$

Subtracting $\int \Gamma_t(u)gdm_t$ and applying (124)

$$\begin{aligned}&\frac{1}{(t-s)}(\int \Gamma_t(P_{t,s}u,u)gdm_t - \int \Gamma_t(u)gdm_t) \\ =&\frac{1}{(t-s)}(-\int \psi P^*_{t,s}(g\Delta_t u)dm_s + \int \psi(g\Delta_t u)dm_t) \\ +&\frac{1}{(t-s)}(-\int P^*_{t,s}(\Gamma_t(g,u))udm_s + \int \Gamma_t(u,g)udm_t).\end{aligned}$$

Letting $s \nearrow t$ we have since $g \in \mathcal{F} \cap L^\infty(X,m_t)$ and $\Delta_t u \in \mathrm{Lip}_b(X)$, $g\Delta_t u \in \mathcal{F} \cap L^1(X,m_t)$

$$\lim_{s\nearrow t} \frac{1}{(t-s)}(-\int uP^*_{t,s}(g\Delta_t u)dm_s + \int u(g\Delta_t u)dm_t) = \int \Gamma_t(u,g\Delta_t u)dm_t$$

by virtue of Lemma 3.5. In order to determine

$$\lim_{s\nearrow t} \frac{1}{(t-s)}(-\int P^*_{t,s}(\Gamma_t(g,u))udm_s + \int \Gamma_t(u,g)udm_t),$$

we need to argue whether $\Gamma_t(g,u) \in \mathcal{F}$. But this is the case, since, due to our static RCD(K,∞) assumption, we may apply Theorem 3.4 in [55] and obtain

$$\Gamma_t(\Gamma_t(g,u)) \leq 2(\gamma_2(u) - K\Gamma_t(u))\Gamma_t(g) + 2(\gamma_2(g) - K\Gamma_t(g))\Gamma_\tau(u) \quad m_t\text{-a.e.},$$

where $\gamma_2(u), \gamma_2(g) \in L^1(X,m_t)$. Our regularity assumptions on u and g provide that the right hand side is in $L^1(X,m_t)$ and consequently Lemma 3.5 implies

$$\lim_{s\nearrow t} \frac{1}{(t-s)}(-\int P^*_{t,s}(\Gamma_t(g,u))udm_s + \int \Gamma_t(u,g)udm_t) = \int \Gamma_t(\Gamma_t(g,u),u)dm_t.$$

Combining these observations we find

$$\begin{aligned}&\lim_{s\nearrow t} \frac{1}{(t-s)}(\int \Gamma_t(P_{t,s}u,u)gdm_t - \int \Gamma_t(u)gdm_t) \\ =& \int \Gamma_t(u,g\Delta_t u)dm_t + \int \Gamma_t(\Gamma_t(g,u),u)dm_t = -\int (\Delta_t u)^2 g + \Gamma_t(g,u)\Delta_t udm_t.\end{aligned} \tag{125}$$

Hence from (121), (122), (123) and (125)

$$\frac{1}{2}\int (\partial_t \Gamma_t)(u)gdm_t + \frac{1}{2}\int \Gamma_t(\Gamma_t(u),g)dm_t \leq \int (\Delta_t u)^2 g + \Gamma_t(g,u)\Delta_t udm_t.$$

Let now $g \in Dom(\Delta_t) \cap L^\infty(X,m)$ with $g \geq 0$ and $u \in Dom(\Delta_t) \cap L^\infty(X,m_t)$ with $\Gamma_t u \in L^\infty(X,m_t)$. Then from the above argumentation we obtain

$$\begin{aligned}&\frac{1}{2}\int \Gamma_t(h^t_\varepsilon u)\Delta_t g\, dm_t + \int (\Delta_t(h^t_\varepsilon u))^2 g + \Gamma_t(g,h^t_\varepsilon u)\Delta_t(h^t_\varepsilon u)dm_t \\ &\geq \frac{1}{2}\int (\partial_t \Gamma_t)(h^t_\varepsilon u_n)gdm_t.\end{aligned}$$

Since $(\partial_t \Gamma_t)(u)(x) = -2H_t(x)\Gamma_t(u)(x)$ for m-a.e. $x \in X$ and $|H_t(x)| \le C$, we obtain the assertion by letting $\varepsilon \to 0$ with taking into account that

$$||h^t_\varepsilon u - u||_{\mathcal{F}} \to 0 \text{ as } \varepsilon \to 0 \text{ and } \Delta_t h^t_\varepsilon u = h^t_\varepsilon \Delta_t u.$$

□

3.2.2 Self-improvement of the Gradient Estimate

Quasi-regular Dirichlet Forms

We follow the approach in [55] and briefly recall the notion of quasi-regular Dirichlet forms developed in [44] and [21]. We denote by $\mathcal{F} = \{u \in L^2(X,m) | \mathcal{E}(u) < \infty\}$ the domain of a Dirichlet form $\mathcal{E}\colon L^2(X,m) \to [0,\infty]$, where X is a Polish space and m is a σ-finite Borel measure. $\mathcal{F}$ is a Hilbert space with norm $||u||^2_{\mathcal{F}} = ||u||^2_{L^2(X,m)} + \mathcal{E}(u)$. If F is a closed set in X we denote

$$\mathcal{F}_F := \{u \in \mathcal{F} | u(x) = 0 \text{ for } m\text{-a.e. } x \in X \setminus F\}.$$

Definition 3.10. *Given a Dirichlet form $\mathcal{E}$ on a Polish space X, an $\mathcal{E}$-nest is an increasing sequence of closed subsets $(F_k)_{k\in\mathbb{N}} \subset X$ such that $\cup_{k\in\mathbb{N}}\mathcal{F}_{F_k}$ is dense in $\mathcal{F}$.*
A set $N \subset X$ is $\mathcal{E}$-polar if there is an $\mathcal{E}$-nest $(F_k)_{k\in\mathbb{N}}$ such that $N \subset X \setminus \cup_{k\in\mathbb{N}} F_k$.
If a property holds in a complement of an $\mathcal{E}$-polar set we say that it holds $\mathcal{E}$-quasi-everywhere ($\mathcal{E}$-q.e.).
A function $u\colon X \to \mathbb{R}$ is said to be $\mathcal{E}$-quasi-continuous if there exists an $\mathcal{E}$-nest $(F_k)_{k\in\mathbb{N}}$ such that every restriction $f_{|F_k}$ is continuous on F_k.

The Dirichlet form $\mathcal{E}$ is said to be quasi-regular if the following three properties hold.

1. *There exists an $\mathcal{E}$-nest $(F_k)_{k\in\mathbb{N}}$ consisting of compact sets.*
2. *There exists a dense subset of $\mathcal{F}$ whose elements have $\mathcal{E}$-quasi-continuous representatives.*
3. *There exists an $\mathcal{E}$-polar set $N \subset X$ and a countable collection of $\mathcal{E}$-quasi-continuous functions $(f_k)_{k\in\mathbb{N}} \subset \mathcal{F}$ separating the points of $X \setminus N$.*

For every $u \in \mathcal{F}$ the quasi-regularity implies that u admits an $\mathcal{E}$-quasi-continuous representative $\tilde{u}$. The representative is unique $q.e.$ and

$$\text{if } u \in \mathcal{F} \text{ with } |u| \le C\, m\text{-a.e., then } |\tilde{u}| \le C \text{ q.e..} \tag{126}$$

The following Lemma is taken from [55, Lemma 2.6].

Lemma 3.11. *Let $\mathcal{E}$ be a strongly local, quasi-regular Dirichlet form with linear generator Δ. Let $\psi \in L^1(X,m) \cap L^\infty(X,m)$ nonnegative and $\varphi \in L^1(X,m) \cap L^2(X,m)$ such that*

$$\int_X \psi \Delta g \, dm \ge -\int_X \varphi g \, dm$$

for any nonnegative $g \in \mathcal{F} \cap L^\infty(X, m)$ *with* $\Delta g \in L^\infty(X, m)$. *Then* $\psi \in \mathcal{F}$ *with*

$$\mathcal{E}(\psi) \leq \int_X \psi\varphi \, dm, \qquad \int \varphi \, dm \geq 0,$$

and there exists a unique finite Borel measure $\mu := \mu_+ - \varphi m$ *with* $\mu_+ \geq 0$, $\mu_+(X) \leq \int \varphi \, dm$ *such that every* $\mathcal{E}$*-polar set is* $|\mu|$*-negligible, the q.c. representative of any function in* $\mathcal{F}$ *belongs to* $L^1(X, |\mu|)$ *and*

$$-\mathcal{E}(\psi, g) = -\int \Gamma(\psi, g) \, dm = \int \tilde{g} \, d\mu \text{ for every } g \in \mathcal{F}.$$

We denote by $\Delta^* u$ the measured valued Laplacian, i.e. the signed measure $\mu = \mu_+ - \mu_-$ such that

$$\mathcal{E}(u, \varphi) = \int \tilde{\varphi} \, d\mu \text{ for every } \varphi \in \mathcal{F}. \tag{127}$$

Contraction Estimates for the Heat Flows $P_{t,s}$ and $\hat{P}_{t,s}$

For each $t \in I$ we define the Hessian

$$H_t[u](g, h) := \frac{1}{2}\Big(\Gamma_t(g, \Gamma_t(u, h)) + \Gamma_t(h, \Gamma_t(u, g)) - \Gamma_t(u, \Gamma_t(g, h))\Big).$$

Recall that on a family of closed Riemannian manifolds (M, g_t) we obtain the equality

$$H_t[u](g, h) = \langle \nabla_t^2 u \nabla_t g, \nabla_t h \rangle_{g_t}.$$

Further note that $|\langle \nabla_t^2 u \nabla_t g, \nabla_t h \rangle_{g_t}| \leq |\nabla_t^2 u|_{HS} |\nabla_t g| |\nabla_t h|$, where $|\cdot|_{HS}$ denotes the Hilbert-Schmidt norm. If the manifold has Ricci curvature bounded from below by some $K \in \mathbb{R}$ then with $||\cdot||_2 = ||\cdot||_{L^2}$ and $K_- = \max\{-K, 0\}$

$$|||\nabla_t^2 u|_{HS}||_2^2 \leq (1 + K_-/2)(||\Delta_t u||_2^2 + ||u||_2^2).$$

We define the distribution valued Γ_2-operator

$$\Gamma_{2,t}(u) \colon \mathcal{F} \cap L^\infty \cap L^1 \to \mathbb{R}$$

as in Chapter 2.

Definition 3.12. *For each* $u \in Dom(\Delta_t)$ *such that* $u, \Gamma_t(u) \in L^\infty(X, m_t)$ *we define*

$$\Gamma_{2,t}(u)(g) = \int -\frac{1}{2}\Gamma_t(\Gamma_t(u), g) \, dm_t + \int (g(\Delta_t u)^2 + \Gamma_t(g, u)\Delta_t u) \, dm_t,$$

where $g \in \mathcal{F}$ *such that* $g \in L^1(X, m_t) \cap L^\infty(X, m_t)$.

Note that thanks to the static $\mathrm{RCD}(K, N)$-condition the domain of the Laplacian coincides with the domain of the Hessian, i.e. $Dom(\Delta_t) = W^{2,2}(X, d_t, m_t)$, and

$$|\Gamma_{2,t}(u)(g)| \leq ||g||_\infty ||\Delta_t u||_2^2 + C||\sqrt{\Gamma_t(u)}||_\infty ||\sqrt{\mathcal{E}_t(g)}(||\Delta_t u||_2 + ||u||_2),$$

cf. section 2.8 in Chapter 2. Moreover, each $\mathcal{E}_t = 2\mathrm{Ch}_t$ defines a quasi-regular Dirichlet form ([55, Theorem 4.1]).

Proposition 3.13. *Suppose that Bochner's inequality holds at time $t \in I$. Then for every $u \in Dom(\Delta_t)$ with $u, \Gamma_t(u) \in L^\infty(X, m_t)$*

1. $\Gamma_t(u) \in \mathcal{F}$ *with*

$$\frac{1}{2}\mathcal{E}_t(\Gamma_t(u)) \leq L||\Gamma_t(u)||_\infty \mathcal{E}_t(u) + ||\Gamma_t(u)||_\infty ||\Delta_t u||_2^2 \\ + C||\Delta_t u||_2 \sqrt{||\Gamma_t(u)^2||_\infty (||\Delta_t u||_2^2 + ||u||_2^2)}.$$

2. *There exists a finite nonnegative Borel measure μ_+ such that every $\mathcal{E}_t$-polar set is μ_+-negligible and for each $g \in \mathcal{F}$ the $\mathcal{E}_t$-q.c. representative $\tilde{g} \in L^1(X, \mu_+)$ with*

$$2\Gamma_{2,t}(u)(g) = \int g(\partial_t \Gamma_t)(u)\, dm_t + \int \tilde{g}\, d\mu_+.$$

   In particular $\Gamma_{2,t}(u)$ is a finite Borel measure with

$$2\Gamma_{2,t}(u) = (\partial_t \Gamma_t)(u) m + \mu_+.$$

Proof. Let $u_\varepsilon = h_\varepsilon^t u$. Choosing $\psi = \Gamma_t(u_\varepsilon)$ and $\varphi = -(\partial_t \Gamma_t)(u_\varepsilon) - 2\Gamma_t(u_\varepsilon, \Delta_t u_\varepsilon)$ in Lemma 3.11 and applying Bochner's inequality together with the Leibniz rule yields

$$\mathcal{E}_t(\Gamma_t(u_\varepsilon)) \leq -\int \Gamma_t(u_\varepsilon)((\partial_t \Gamma_t)(u_\varepsilon) + 2\Gamma_t(u_\varepsilon, \Delta_t u_\varepsilon))\, dm_t.$$

Applying the Leibniz rule once again we obtain

$$\mathcal{E}_t(\Gamma_t(u_\varepsilon)) \leq -\int (\Gamma_t(u_\varepsilon)(\partial_t \Gamma_t)(u_\varepsilon) - 2(\Delta_t u_\varepsilon)^2 \Gamma_t(u_\varepsilon) - 2\Gamma_t(u_\varepsilon, \Gamma_t(u_\varepsilon))\Delta_t u_\varepsilon)\, dm_t.$$

Note that as $\varepsilon \to 0$, $\Gamma(u_\varepsilon) \to \Gamma(u)$ pointwise, in L^1 and in the weak* L^∞ topology. The latter is due to the fact that $\Gamma(u_\varepsilon - u)$ is uniformly bounded and converges to 0 in L^1. Moreover by the uniform boundedness of $\Gamma(u_\varepsilon)$ in L^∞ we obtain that $\Gamma(u_\varepsilon) \to \Gamma(u)$ in L^2. Hence we find

$$\mathcal{E}_t(\Gamma_t(u)) \leq \liminf_{\varepsilon \to 0} \mathcal{E}_t(\Gamma_t(u_\varepsilon))$$

and

$$\int \Gamma_t(u)(\partial_t \Gamma_t)(u)\, dm_t = \int \Gamma_t(u)^2 e^{H_t}\, dm_t \\ = \lim_{\varepsilon \to 0} \int \Gamma_t(u_\varepsilon)^2 e^{H_t}\, dm_t = \lim_{\varepsilon \to 0} \int \Gamma_t(u_\varepsilon)(\partial_t \Gamma_t)(u_\varepsilon)\, dm_t,$$

while

$$\int (\Delta_t u)^2 \Gamma_t(u)\, dm_t = \lim_{\varepsilon \to 0} \int (h_\varepsilon^t \Delta_t u)^2 \Gamma_t(u_\varepsilon)\, dm_t = \lim_{\varepsilon \to 0} \int (\Delta_t u_\varepsilon)^2 \Gamma_t(u_\varepsilon)\, dm_t.$$

In order to show that

$$\lim_{\varepsilon\to 0}\int \Gamma_t(u_\varepsilon,\Gamma_t(u_\varepsilon))\Delta_t u_\varepsilon\, dm_t = \int \Gamma_t(u,\Gamma_t(u))\Delta_t u\, dm_t,$$

we show that $\Gamma_t(u_\varepsilon,\Gamma_t(u_\varepsilon))$ weakly converges to $\Gamma_t(u,\Gamma_t(u))$ in L^2. Take a sufficiently smooth testfunction φ ($\varphi \in \mathcal{F}\cap L^\infty$), then we easily deduce

$$\begin{aligned}\int \Gamma_t(u_\varepsilon,\Gamma_t(u_\varepsilon))\varphi\, dm_t &= -\int \Delta_t u_\varepsilon \Gamma_t(u_\varepsilon)\varphi\, dm_t - \int \Gamma_t(u_\varepsilon,\varphi)\Gamma_t(u_\varepsilon)\, dm_t\\ &\to -\int \Delta_t u \Gamma_t(u)\varphi\, dm_t - \int \Gamma_t(u,\varphi)\Gamma_t(u)\, dm_t\end{aligned}$$

by the strong L^2 convergence of $\Delta_t u_\varepsilon$, the weak*-L^∞ convergence of $\Gamma(u_\varepsilon)$ and the L^1 convergence of $\Gamma(u_\varepsilon,\varphi)$. Moreover $||\Gamma_t(u_\varepsilon,\Gamma_t(u_\varepsilon))||_2$ is uniformly bounded in ε since

$$\begin{aligned}\int |\Gamma_t(u_\varepsilon,\Gamma_t(u_\varepsilon))|^2\, dm_t &\le 4||\Gamma_t(u_\varepsilon)^2||_\infty C(||\Delta_t u_\varepsilon||_2^2 + ||u_\varepsilon||_2^2)\\ &\le C||\Gamma_t(u)^2||_\infty||(||\Delta_t u||_2^2 + ||u||_2^2))\end{aligned}$$

since the domain of the Laplacian coincides with the domain of the Hessian, cf. Section 2.8 in Chapter 2, [27]. Consequently we obtain that $\Gamma_t(u_\varepsilon,\Gamma_t(u_\varepsilon))$ weakly converges to $\Gamma_t(u,\Gamma_t(u))$ in L^2 since $\mathcal{F}\cap L^\infty$ is dense in L^2 [5, Theorem 4.5].

We conclude

$$\frac{1}{2}\mathcal{E}_t(\Gamma_t(u)) \le -\int \frac{1}{2}\Gamma_t(u)(\partial_t\Gamma_t)(u) - \Gamma_t(u)(\Delta_t u)^2 - \Gamma_t(u,\Gamma_t(u))\Delta_t u\, dm_t$$

$$\le L||\Gamma_t(u)||_\infty \mathcal{E}_t(u) + ||\Gamma_t(u)||_\infty||\Delta_t u||_2^2 + C||\Delta_t u||_2\sqrt{||\Gamma_t(u)^2||_\infty(||\Delta_t u||_2^2 + ||u||_2^2)}.$$

We show the second claim again by using the semigroup mollification $u_\varepsilon := h_\varepsilon^t u$. By Lemma 3.11 we deduce that

$$\begin{aligned}&\int g\, d\Delta_t^*\Gamma_t(u_\varepsilon) - \int \tilde{g}2\Gamma_t(u_\varepsilon,\Delta_t u_\varepsilon)\, dm_t\\ &\quad= \int \tilde{g}\, d\mu_+(u_\varepsilon) + \int \tilde{g}(\partial_t\Gamma_t)(u_\varepsilon)\, dm_t,\end{aligned}$$

where Δ_t^* is the measure valued Laplacian, and $\mu_+(u_\varepsilon)$ the nonnegative Borel measure with $\mu_+(u_\varepsilon)(X) \le \int (\Delta_t u_\varepsilon)^2 + \frac{1}{2}(\partial_t\Gamma_t)(u_\varepsilon)\, dm_t$. Hence, since $g = \tilde{g}$ q.e.

$$\begin{aligned}&\int g\, d\mu_+(u_\varepsilon)\\ &= \int -\Gamma_t(\Gamma_t(u_\varepsilon),g)\, dm_t + \int 2g(\Delta_t u_\varepsilon)^2 + 2\Gamma_t(g,u_\varepsilon)(\Delta_t u_\varepsilon)\, dm_t - \int g(\partial_t\Gamma_t)(u_\varepsilon)\, dm_t.\end{aligned}$$

Note that the right hand side converges as $\varepsilon\to 0$ since $\Gamma(u_\varepsilon)\to\Gamma(u)$ weakly in $\mathcal{F}$. Indeed, take a test function $\varphi \in Dom(\Delta_t)$. Then

$$\lim_{\varepsilon\to 0}\int \Gamma_t(\Gamma_t(u_\varepsilon),\varphi)\, dm_t = -\lim_{\varepsilon\to 0}\int \Gamma_t(u_\varepsilon)\Delta_t\varphi\, dm_t = \int \Gamma_t(\Gamma_t(u),\varphi)\, dm_t.$$

Since $\mathcal{E}_t(\Gamma_t(u_\varepsilon))$ is uniformly bounded in ε by the first claim and $Dom(\Delta_t)$ is dense in $\mathcal{F}$ we deduce that

$$\lim_{\varepsilon\to 0}\int \Gamma_t(\Gamma_t(u_\varepsilon),g)\,dm_t = \int \Gamma_t(\Gamma_t(u),g)\,dm_t \qquad \forall\, g\in\mathcal{F}.$$

Define the linear functional $\tilde{\mu}_+(u)\colon \mathcal{F}\cap L^\infty\to\mathbb{R}$ by

$$\tilde{\mu}_+(u)(g) := \lim_{\varepsilon\to 0}\int g\,d\mu_+(u_\varepsilon).$$

Note that if $g\geq 0$ we have $\tilde{\mu}_+(u)(g)\geq 0$ by the Bochner inequality. The Hahn-Banach theorem implies that there exists a linear functional $M\colon\mathcal{F}\to\mathbb{R}$ such that $M(g)=\mu_+(u)(g)$ for all $g\in\mathcal{F}\cap L^\infty$ and $M(g)\geq 0$ for all $g\in\mathcal{F}$ such that $g\geq 0$ a.e.. Moreover, if $g\in\mathcal{F}$ with $g\leq 1$ m-a.e.

$$M(g)=\mu_+(u)(g)=\lim_{\varepsilon\to 0}\int g\,d\mu_+(u_\varepsilon)\leq \mu_+(u_\varepsilon)(X)\leq \int (\Delta_t u)^2 + C\Gamma_t(u)\,dm_t.$$

Thus by Proposition 2.5 in [55] there exists a unique finite and nonnegative Borel measure μ_+ in X such that every $\mathcal{E}_t$-polar set is μ_+-negligible and for each $g\in\mathcal{F}$ the $\mathcal{E}_t$-q.c. representative $\tilde{g}\in L^1(X,\mu_+)$ with

$$M(g)=\int \tilde{g}\,d\mu_+.$$

Consequently

$$2\Gamma_{2,t}(u)(g)=\int g(\partial_t\Gamma_t)(u)\,dm_t+\int\tilde{g}\,d\mu_+,$$

and hence $\Gamma_{2,t}$ is measure valued with $2\Gamma_{2,t}(u)=(\partial_t\Gamma_t)(u)\,m_t+\,\mu_+$. $\square$

By virtue of Lebesgue's decomposition theorem we denote by $\gamma_{2,t}(u)\in L^1(X,m_t)$ the density wrt m_t

$$\Gamma_{2,t}(u)=\gamma_{2,t}(u)m_t+\Gamma_2^\perp(u),\quad \Gamma_2^\perp(u)\perp m_t,$$

and thus by the above Lemma

$$\gamma_{2,t}(u)\geq\frac{1}{2}(\partial_t\Gamma_t)(u)\ m\text{-a.e. and }\Gamma_{2,t}^\perp(u)\geq 0. \tag{128}$$

We define for $u,h\in Dom(\Delta_t)$ such that $\Gamma_t(u),\Gamma_t(h)\in L^\infty(X,m_t)$

$$\Gamma_{2,t}(u,h)(g):=\frac{1}{4}\Gamma_{2,t}(u+h)(g)-\frac{1}{4}\Gamma_{2,t}(u-h)(g),$$

where $g\in\mathcal{F}\cap L^\infty$. Note that the right-hand side is well-defined since the domain of the Laplacian and the Hessian coincide and

$$\Gamma_{2,t}(u,h)(g)=\int -\frac{1}{2}\Gamma_t(g,\Gamma_t(u,h))+g\Delta_t u\Delta_t h+\frac{1}{2}\Delta_t h\Gamma_t(u,g)+\frac{1}{2}\Delta_t u\Gamma_t(h,g)\,dm_t$$

Similarly,

$$\gamma_{2,t}(u,h):=\frac{1}{4}\gamma_{2,t}(u+h)-\frac{1}{4}\gamma_{2,t}(u-h).$$

The following Lemma is an adaptation of Lemma 3.3 in [55].

Lemma 3.14. *Let* $\bar{u} = (u_i)_{i=1}^n$ *with* $u_i \in Dom(\Delta_t)$ *such that* $u, \Gamma_t(u) \in L^\infty(X, m_t)$ *and let* $\Psi \in \mathcal{C}^3(\mathbb{R}^n)$ *with* $\Psi(0) = 0$. *Then*

$$\begin{aligned}\Gamma_{2,t}(\Psi(\bar{u})) = &\sum_{i,j} \Gamma_{2,t}(u_i, u_j)(\partial_i \Psi)(\bar{u})(\partial_j \Psi)(\bar{u}) \\ &+2 \sum_{i,j,k} (\partial_i \Psi)(\bar{u})(\partial_{jk} \Psi)(\bar{u}) H_t[u_i](u_j, u_k)\, m_t \\ &+ \sum_{i,j,k,h} (\partial_{ik} \Psi)(\bar{u})(\partial_{jh} \Psi)(\bar{u}) \Gamma_t(u_i, u_j)\Gamma_t(u_k, u_h)\, m_t.\end{aligned}$$

In particular m_t*-a.e.*

$$\begin{aligned}\gamma_{2,t}(\Psi(\bar{u})) = &\sum_{i,j} \gamma_{2,t}(u_i, u_j)(\partial_i \Psi)(\bar{u})(\partial_j \Psi)(\bar{u}) \\ &+2 \sum_{i,j,k} (\partial_i \Psi)(\bar{u})(\partial_{jk} \Psi)(\bar{u}) H_t[u_i](u_j, u_k) \\ &+ \sum_{i,j,k,h} (\partial_{ik} \Psi)(\bar{u})(\partial_{jh} \Psi)(\bar{u}) \Gamma_t(u_i, u_j)\Gamma_t(u_k, u_h).\end{aligned}$$

Proof. Note that $\Psi(\bar{u}) \in Dom(\Delta_t)$ with $\Gamma_t(u) \in L^\infty$ since

$$\begin{aligned}\Gamma_t(\Psi(\bar{u})) &= \sum_{i,j} \partial_i \Psi(\bar{u}) \partial_j \Psi(\bar{u}) \Gamma_t(u_i, u_j) \in L^1 \cap L^\infty, \\ \Delta_t(\Psi(\bar{u})) &= \sum_i \partial_i \Psi(\bar{u}) \Delta_t u_i + \sum_{i,j} \partial_{ij} \Psi(\bar{u}) \Gamma_t(u_i, u_j) \in L^2.\end{aligned}$$

Thus by definition for each $g \in \mathcal{F} \cap L^\infty$

$$2\Gamma_{2,t}(\Psi(\bar{u}))(g) = \int -\Gamma_t(g, \Gamma_t(\Psi(\bar{u}))) + 2g(\Delta_t \Psi(\bar{u}))^2 + 2\Gamma(g, \Psi(\bar{u}))\Delta_t \Psi(\bar{u})\, dm_t.$$

We calculate using the notation $\psi = \Psi(\bar{u})$, $\psi_i = \partial_i \Psi(\bar{u})$ and $\psi_{ij} = \partial_{ij} \Psi(\bar{u})$ for the first term

$$\begin{aligned}&\int -\Gamma_t(g, \Gamma_t(\Psi(\bar{u})))\, dm_t \\ =&\sum_{i,j} \Big\{ \int -\Gamma_t(g\psi_i\psi_j, \Gamma_t(u_i, u_j))\, dm \\ &\qquad + \int g\Big(\Gamma_t(u_i, u_j)\Delta_t(\psi_i\psi_j) + 2\Gamma_t(\psi_i\psi_j, \Gamma_t(u_i, u_j))\Big)\, dm_t\Big\} \\ =&\sum_{i,j} \int -\Gamma_t(g\psi_i\psi_j, \Gamma_t(u_i, u_j))\, dm + \int 2g\Big(I + II\Big)\, dm_t,\end{aligned}$$

where

$$I = \sum_{i,j,k,h} \Gamma_t(u_i, u_j)\Big(\psi_i(\psi_{jk}\Delta_t u_k + \psi_{jkh}\Gamma_t(u_k, u_h)) + \psi_{ik}\psi_{jh}\Gamma_t(u_k, u_h)\Big)$$

and

$$II = \sum_{i,j,k} \psi_i \psi_{jk} \Big(\Gamma_t(u_k, \Gamma_t(u_j, u_i)) + \Gamma_t(u_j, \Gamma_t(u_i, u_k))\Big).$$

On the other hand

$$
\begin{aligned}
&\int 2g(\Delta_t\Psi(\bar{u}))^2 + 2\Gamma_t(g,\Psi(\bar{u}))\Delta_t\Psi(\bar{u})\,dm_t\\
=&\sum_{i,j} 2\int \Big(\Delta_t u_i \Delta_t u_j g\psi_i\psi_j + \Gamma_t(u_i, g\psi_i\psi_j)\Delta_t u_j\Big)\,dm_t\\
&-\sum_{i,j,k,h}\int 2g\Big(\psi_i\Delta_t u_k\psi_{kj}\Gamma_t(u_i,u_j) + \psi_i\Gamma_t(u_k,u_h)\psi_{khj}\Gamma_t(u_i,u_j)\\
&\qquad\qquad + \psi_i\psi_{jk}\Gamma_t(u_i,\Gamma_t(u_j,u_k))\Big)\,dm_t.
\end{aligned}
$$

Adding up and collecting terms yields

$$
\begin{aligned}
&2\Gamma_{2,t}(\Psi(\bar{u}))(g)\\
=&\sum_{i,j}\int\Big(-\Gamma_t(g\psi_i\psi_j,\Gamma_t(u_i,u_j)) + 2g\psi_i\psi_j\Delta_t u_i\Delta_t u_j + 2\Gamma_t(u_i,g\psi_i\psi_j)\Delta_t u_j\Big)dm_t\\
+&\sum_{i,j,k}\int 2g\psi_i\psi_{jk}\Big(\Gamma_t(u_k,\Gamma_t(u_i,u_j)) + \Gamma_t(u_j,\Gamma_t(u_i,u_k)) - \Gamma_t(u_i,\Gamma_t(u_j,u_k))\Big)dm_t\\
+&\sum_{i,j,k,h}\int 2g\psi_{ik}\psi_{jk}\Gamma(u_k,u_h)\Gamma(u_i,u_j)\,dm_t\\
=&2\sum_{i,j}\Gamma_{2,t}(u_i,u_j)(g\psi_i\psi_j) + \sum_{i,j,k}\int 4g\psi_i\psi_{jk}(H_t[u_i](u_k,u_j))\,dm_t\\
&+\sum_{i,j,k,h}\int 2g\psi_{ik}\psi_{jk}\Gamma_t(u_k,u_h)\Gamma_t(u_i,u_j)\,dm_t
\end{aligned}
$$

for each $g\in\mathcal{F}\cap L^\infty$.

For arbitrary $g\in\mathcal{F}$, set $g^n := g\wedge n$. Then, by dominated convergence (recall that $\tilde{g}\in L^1(X,\mu_+)$)

$$
\begin{aligned}
\lim_{n\to\infty}\int g^n\,d\Gamma_{2,t}(\Psi(\bar{u})) &= \lim_{n\to\infty}\Big(\int g^n(\partial_t\Gamma_t)(\Psi(\bar{u}))\,dm_t + \int \tilde{g^n}\,d\mu_+\Big)\\
&= \int g\,d\Gamma_{2,t}(\Psi(\bar{u})).
\end{aligned}
$$

Similarly we can pass to the limit for the other integrals and obtain for all $g\in\mathcal{F}$

$$
\begin{aligned}
2\Gamma_{2,t}(\Psi(\bar{u}))(g) =& 2\sum_{i,j}\Gamma_{2,t}(u_i,u_j)(g\psi_i\psi_j) + \sum_{i,j,k}\int 4g\psi_i\psi_{jk}(H_t[u_i](u_k,u_j))\,dm_t\\
&+\sum_{i,j,k,h}\int 2g\psi_{ik}\psi_{jk}\Gamma_t(u_k,u_h)\Gamma_t(u_i,u_j)\,dm_t,
\end{aligned}
$$

and hence the result. $\square$

Proposition 3.15. *Suppose that Bochner's inequality holds at time t. Then for every $u\in Dom(\Delta_t)\cap L^\infty(X,m_t)$ such that $\Gamma_t(u)\in L^\infty(X,m_t)$*

$$
\Gamma_t(\Gamma_t(u)) \le 4\big(\gamma_{2,t}(u) - \frac{1}{2}\partial_t\Gamma_t(u)\big)\Gamma_t(u).
$$

Proof. We choose the same polynomial $\Psi\colon \mathbb{R}^3 \to \mathbb{R}$ as in [55] by

$$\Psi(\bar{u}) := \lambda u_1 + (u_2 - a)(u_3 - b) - ab, \qquad \lambda, a, b \in \mathbb{R},$$

where $\bar{u} = (u_1, u_2, u_3)$, where each $u_i \in Dom(\Delta_t) \cap L^\infty(X, m_t)$ with $\Gamma_t(u_i) \in L^\infty(X, m_t)$. We apply Lemma 3.13 and obtain

$$\gamma_{2,t}(\Psi(\bar{u})) \geq \frac{1}{2}(\partial_t \Gamma_t)(\Psi(\bar{u})) \quad m\text{-a.e. in } X, \tag{129}$$

where both sides of the inequality depend on $\lambda, a, b \in \mathbb{R}$. Choosing λ, a, b in a dense and countable subset D of $\mathbb{R}$ yields that (129) holds m-a.e. for all λ, a, b in D. Since

$$(\partial_t \Gamma_t)(\Psi(\bar{u})) = \sum_{i,j} \partial_i \Psi(\bar{u}) \partial_j \Psi(\bar{u}) (\partial_t \Gamma_t)(u_i, u_j),$$

and

$$\begin{aligned}\gamma_{2,t}(\Psi(\bar{u})) = &\sum_{i,j} \partial_i \Psi(\bar{u}) \partial_j \Psi(\bar{u}) \gamma_{2,t}(u_i, u_j) + 2 \sum_{i,j,k} \partial_i \Psi(\bar{u}) \partial_{jk} \Psi(\bar{u}) H_t[u^i](u_j, u_k) \\ &+ \sum_{i,j,k,h} \partial_{ik} \Psi(\bar{u}) \partial_{jh} \Psi(\bar{u}) \Gamma_t(u_i, u_j) \Gamma_t(u_k, u_h),\end{aligned}$$

cf. [55, Lemma 3.3], both sides are continuous in λ, a, b, and hence we conclude that (129) holds for all λ, a, b in $\mathbb{R}$.

Thus, for m-a.e. $x \in X$ we may set $a := u_2(x)$, $b := u_3(x)$ so that

$$\begin{gathered}\partial_1 \Psi(\bar{u})(x) = \lambda, \quad \partial_2 \Psi(\bar{u})(x) = 0 = \partial_3 \Psi(\bar{u})(x) \\ \partial_{23} \Psi(\bar{u})(x) = 1 = \partial_{32} \Psi(\bar{u})(x), \quad \partial_{ij} \Psi(\bar{u})(x) = 0 \text{ else,}\end{gathered}$$

m-a.e., and exploiting (129) yields

$$\lambda^2 \gamma_{2,t}(u_1) + 4\lambda H_t[u^1](u_2, u_3) + 2\Big(\Gamma_t(u_2, u_3)^2 + \Gamma_t(u_2)\Gamma_t(u_3)\Big) \geq \frac{1}{2}\lambda^2 (\partial_t \Gamma_t)(u_1).$$

Using Cauchy-Schwartz inequality $\Gamma_t(u_2, u_3)^2 \leq \Gamma_t(u_2)\Gamma_t(u_3)$ this can be transformed into

$$\lambda^2 \Big(\gamma_{2,t}(u_1) - \frac{1}{2}(\partial_t \Gamma_t)(u_1)\Big) + 4\lambda H_t[u^1](u_2, u_3) + 4\Gamma_t(u_2)\Gamma_t(u_3) \geq 0,$$

and since λ is arbitrary [27, Lemma 3.3.6] we obtain

$$(H_t[u_1](u_2, u_3))^2 \leq \Big(\gamma_{2,t}(u_1) - \frac{1}{2}(\partial_t \Gamma_t)(u_1)\Big)\Gamma_t(u_2)\Gamma_t(u_3).$$

From the definition of the Hessian we deduce that

$$H_t[u_1](u_2, u_3) + H_t[u_2](u_1, u_3) = \Gamma_t(\Gamma_t(u_1, u_2), u_3)$$

and consequently

$$|\Gamma_t(\Gamma_t(u_1, u_2), u_3)| \leq \sqrt{\Gamma_t(u_3)} \Big(\sqrt{\gamma_{2,t}(u_1) - \frac{1}{2}(\partial_t \Gamma_t)(u_1)} \sqrt{\Gamma_t(u_2)} \tag{130}$$

$$+ \sqrt{\gamma_{2,t}(u_2) - \frac{1}{2}(\partial_t \Gamma_t)(u_2)} \sqrt{\Gamma_t(u_1)} \Big). \tag{131}$$

We obtain (130) for arbitrary $u_3 \in \mathcal{F} \cap L^\infty(X, m_t)$ by approximating u_3 by a sequence u_3^n converging in energy with

$$\Gamma_t(u_3^n) \to \Gamma_t(u), \qquad \Gamma_t(u_3^n, \Gamma_t(u_1, u_2)) \to \Gamma_t(u_3, \Gamma_t(u_1, u_2))$$

pointwise and in $L^1(X, m_t)$, cf. Theorem 3.4 in [55] Hence we may choose $u_3 = \Gamma_t(u_1, u_2)$, and obtain the result choosing $u_1 = u_2$.

□

Now we are ready to prove Theorem 3.2. We will assume that

$$u_r \in \mathrm{Lip}(X) \text{ for all } r \in (s,t) \text{ with } \sup_{r,x} \mathrm{lip}_r u_r(x) < \infty. \tag{132}$$

Proof of Theorem 3.2. Define for each $\varepsilon > 0$ the concave and smooth function $\omega_\varepsilon(\cdot) := (\varepsilon + \cdot)^\alpha - \varepsilon^\alpha$. Note that this function satisfies

$$2\omega_\varepsilon'(r) + 4r\omega_\varepsilon''(r) \geq 0. \tag{133}$$

For each $s, t \in (0, T)$ under consideration as well as $u \in \mathrm{Lip}(X)$ and $g \in \mathcal{F} \cap L^\infty$ with $g \geq 0$, we set $u_r = P_{r,s}u$, $g_r = P_{t,r}^* g$ for $r \in [s,t]$. Note that for a.e. $r \in [s,t]$ $u_r \in Dom(\Delta_r)$ and $u, \Gamma_r(u) \in L^\infty(X, m_r)$.

We consider the function

$$h_r^\varepsilon := \int g_r \omega_\varepsilon(\Gamma_r(u_r)) dm_r.$$

Choose $s \leq \sigma < \tau \leq t$ and $\delta > 0$ sufficiently small that $\sigma \leq \tau - \delta$ such that

$$h_\tau^\varepsilon \leq \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_{\tau-\delta}^{\tau} h_r dr \quad \text{and} \quad h_\sigma^\varepsilon \geq \limsup_{\delta \searrow 0} \frac{1}{\delta} \int_{\sigma}^{\sigma+\delta} h_r dr.$$

Note that by Lebesgue's density theorem, this is true at least for a.e. $\sigma \geq s$ and for a.e. $\tau \leq t$. Then from

$$\int_{\tau-\sigma}^{\tau} h_r\, dr - \int_{\sigma}^{\sigma+\delta} h_r\, dr = \int_{\sigma}^{\tau-\delta} (h_{r+\delta} - h_r)\, dr,$$

and the concavity of ω_ε we deduce

$$\begin{aligned}
h_\tau^\varepsilon - h_\sigma^\varepsilon \leq & \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_{\sigma}^{\tau-\delta} \big[h_{r+\delta} - h_r\big] dr \\
\leq & \limsup_{\delta \searrow 0} \frac{1}{\delta} \int_{\sigma}^{\tau-\delta} \int_X \omega_\varepsilon(\Gamma_{r+\delta}(u_{r+\delta})) d(\mu_{r+\delta} - \mu_r)\, dr \\
& + \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_{\sigma}^{\tau-\delta} \int_X g_r \omega_\varepsilon'(\Gamma_r(u_r)) \Big[\Gamma_{r+\delta}(u_r) - \Gamma_r(u_r)\Big] dm_r\, dr \\
& + \limsup_{\delta \searrow 0} \frac{1}{\delta} \int_{\sigma}^{\tau-\delta} \int_X g_r \omega_\varepsilon'(\Gamma_r(u_r)) \Gamma_{r+\delta}(u_{r+\delta}, u_{r+\delta} - u_r)\, dm_r \\
& + \limsup_{\delta \searrow 0} \frac{1}{\delta} \int_{\sigma}^{\tau-\delta} \int_X g_r \omega_\varepsilon'(\Gamma_r(u_r)) \Gamma_{r+\delta}(u_{r+\delta} - u_r, u_r)\, dm_r\, dr \\
=: & (I) + (II) + (III') + (III'').
\end{aligned}$$

Let us denote with a slight abuse of notation $\hat{g}_r = g_r\omega'_\varepsilon(\Gamma_r(u_r))$. Note that $\hat{g} \in L^1 \cap L^\infty(X)$ and $\hat{g} \in \mathcal{F}$. Each of the four terms will be considered separately. Since $r \mapsto \mu_r$ is a solution to the dual heat equation, we obtain

$$\begin{aligned}(I) =& \limsup_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau-\delta} \int_X \omega_\epsilon(\Gamma_{r+\delta}(u_{r+\delta})) \cdot \Big(- \int_r^{r+\delta} \Delta_q g_q \, dm_q \, dq \Big) dr \\ =& - \liminf_{\delta \searrow 0} \int_{\sigma+\delta}^{\tau} \int_X \omega_\varepsilon(\Gamma_r(u_r)) \Big(\frac{1}{\delta} \int_{r-\delta}^{r} \Delta_q g_q e^{-f_q} \, dq \Big) dm_\diamond \, dr \\ =& - \int_\sigma^\tau \int_X \omega_\varepsilon(\Gamma_r(u_r)) \cdot \Delta_r g_r \, dm_r \, dr\end{aligned}$$

due Lebesgue's density theorem applied to $r \mapsto \Delta_r g_r e^{-f_r}$. Note that the latter function is in L^2 (Theorem 3.6) and the function $r \mapsto \omega_\varepsilon(\Gamma_r(u_r))$ is in L^∞ thanks to (132).

The second term can estimated according to Proposition 3.8:

$$\begin{aligned}(II) =& \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau-\delta} \int_X \hat{g}_r \Big[\Gamma_{r+\delta}(u_r) - \Gamma_r(u_r) \Big] dm_r \, dr \\ =& \int_\sigma^\tau \int_X \hat{g}_r \, (\partial_r \Gamma_r)(u_r) dm_r dr.\end{aligned}$$

The term (III') is transformed as follows

$$\begin{aligned}&(III') \\ =& \limsup_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau-\delta} \int_X \hat{g}_{r+\delta} \Gamma_{r+\delta}(u_{r+\delta}, u_{r+\delta} - u_r) \, dm_{r+\delta} \, dr \\ =& - \liminf_{\delta \searrow 0} \int_\sigma^{\tau-\delta} \int_X \Big(\Gamma_{r+\delta}(\hat{g}_{r+\delta}, u_{r+\delta}) + \hat{g}_{r+\delta} \, \Delta_{r+\delta} u_{r+\delta} \Big) \Big(\frac{1}{\delta} \int_r^{r+\delta} \Delta_q u_q \, dq \Big) dm_{r+\delta} \, dr \\ =& - \int_\sigma^\tau \int_X \Big(\Gamma_r(\hat{g}_r, u_r) + \hat{g}_r \, \Delta_r u_r \Big) \cdot \Delta_r u_r \, dm_r \, dr.\end{aligned}$$

Here again we used Lebesgue's density theorem (applied to $r \mapsto \Delta_r u_r$) and the 'nearly continuity' of $r \mapsto \hat{g}_r$ as map from (s,t) into $L^2(X,m)$ and as map into $\mathcal{F}$ (Lusin's theorem). Moreover, we used the boundedness (uniformly in r and x) of g_r and of $\Gamma_r(u_r)$ as well as the square integrability of $\Delta_r u_r$.

Similarly, the term (III'') will be transformed:

$$\begin{aligned}(III'') =& \limsup_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau-\delta} \int_X \hat{g}_r \Gamma_r(u_{r+\delta} - u_r, u_r) \, dm_r \, dr \\ =& - \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau-\delta} \int_X \Big(\Gamma_r(\hat{g}_r, u_r) + \hat{g}_r \Delta_r u_r \Big) \cdot \Big(\int_r^{r+\delta} \Delta_q u_q \, dq \Big) dm_r \, dr \\ =& - \int_\sigma^\tau \int_X \Big(\Gamma_r(\hat{g}_r, u_r) + \hat{g}_r \, \Delta_r u_r \Big) \cdot \Big(\Delta_r u_r \Big) dm_r \, dr.\end{aligned}$$

We therefore obtain

$$
\begin{aligned}
& h^\varepsilon_\tau - h^\varepsilon_\sigma = (I) + (II) + (III') + (III'') \\
\le & \int_\sigma^\tau \int_X \Big[-\omega_\varepsilon(\Gamma_r(u_r)) \cdot \Delta_r g_r + \hat{g}_r\, (\partial_r \Gamma_r)(u_r) - 2\big(\Gamma_r(\hat{g}_r, u_r) + \hat{g}_r\, \Delta_r u_r\big)\, \Delta_r u_r \Big] dm_r\, dr \\
= & \int_\sigma^\tau \int \Big[\Gamma_r(\Gamma_r(u_r), \hat{g}_r) - \Gamma_r(\Gamma_r(u_r)) \omega''_\varepsilon(\Gamma_r(u_r)) g_r + \hat{g}_r\, (\partial_r \Gamma_r)(u_r) \\
& \qquad - 2\big(\Gamma_r(\hat{g}_r, u_r) + \hat{g}_r\, \Delta_r u_r\big)\, \Delta_r u_r \Big] dm_r\, dr \\
= & \int_\sigma^\tau -2\Gamma_{2,r}(u_r)(\hat{g}_r)\, dr - \int_\sigma^\tau \int \Big[\Gamma_r(\Gamma_r(u_r)) \omega''_\varepsilon(\Gamma_r(u_r)) g_r + \hat{g}_r\, (\partial_r \Gamma_r)(u_r) \Big] dm_r\, dr.
\end{aligned}
$$

Applying (128), Proposition 3.15, (133) and taking into account the concavity of ω_ε we further deduce for a.e. $r \in [s,t]$,

$$
\begin{aligned}
& h^\varepsilon_\tau - h^\varepsilon_\sigma \\
\le & \int_\sigma^\tau \int_X \Big[-2\gamma_{2,r}(u_r)\hat{g}_r + \hat{g}_r\, (\partial_r \Gamma_r)(u_r) - \Gamma_r(\Gamma_r(u_r)) \omega''_\varepsilon(\Gamma_r(u_r)) g_r \Big]\, dm_r\, dr \\
\le & \int_\sigma^\tau \int_X \Big[-g_r \Big(\gamma_{2,r}(u_r) - \frac{1}{2}(\partial_r \Gamma_r)(u_r) \Big) \Big(2\omega'_\varepsilon(\Gamma_r(u_r)) + 4\omega''_\varepsilon(\Gamma_r(u_r)) \Gamma_r(u_r) \Big) \Big]\, dm_r\, dr \\
\le & 0.
\end{aligned}
$$

Hence we showed that, given u and g, there exists exceptional sets (which are null sets) for τ and σ outside of these sets

$$
\int_X \omega_\varepsilon(\Gamma_\tau(P_{\tau,\sigma} u)) g\, dm_\tau - \int_X P_{\tau,\sigma} \omega_\varepsilon(\Gamma_\sigma(u))\, g\, dm_\tau \le 0 \tag{134}
$$

holds. Choosing g's from a dense countable set one may achieve that the exceptional sets for σ and τ in (134) do not depend on g. Next we may assume that $\sigma, \tau \in [s,t]$ with $\sigma < \tau$ is chosen such that (134) simultaneously holds for all u from a dense countable set $\mathcal{C}_1$ in $\mathrm{Lip}(X)$. We approximate arbitrary $u \in \mathrm{Lip}(X)$ by $u_n \in \mathcal{C}_1$ in energy and in L^2 such that $\sqrt{\Gamma_\tau(P_{\tau,\sigma} u_n)} \rightharpoonup G$ in L^2, for some $G \in L^2(X)$. This is possible since $||\sqrt{\Gamma_\tau(P_{\tau,\sigma} u_n)}||_{L^2(X)}$ is uniformly bounded. Then we have on the one hand

$$
\limsup_{n\to\infty} \int_X P_{\tau,\sigma} \omega_\varepsilon(\Gamma_\sigma(u_n))\, g\, dm_\tau \le \int_X P_{\tau,\sigma} \omega_\varepsilon(\Gamma_\sigma(u))\, g\, dm_\tau \tag{135}
$$

since

$$
\begin{aligned}
& \int_X P_{\tau,\sigma} \omega_\varepsilon(\Gamma_\sigma(u_n))\, g\, dm_\tau - \int_X P_{\tau,\sigma} \omega_\varepsilon(\Gamma_\sigma(u))\, g\, dm_\tau \\
\le & \int_X P^*_{\tau,\sigma} g\, \omega'_\varepsilon(\Gamma_\sigma(u))(\Gamma_\sigma(u_n) - \Gamma_\sigma(u))\, dm_\sigma \\
\le & ||P^*_{\tau,\sigma} g\, \omega'_\varepsilon(\Gamma_\sigma(u))||_{L^\infty(X)} \left| \int_X \Gamma_\sigma(u_n) - \Gamma_\sigma(u)\, dm_\sigma \right|.
\end{aligned}
$$

On the other hand we find

$$
\liminf_{n\to\infty} \int_X \omega_\varepsilon(\Gamma_\tau(P_{\tau,\sigma} u_n)) g\, dm_\tau \ge \int_X \omega_\varepsilon(\Gamma_\tau(P_{\tau,\sigma} u)) g\, dm_\tau. \tag{136}
$$

Indeed, since $P_{\tau,\sigma}u_n \to P_{\tau,\sigma}u$ and $\sqrt{\Gamma(P_{\tau,\sigma}u_n)} \rightharpoonup G$ in $L^2(X)$ we know $\Gamma(P_{\tau,\sigma}u) \leq G^2$ m-a.e. and hence

$$\begin{aligned}
&\int_X \omega_\varepsilon(\Gamma_\tau(P_{\tau,\sigma}u_n))g\,dm_\tau - \int_X \omega_\varepsilon(\Gamma_\tau(P_{\tau,\sigma}u))g\,dm_\tau \\
&= \int_X \tilde{\omega}_\varepsilon(\sqrt{\Gamma_\tau(P_{\tau,\sigma}u_n)})g\,dm_\tau - \int_X \tilde{\omega}_\varepsilon(\sqrt{\Gamma_\tau(P_{\tau,\sigma}u)})g\,dm_\tau \\
&\geq \int_X \tilde{\omega}'_\varepsilon(\sqrt{\Gamma_\tau(P_{\tau,\sigma}u)})(\sqrt{\Gamma_\tau(P_{\tau,\sigma}u_n)} - \sqrt{\Gamma_\tau(P_{\tau,\sigma}u)})g\,dm_\tau \\
&\geq \int_X \tilde{\omega}'_\varepsilon(\sqrt{\Gamma_\tau(P_{\tau,\sigma}u)})(\sqrt{\Gamma_\tau(P_{\tau,\sigma}u_n)} - G)g\,dm_\tau,
\end{aligned}$$

where $\tilde{\omega}(r) = \omega(r^2)$, which is convex and monotone. Combining (134), (135) and (136) yields

$$\begin{aligned}
&\int_X \omega_\varepsilon(\Gamma_\tau(P_{\tau,\sigma}u))g\,dm_\tau - \int_X P_{\tau,\sigma}\omega_\varepsilon(\Gamma_\sigma(u))\,g\,dm_\tau \\
&\leq \liminf_n \int_X \omega_\varepsilon(\Gamma_\tau(P_{\tau,\sigma}u_n))g\,dm_\tau - \limsup_n \int_X P_{\tau,\sigma}\omega_\varepsilon(\Gamma_\sigma(u_n))\,g\,dm_\tau \\
&\leq \liminf_n \left(\int_X \omega_\varepsilon(\Gamma_\tau(P_{\tau,\sigma}u_n))g\,dm_\tau - \int_X P_{\tau,\sigma}\omega_\varepsilon(\Gamma_\sigma(u_n))\,g\,dm_\tau\right) \leq 0.
\end{aligned}$$

Letting $\varepsilon \to 0$ we showed that

$$\int_X (\Gamma_\tau(P_{\tau,\sigma}u))^\alpha g\,dm_\tau \leq \int_X P_{\tau,\sigma}(\Gamma_\sigma(u)^\alpha)\,g\,dm_\tau. \tag{137}$$

Since $\mathrm{Lip}(X)$ is dense in $\mathcal{F}$ we can extend (137) to arbitrary $u \in \mathcal{F}$. Since g is arbitrary we obtain the result. □

3.3 Application to Super-Ricci flows and Couplings of Brownian Motions

In this section we apply the previous results to super-Ricci flows as defined in Chapter 2. We recall that the defining property is the relative entropy $S\colon I \times \mathcal{P}(X) \to (-\infty, \infty]$ given by

$$S_t(\mu) = \int \rho \log \rho \, dm_t$$

whenever $\mu = \rho m_t$, and $S_t(\mu) = \infty$ otherwise.

We proved the following.

Definition 3.16. *We say that (X, d_t, m_t) is a super-Ricci flow if one of the following equivalent assertions holds*

i) For a.e. $t \in (0,T)$ and every W_t-geodesic $(\mu^a)_{a\in[0,1]}$ in $\mathcal{P}(X)$ with $\mu^0, \mu^1 \in$ Dom(S)

$$\partial_a^+ S_t(\mu^a)\big|_{a=1-} - \partial_a^- S_t(\mu^a)\big|_{a=0+} \geq -\frac{1}{2}\partial_t^- W_{t-}^2(\mu^0,\mu^1) \tag{138}$$

('dynamic convexity').

ii) For all $0 \le s < t \le T$ *and* $\mu, \nu \in \mathcal{P}(X)$

$$W_s(\hat{P}_{t,s}\mu, \hat{P}_{t,s}\nu) \le W_t(\mu, \nu) \tag{139}$$

('transport estimate').

iii) For all $u \in Dom(\mathcal{E})$ *and all* $0 < s < t < T$

$$\big|\nabla_t(P_{t,s}u)\big|_*^2 \le P_{t,s}\big(|\nabla_s u|_*^2\big) \tag{140}$$

('gradient estimate').

iv) For all $0 < s < t < T$ *and for all* $u_s, g_t \in \mathcal{F}$ *with* $g_t \ge 0$, $g_t \in L^\infty$, $u_s \in \mathrm{Lip}(X)$ *and for a.e.* $r \in (s,t)$

$$\mathbf{\Gamma}_{2,r}(u_r)(g_r) \ge \frac{1}{2}\int \overset{\bullet}{\Gamma}_r\,(u_r)g_r dm_r \tag{141}$$

('dynamic Bochner inequality' or 'dynamic Bakry-Emery condition') where $u_r = P_{r,s}u_s$ *and* $g_r = P^*_{t,r}g_t$. *Moreover, the regularity assumption* (132) *is satisfied.*

The following corollary is a consequence of Theorem 3.1 and Theorem 3.2. In particular, choosing $\mu = \delta_x$ and $\nu = \delta_y$ for some arbitrary $x, y \in X$, Corollary 3.17 implies for $p = \infty$

$$W_{\infty,s}(\hat{P}_{t,s}\delta_x, \hat{P}_{t,s}\delta_y) \le d_t(x,y). \tag{142}$$

Corollary 3.17. *Suppose that* $(X, d_t, m_t)_{t\in I}$ *is a super-Ricci flow satisfying the assumptions in Theorem 3.1. Then*

i) for every $u \in \mathcal{F} \cap L^\infty(X, m_s)$ *and every* $\beta \in [1,2]$

$$|\nabla_t P_{t,s}u|_*^\beta \le P_{t,s}(|\nabla_s u|_*^\beta), \tag{143}$$

ii) for every $\mu, \nu \in \mathcal{P}(X)$ *and every* $p \in [1, \infty]$

$$W_{p,s}(\hat{P}_{t,s}\mu, \hat{P}_{t,s}\nu) \le W_{p,t}(\mu, \nu). \tag{144}$$

Proof. Note that, taking into account $\Gamma(u) = |\nabla u|_*^2$ due to our static Riemannian curvature bound, (143) holds at least for a.e. $s \le t$ by Definition 3.16, Theorem 3.1 and Theorem 3.2. Then applying Kuwada's duality [36, Theorem 2.2] implies that (144) holds at all these time instances. Indeed, (143) implies that for all $u \in \mathrm{Lip}_b(X)$, $|\nabla_t P_{t,s}u|_* \le P_{t,s}(|\nabla_s u|_*^\beta)^{1/\beta}$ and thus by Proposition 3.11 in [7] $\mathrm{lip}_t P_{t,s}u \le P_{t,s}(|\nabla_s u|_*^\beta)^{1/\beta}$. We obtain

$$\mathrm{lip}_t(P_{t,s}u) \le P_{t,s}(\mathrm{lip}_s(u)^\beta)^{1/\beta}$$

by virtue of $|\nabla u|_* \le \mathrm{lip}\,u$ (Lemma 4.4 in [5]) and the monotonicity of the functions $P_{t,s}$, r^β and $r^{1/\beta}$. We deduce from Theorem 2.2 in [36] for a.e. $s \le t$

$$W_{p,s}(\hat{P}_{t,s}\mu, \hat{P}_{t,s}\nu) \le W_{p,t}(\mu, \nu),$$

where p is the Hölder conjugate of β; $1/p + 1/\beta = 1$. Since both sides of (144) are continuous in s and t (see Lemma 3.18 below), we obtain that (144) holds for all times $s \le t$ and thus also (143) holds for all times by Theorem 2.2 in [36]. The same applies to $p = 1$ in (143) by noting that $\mathrm{lip}_t(P_{t,s}u) \le P_{t,s}(\mathrm{lip}_s(u)^\beta)^{1/\beta}$ for all $\beta \ge 2$ by virtue of Jensen's inequality. □

Lemma 3.18. *We obtain the following continuity estimate for the heat flow $\mu_s = \hat{P}_{t,s}\mu$, where $\mu \in \mathcal{P}(X)$. There exist constants $c, c' > 0$ depending only on K, N and L such that*

$$W_{p,t}(\mu_s, \mu_{s'})^p \leq c|s-s'|^{p/2} e^{c'|s-s'|/2}.$$

for all $0 \leq s, s' \leq t$.

Proof. Assume $0 < s < s' < t$. Then by $\mu_s = \hat{P}_{s',s}\mu_{s'}$ we estimate

$$W_{p,t}(\mu_s, \mu_{s'})^p \leq \int\int d_t^p(x,y) p_{s',s}(x,y)\, dm_s(y)\, d\mu_{s'}(x).$$

By virtue of the Gaussian upper bounds ([41, Section 4]) and the Bishop Gromov volume comparison in RCD(K,N) spaces ([57, Theorem 2.3]) we obtain for $\sigma = s' - s$ and $B_t(r,x)$ denoting the ball of radius r around x in the metric space (X, d_t)

$$\begin{aligned} p_{s',s}(x,y) &\leq \frac{C}{m_t(B_t(\sqrt{\sigma},x))} \cdot \exp\Big(-\frac{d_t^2(x,y)}{C\sigma}\Big) \\ A(R,x) &\leq \Big(\frac{R}{r}\Big)^{N-1} \cdot e^{R\sqrt{|K|(N-1)}} \cdot A(r,x) \end{aligned}$$

for $R \geq r$ where $A(r,x) = \partial_{r+} m_t(B_t(r,x))$ and thus (by integrating from 0 to $\sqrt{\sigma}$)

$$A(R,x) \leq N \frac{R^{N-1}}{\sigma^{N/2}} \cdot e^{R\sqrt{|K|(N-1)}} \cdot m_t(B_t(\sqrt{\sigma},x))$$

for $R \geq \sqrt{\sigma}$. Then estimating further yields (with varying constants)

$$\begin{aligned} &\int\int d_t^p(x,y) p_{s',s}(x,y)\, dm_s(y)\, d\mu_{s'}(x) \\ &\leq \int_X \Big[\frac{C}{m_t(B_t(\sqrt{\sigma},x))} \cdot \int_X d_t^p(x,y) \cdot \exp\Big(-\frac{d_t^2(x,y)}{C\sigma}\Big) dm_t(y)\Big] d\mu_{s'}(x) \\ &\leq C\sigma^{p/2} + C\int_X \int_{\sqrt{\sigma}}^{\infty} R^p \cdot \exp\Big(-\frac{R^2}{C\sigma}\Big) N \frac{R^{N-1}}{\sigma^{N/2}} \cdot e^{R\sqrt{|K|(N-1)}}\, dR\, d\mu_{s'}(x) \\ &\leq C\sigma^{p/2} + c''\sigma^{p/2} e^{c'\sigma/2} \leq c\sigma^{p/2} e^{c'\sigma/2}. \end{aligned} \tag{145}$$

□

Brownian motions

In the remainder of this section we follow the approach in [58] and construct couplings of two Brownian motions $(X_s^1)_{s\leq t}, (X_s^2)_{s\leq t}$ on X such that the distance d_s between X_s^1 and X_s^2 does not increase.

Definition 3.19. *Let $\mu \in \mathcal{P}(X)$ and $t \in I$. We call a stochastic process $(X_s)_{s\leq t}$ on a probability space $(\Omega, \Sigma, \mathbb{P})$ with values in X a Brownian motion on X with initial distribution μ if the process is sample-continuous and if for all $s \leq t$*

$$\mathbb{P}[X_s \in A] = \hat{P}_{t,s}(\mu)(A) = \int_X \int_A p_{t,s}(x,y)\, dm_s(y)\, d\mu(x).$$

Remark. *Let us remark that the Brownian motion defined here has time-dependent generator Δ_s instead of $\frac{1}{2}\Delta_s$. This is only for convenience and the stochastic process with generators $(\frac{1}{2}\Delta_s)_{s\leq t/2}$ is given by $(\tilde{X}_s)_{s\leq t/2}$, where $\tilde{X}_s := X_{2s}$.*

In order to prove existence of a Brownian motion we consider for fixed $t \in I$ the finite subset $J = \{t_1, \cdots, t_r\}$ of $(0,t]$ and the finite dimensional distribution P^μ_J, where $\mu \in \mathcal{P}(X)$, defined by

$$\begin{aligned} &P^\mu_J(B_r \times \ldots \times B_1) \\ :=& \int_X \int_{B_r} \cdots \int_{B_1} p_{t_2,t_1}(x_{t_2}, x_{t_1})\, dm_{t_1}(x_{t_1}) \ldots p_{t,t_r}(x, x_{t_r})\, dm_{t_r}(x_{t_r})\, d\mu(x). \end{aligned}$$

The family of probability measures $\{P^\mu_J | J \text{ finite } \subset (0,t]\}$ defines a projective family, hence the Kolmogorov extension theorem [16, Theorem 35.5] implies that there exists a unique probability measure $P^\mu_{(0,t]}$ on $(X^{(0,t]}, \mathcal{B}(X)^{(0,t]})$ such that $(\pi_J)_\# P^\mu_{(0,t]} = P^\mu_J$. Here, π_J denotes the projection $\omega \mapsto (\omega(t_1), \ldots, \omega(t_r))$ from $X^{(0,t]}$ to X^r.

For every $s \in (0,t]$ the map $\pi_s \colon \omega \mapsto \omega(s)$ from $X^{(0,t]}$ to X is a stochastic process with finite-dimensional distributions $(P^\mu_J)_J$. The following Proposition yields existence of a continuous modification $(X_s)_{s\leq t}$, and hence a Brownian motion.

Proposition 3.20. *For each $t \in I$ and each $\mu \in \mathcal{P}(X)$ there exists a Brownian motion on X with initial distribution μ, which is unique in law.*

Proof. We need to show that there exists positive constants $\alpha, \beta, c > 0$ such that the above mentioned process π_s satisfies

$$E[d(\pi_{s'}, \pi_s)^\alpha] \leq c|s - s'|^{1+\beta} \tag{146}$$

for all $s', s \in (0,t]$. Then the Kolmogorov continuity theorem [16, Theorem 39.3] implies that there exists a modification $(X_s)_{s\leq t}$ such that the map $s \mapsto X_s(\omega)$ is continuous for $P^\mu_{(0,t]}$-a.e ω. Hence the process $(X_s)_{s\leq t}$ on the probability space $(X^{(0,t]}, \mathcal{B}(X)^{(0,t]}, P^\mu_{(0,t]})$ yields the desired properties. For $\alpha > 2$ (146) follows from (145) in the proof of Lemma 3.18.

Since all finite-dimensional distributions are uniquely determined this process is unique in law. $\square$

Couplings of Brownian motions

We introduce the σ-field $\mathcal{B}^u(X^2) := \bigcap_{\nu\in\mathcal{P}(X^2)} \mathcal{B}^\nu(X^2)$ of universally measurable subsets of X^2, i.e. the intersection of all $\mathcal{B}^\nu(X^2)$, where ν runs through the set $\mathcal{P}(X^2)$ and where $\mathcal{B}^\nu(X^2)$ denotes the completion of the Borel σ-field on X^2 w.r.t. $\nu \in \mathcal{P}(X^2)$. Let $\mathrm{D} := \{k2^{-n} | k, n \in \mathbb{N}\} \cap (0,t]$ denote the set of nonnegative dyadic number s in $(0,t]$ and $\mathrm{D}_n := \{k2^{-n} | k \in \mathbb{N}\} \cap (0,t]$ for fixed $n \in \mathbb{N}$.

In the remainder we will asume that the transport estimate (144) holds for all $p \in [1,\infty]$.

Lemma 3.21. *For each $s \leq t$ there exists a Markov kernel $q^*_{t,s}$ on $(X^2, \mathcal{B}^u(X^2))$ with the following properties:*

i) For each $(x,y) \in X^2$ the probability measure $q^*_{t,s}((x,y),\cdot)$ is a coupling of the probability measures $p_{t,s}(x,\cdot)$ and $p_{t,s}(y,\cdot)$.

ii) For each $(x,y) \in X^2$ and $q^*_{t,s}((x,y),\cdot)$-a.e. $(x',y') \in X^2$

$$d_s(x',y') \le d_t(x,y).$$

Proof. By virtue of the transport estimate (142) there exists at least one probability measures with properties *i)* and *ii)*. Indeed, define $\mu_s = \hat{P}_{t,s}\delta_x$, $\nu_s = \hat{P}_{t,s}\delta_y$ and let $\gamma_p \in \Pi(\mu_s,\nu_s)$ such that $W_{p,s}(\mu_s,\nu_s) = ||d_s||_{L^p(\gamma_p)}$. Since $\gamma_p \in \Pi(\mu_s,\nu_s)$, $(\gamma_p)_{p\in\mathbb{N}}$ is tight ([62, Lemma 4.4]) and hence there exists a subsequence p_k and a probability measure γ such that γ_{p_k} weakly converges to γ. Since $\Pi(\mu_s,\nu_s)$ is closed we obtain that $\gamma \in \Pi(\mu_s,\nu_s)$. Moreover, since $d_s \wedge R \in \mathcal{C}_b(X \times X)$

$$||d_s \wedge R||_{L^p(\gamma)} = \lim_{k\to\infty} ||d_s \wedge R||_{L^p(\gamma_{p_k})} \le \lim_{k\to\infty} ||d_s||_{L^{p_k}(\gamma_{p_k})} \le d_t(x,y),$$

where the second inequality follows from the Hölder inequality and the last from Corollary 2.15. Letting $R \to \infty$ and $p \to \infty$, we obtain

$$||d_s||_{L^\infty(\gamma)} \le d_t(x,y).$$

Hence the set of all these couplings γ is non-empty and satisfies *i)* and *ii)*. Moreover, for given $x,y \in X$ this set is closed w.r.t. weak convergence in $\mathcal{P}(X^2)$. According to a measurable selection theorem [17, Theorem 6.9.2] we may choose a coupling $q^*_{t,s}((x,y),\cdot)$ such that the map

$$(x,y) \mapsto q^*_{t,s}((x,y),\cdot), \qquad (X^2,\mathcal{B}^u(X^2)) \to (\mathcal{P}(X^2),\mathcal{B}(\mathcal{P}(X^2)))$$

is measurable. $\square$

Lemma 3.22. *For each $n \in \mathbb{N}$ and $s,s' \in \mathrm{D}_n$ there exists a Markov kernel $q^{(n)}_{s,s'}$ on $(X^2,\mathcal{B}^u(X^2))$ with the following properties:*

i) For each $(x,y) \in X^2$ the probability measure $q^{(n)}_{s,s'}((x,y),\cdot)$ is a coupling of $p_{s,s'}(x,\cdot)$ and $p_{s,s'}(y,\cdot)$.

ii) For each $(x,y) \in X^2$

$$d_{s'}(x',y') \le d_s(x,y)$$

for $q^{(n)}_{s,s'}((x,y),\cdot)$-a.e. (x',y').

Proof. For $s = l2^{-n}$ and $s' = k2^{-n}$ with $l \ge k$ we put

$$q^{(n)}_{s,s'} := q^*_{(k+1)2^{-n},s'} \circ \ldots \circ q^*_{s,(l-1)2^{-n}}.$$

Obviously we have for $r \in \mathrm{D}_n$ such that $s' \le r \le s$,

$$q^{(n)}_{r,s'} \circ q^{(n)}_{s,r} = q^{(n)}_{s,s'} \tag{147}$$

and the properties *i)* and *ii)* hold by iteration, cf. Lemma 2.3 in [58]. $\square$

We fix a distribution $\nu \in \mathcal{P}(X^2)$ with marginals ν_1 and ν_2. Similarly as before for any finite subset $J = \{t_1, \ldots, t_r\}$ of D_n we consider the finite-dimensional distribution $Q_J^{(n)}$ on $(X^2)^{|J|}$

$$\begin{aligned}
&Q_J^{(n)}(A_r \times \ldots \times A_1)\\
&= \int_{X^2} \int_{A_r} \cdots \int_{A_1} q_{t_2,t_1}^{(n)}((x_2,y_2), d(x_1,y_1)) \ldots q_{t,t_r}^{(n)}((x,y), d(x_r,y_r))\nu(d(x,y)),
\end{aligned}$$

where $q^*_{t,t_r} = q^{(n)}_{l2^{-n},t_r} \circ q^*_{t,l2^{-n}}$ whenever $l2^{-n} < t < (l+1)2^{-n}$.

Lemma 3.23. *For fixed finite $J \subset \mathrm{D}_m$ the family $\{Q_J^{(n)} | n \in \mathbb{R}, n \geq m\}$ is a tight family of probability measures on $(X^2)^{|J|}$.*

Proof. Let $J = \{t_1, \ldots, t_r\}$ with each $t_i \in \mathrm{D}_m$. The families $\{\hat{P}_{t,t_i}(\nu_1) | i = 1, \ldots, r\}$ and $\{\hat{P}_{t,t_i}(\nu_2) | i = 1, \ldots, r\}$ are tight by virtue of Prokhorov's theorem, see e.g. [17]. This means that given $\varepsilon > 0$ there exist compact sets $B_1, B_2 \subset X$ such that for all $i = 1, \ldots, r$

$$\hat{P}_{t,t_i}(\nu_1)(X \setminus B_1) < \varepsilon, \quad \hat{P}_{t,t_i}(\nu_2)(X \setminus B_2) < \varepsilon.$$

Applying $A_1 \times A_2 \subset X \times A_2 \cup A_1 \times X$ and (147) yields for the compact set $\vec{B} = (B_1 \times B_2)^r$ and $n \in \mathbb{N}$

$$\begin{aligned}
Q_J^{(n)}((X^2)^r \setminus \vec{B}) &\leq \sum_{i=1}^r Q_{t,t_i}^{(n)}(X^2 \setminus B_1 \times B_2)\\
&\leq \sum_{i=1}^r \left[Q_{t,t_i}^{(n)}((X \setminus B_1) \times X) + Q_{t,t_i}^{(n)}((X \times (X \setminus B_2)) \right]\\
&= \sum_{i=1}^r \left[\hat{P}_{t,t_i}(\nu_1)(X \setminus B_1) + \hat{P}_{t,t_i}(\nu_2)(X \setminus B_2) \right]\\
&\leq 2r\varepsilon,
\end{aligned}$$

where the last two inequalities follow from *i)* of Lemma 3.22 and the tightness of $\{\hat{P}_{t,t_i}(\nu_j)\}_i$ respectively. Hence the family $\{Q_J^{(n)} | n \in \mathbb{R}, n \geq m\}$ is tight. $\square$

For $J = \{t_1, \ldots, t_r\}$ as above we set

$$\vec{e_1} \colon (X^2)^r \to X^r, \quad ((x_1,y_1), \ldots, (x_r,y_r)) \mapsto (x_1, \ldots, x_r),$$

and similarly for $\vec{e_2}$.

Proposition 3.24. *There exists a projective family $\{Q_J^\nu | J \text{ finite } \subset \mathrm{D}\}$ of probability measures and a subsequence $(n_l)_{l \in \mathbb{N}}$ such that for each finite $J \subset \mathrm{D}$*

i) $Q_J^{(n_l)} \to Q_J^\nu$ *weakly in* $\mathcal{P}((X^2)^{|J|})$ *as* $l \to \infty$,

ii) and $(\vec{e_1})_\# Q_J^\nu = P_J^{\nu_1}$, $(\vec{e_2})_\# Q_J^\nu = P_J^{\nu_2}$.

In particular there exists a probability measure $Q_{\mathrm{D}}^\nu \in \mathcal{P}((X^2)^{\mathrm{D}})$ such that for all finite $J \subset \mathrm{D}$

$$(\pi_J)_\# Q_{\mathrm{D}}^\nu = Q_J^\nu$$

and

$$(\vec{e_1})_\# Q_{\mathrm{D}}^\nu = P_{\mathrm{D}}^{\nu_1}, \quad (\vec{e_2})_\# Q_{\mathrm{D}}^\nu = P_{\mathrm{D}}^{\nu_2}.$$

Proof. Lemma 3.23 yields for each fixed J the existence of a weakly converging subsequence $Q_J^{(n_l)}$ by virtue of Prokhorov's theorem. By a diagonal argument we may choose a subsequence such that $Q_J^{(n_l)}$ weakly converges for all finite $J \subset \mathrm{D}$. Note that

$$(\vec{e_1})_\# Q_J^{(n_l)} = P_J^{\nu_1}, \quad (\vec{e_2})_\# Q_J^{(n_l)} = P_J^{\nu_2}$$

and hence the same holds true for the limit. We obtain the last assertion by applying Kolmogorov's extension theorem. □

The next theorem is in particular true for super-Ricci flows satisfying additionally (117) and (118).

Theorem 3.25. *Let* $(X, d_t, m_t)_{t\in I}$ *be a family of RCD*(K,N) *spaces such that* (113) *and* (114) *hold. Moreover we assume that the transport estimate* (144) *holds for every* $p \in [1,\infty]$*. Then, for each* $x, y \in X$ *there exists a continuous stochastic process* $(X_s)_{s\le t}$ *such that* $(X_s)_{s\le t}$ *is a coupling of the Brownian motions* $(X_s^1)_{s\le t}$ *and* $(X_s^2)_{s\le t}$ *with values in* X *and initial distributions* δ_x *and* δ_y *respectively and it satisfies for* $Q_{\mathrm{D}}^{(\delta_x,\delta_y)}$*-a.e. path*

$$d_s(X_s^1, X_s^2) \le d_t(x,y),$$

for each $s \le t$.

Proof. Set $\nu = (\nu_1, \nu_2) = (\delta_x, \delta_y)$. Consider the coordinate process $\pi_s = (\pi_s^1, \pi_s^2)\colon (X^2)^{\mathrm{D}} \to X^2$. Under Q_{D}^ν the process $(\pi_s^1)_{s\in\mathrm{D}}$ has distribution $P_{\mathrm{D}}^{\nu_1}$ and satisfies the continuity property (146). The corresponding statement holds true for the process $(\pi_s^2)_{s\in\mathrm{D}}$. Hence, the process $\pi_t = (\pi_s^1, \pi_s^2)$ satisfies the Kolmogorov continuity theorem for $\alpha > 2$ since

$$\begin{aligned} E[\hat{d}_t(\pi_s, \pi_{s'})^\alpha] \le & 2^{\alpha/2}\Big(E[d_t(\pi_s^1, \pi_{s'}^1)^\alpha] + E[d_t(\pi_s^2, \pi_{s'}^2)^\alpha]\Big) \\ \le & c2^{\alpha/2}|s-s'|^{\alpha/2}, \end{aligned}$$

with product metric $\hat{d}^2((x^1,y^1),(x^2,y^2)) = d^2(x^1,x^2) + d^2(y^1,y^2)$. Consequently there exists a continuous modification $(X_s)_{s\le t} = (X_s^1, X_s^2)_{s\le t}$ defined by $X_s = \lim_{s'\to s, s\in\mathrm{D}} \pi_{s'}$ for Q_{D}^ν-a.e. ω and all $s \le t$, cf. Lemma 63.5 in [15]. The process $(X_s^i)_{s\le t}$, $i = 1,2$ is a Brownian motion by continuity of $s \mapsto p_{t,s}(x,dy)$.

We need to justify that for Q_{D}^ν-a.e. path

$$d_s(X_s^1, X_s^2) \le d_t(x,y).$$

For each $n \in \mathbb{N}$ let $Q_{\mathrm{D}_n}^{(n)}$ be the projective limit of the family $(Q_J^{(n)})_{J\subset\mathrm{D}_n}$, which exists thanks to the Kolmogorov extension theorem. Consider the coordinate process $(\pi_s^{(n)})_{s\in\mathrm{D}_n} = (\pi_s^{1,(n)}, \pi_s^{2,(n)})_{s\in\mathrm{D}_n}$ from $(X^2)^{\mathrm{D}_n} \to X^2$. Then $Q_{\mathrm{D}_n}^{(n)}$-a.e. we have $d_s(\pi_s^{1,(n)}, \pi_s^{2,(n)}) \le d(x,y)$ by virtue of Lemma 3.22. Applying Proposition 3.24 and *ii)* of Lemma 3.22 we obtain for a subsequence

$$\begin{aligned} & E\big[(d_s(\pi_s^1, \pi_s^2)\wedge R)^p\big]^{1/p} = \lim_{l\to\infty} E\big[(d_s(\pi_s^{1,(n_l)}, \pi_s^{2,(n_l)})\wedge R)^p\big]^{1/p} \\ & \le \lim_{l\to\infty} E\big[(d_t(x,y)\wedge R)^p\big]^{1/p} = d_t(x,y)\wedge R, \end{aligned}$$

for each $s \in \mathrm{D}$. Letting R and p tend to ∞ we find for each $s \in \mathrm{D}$

$$d_s(\pi_s^1, \pi_s^2) \leq d_t(x, y).$$

Since the process $(X_s)_{s \in \mathrm{D}}$ is a modification we get for each $s \in \mathrm{D}$ and Q_{D}^{ν}-a.e. $d_s(X_s^1, X_s^2) \leq d_t(x, y)$. Since $\mathrm{D} \subset (0, t]$ is a dense and countable subset we obtain the result by continuity of $s \mapsto X_s(\omega)$.

$\square$

4 Gradient Flow for the Boltzmann Entropy and Cheeger's Energy on Time-dependent Metric Measure Spaces

In this chapter we study notions of gradient flows on metric spaces where the functional as well as the metric varies in time. Our main focus will be on two cases; Cheeger's energy on the space of L^2-integrable functions as well as the relative entropy on the space of Borel probability measures. Recalling the heat flow and its adjoint introduced in Chapter 2 we show that the first can be equivalently defined as Cheeger's energy gradient flow while the second can be defined as the entropy gradient flow. Let us emphasize that we obtain the existence of both gradient flows via a time-dependent JKO-scheme in a more general framework than the one chosen in Chapter 2.

4.1 Main Results

Gradient Flows on Time-dependent Metric Spaces and their Application to the Entropy on Time-dependent Probability Spaces

Before we treat entropy gradient flows on space of probability measures, we consider the more general case given by some energy functional $E\colon [0,T]\times X \to (-\infty,\infty]$ where X is a topological space endowed with a one-parameter family of complete separable geodesic metrics $(d_t)_t$ indexed by $t\in[0,T]$. We always assume that the map $t\mapsto \log d_t(x,y)$ is Lipschitz continuous, i.e. there exists a constant L such that

$$|\log(d_t(x,y)/d_s(x,y))| \le L|t-s|. \tag{148}$$

Additionally we impose a weak topology σ on X in the sense that d_t is sequentially σ-lower semicontinuous, such that each sequence $(x_n)\subset X$ with $\sup_{n,m} d_t(x_n,x_m)<\infty$ admits a σ-convergent subsequence.

We will say that an absolutely continuous curve is a *dynamic EDI-gradient flow* if

$$E_t(x_t) \le E_0(x_0) - \frac{1}{2}\int_0^t |\dot{x}_r|_r^2\,dr - \frac{1}{2}\int_0^t |\nabla_r E_r|^2(x_r)\,dr + \int_0^t (\partial_r E_r)(x_r)\,dr, \tag{149}$$

where $|\dot{x}_r|_r$ and $|\nabla_r E_r|$ denote the metric speed and the metric slope respectively with respect to the metric d_r. This formula represents a time-dependent version of the so-called *Energy Dissipation inequality*, in short EDI. Note that the dissipation is perturbed by the partial time-derivative of the functional along the curve. There are some technical issues in defining the time-dependent metric speed and the partial time-derivative if the functional is not supposed to be differentiable. We refer to Section 4.3 for the discussion.

In order to prove existence of dynamic gradient flows in the EDI sense we will adapt the minimizing movement scheme introduced by Jordan, Kinderlehrer and Otto in [30] in the following way. We fix a time step $h>0$ and an initial value $\bar{x}$. Recursively we define for every $n\in\mathbb{N}$ such that $nh\le T$ the minimizer

x_n^h by

$$x_0^h := \bar{x}, \qquad x_n^h := \arg\min_y \left(E_{nh}(y) + \frac{1}{2h} d_{nh}^2(x_{n-1}^h, y) \right). \tag{150}$$

Under the assumption that E is uniformly Lipschitz in t, sequentially σ-lower semicontinuous in x and uniformly bounded from below (for the precise assumptions see **A1, A2, A3a/A3b, A4** in Section 4.3) we show weak sequential compactness of the scheme, cf. Proposition 4.23. Since we want to show that the limit curve which we obtain is a dynamic gradient flow in the EDI sense we have to tighten up our assumptions on the functional. The first assumption involves the lower semicontinuity of the slope, which is well-known from the "static" theory of gradient flows, while the second requires upper semicontinuity of the partial time-derivative. Then we obtain the following result, cf. Proposition 4.24.

Theorem A. *We assume additionally to our standing assumptions that the partial time-derivative is upper semicontinuous, i.e. if* $x_n \stackrel{\sigma}{\rightharpoonup} x$ *as* $n \to \infty$ *then* $\limsup_{n\to\infty} \partial_t E_t(x_n) \leq \partial_t E_t(x)$, *and the squared slope is lower semicontinuous, i.e. if* $t_n \to t$ *and* $x_n \stackrel{\sigma}{\rightharpoonup} x$, *then* $|\nabla_t E_t|^2(x) \leq \liminf |\nabla_{t_n} E_t|^2(x_n)$. *Then for every* $\bar{x} \in Dom(E)$ *there exists a dynamic gradient flow in the EDI sense, i.e. a curve* $(x_t)_{0\leq t\leq T}$ *satisfying* (149) *and* $\lim_{t\searrow 0} x_t = \bar{x}$.

Similar to the static setting we ask when do we have equality in (149), which is also called *energy dissipation equality*, in short EDE. To answer this question we additionally assume that the functional is K-geodesically convex, i.e. there exists a constant $K \in \mathbb{R}$ such that for every $t \in [0, T]$ and for any pair of points x, y in the domain $Dom(E)$ there exists a d_t-geodesic $(\gamma_a)_{a\in[0,1]}$ connecting x and y such that for all $a \in [0, 1]$

$$E_t(\gamma_a) \leq (1-a)E_t(\gamma_0) + aE_t(\gamma_1) - K\frac{a(1-a)}{2} d_t^2(\gamma_0, \gamma_1).$$

Furthermore we have to impose an additional restraint on the partial time-derivative, i.e. for almost every $t \in [0, T]$

$$\liminf_{n\to\infty} \frac{E_{t_n}(x_n) - E_t(x_n)}{t_n - t} \geq \partial_t E_t(x), \text{ if } t_n \searrow t, \ x_n \xrightarrow{d} x \text{ as } n \to \infty.$$

We obtain the existence of a dynamic gradient flow in the EDE sense as a corollary of the weak chain rule, cf. Proposition 4.8, and Theorem A.

Theorem B. *Under combination of the previous assumptions, for every* $\bar{x} \in Dom(E)$ *there exists a curve* $(x_t)_{0\leq t\leq T}$ *satisfying*

$$E_t(x_t) + \frac{1}{2}\int_0^t |\dot{x}_r|_r^2 dr + \frac{1}{2}\int_0^t |\nabla_r E_r|^2(x_r) dr = E_0(\bar{x}) + \int_0^t (\partial_r E_r)(x_r) dr, \tag{151}$$

such that $\lim_{t\searrow 0} x_t = \bar{x}$.

Let us remark that uniqueness of the flow is not available on this level of generality. We will say a few more words on dynamic gradient flows in the EVI sense introduced in Chapter 2, which provides uniqueness. Under appropriate conditions we show in Proposition 4.12 that dynamic EVI implies dynamic EDE.

The analysis we described above is designed for the study of entropy gradient flows on time-dependent space of probability measures. Let us describe this application in more detail. We fix again a topological space X equipped with a one-parameter family d_t of geodesic separable complete metrics such that (148) holds and a one-parameter family of Borel measures such that $m_t = e^{-f_t} m$ for some probability measure m and suitable functions f_t satisfying

$$|f_t(x) - f_s(x)| \leq L^*|t - s| \qquad \forall x \in X. \tag{152}$$

Given two probability measures $\mu, \nu \in \mathcal{P}_2(X)$, where $\mathcal{P}_2(X)$ denotes the space of probability measures with finite second moments with respect to any metric d_t, we introduce for every $t \in [0, T]$ the L^2-Kantorovich distance defined by

$$W_t(\mu, \nu) = \inf \left\{ \int_{X \times X} d_t^2(x, y)\, d\pi(x, y) \big| \pi \text{ is a coupling of } \mu \text{ and } \nu \right\}^{1/2}.$$

Let us remark that we suppose that the space X is boundedly compact such that the weak topology on $\mathcal{P}_2(X)$ is adequate for our analysis. The relative entropy S_t on $\mathcal{P}(X)$ is given by

$$S_t(\mu) := \int_X \rho \log \rho \, dm_t,$$

provided that μ has a density ρ with respect to m_t. We assume that each static space (X, d_t, m_t) has Ricci curvature bounded below by some $K \in \mathbb{R}$, i.e. for each t and each μ, ν there exists a W_t-geodesic $(\rho_a)_{a \in [0,1]}$ connecting μ and ν such that

$$S_t(\rho_a) \leq (1 - a) S_t(\mu) + a S_t(\nu) - \frac{K}{2} a(1 - a) W_t^2(\mu, \nu). \tag{153}$$

This assumption is essential for the availability of the lower semicontinuity of the squared slope. In particular it is satisfied if the sequence (X, d_t, m_t) constitutes a super-Ricci flow in the sense of Sturm in [59]. We obtain the following result, see also Theorem 4.31 and Theorem 4.33.

Theorem C. *For every $\bar{\mu} \in \mathcal{P}_2(X)$ in the domain of the relative entropy there exists an absolutely continuous curve $(\mu_t)_{0 \leq t \leq T} \subset \mathcal{P}_2(X)$ satisfying*

$$S_t(\mu_t) + \frac{1}{2} \int_0^t |\dot{\mu}_r|_r^2 dr + \frac{1}{2} \int_0^t |\nabla_r S_r|^2(\mu_r) dr = S_0(\bar{\mu}) + \int_0^t (\partial_r S_r)(\mu_r) dr \tag{154}$$

and $\lim_{t \to 0} \mu_t = \bar{\mu}$. *Moreover this curve is unique.*

In the static metric measure space setting it is a well-known fact that the heat equation can be unambiguously defined as the gradient flow of the entropy or as the gradient flow of the Dirichlet energy. Here we prove a similar result for the forward dual heat flow from Chapter 2 under the assumption that each (X, d_t, m_t) satisfies a Riemannian curvature-dimension condition, cf. Theorem 4.45.

Theorem D. *Let $(\mu_t)_{0 \leq t \leq T}$ be a continuous curve in $\mathcal{P}_2(X)$. Then the following are equivalent:*

1. (μ_t) *is a dynamic gradient flow for the relative entropy in the EDE sense.*

2. (μ_t) *is given by* $\mu_t(dx) = \rho_t(x)m_t(dx)$*, where* ρ_t *is a solution to the adjoint heat equation*

$$\partial_t \rho_t(x) = \Delta_t \rho_t(x) - \rho_t(x)\partial_t f_t(x).$$

Gradient Flows on Time-dependent Hilbert Spaces and their Application to Cheeger's Energy on the Time-dependent Space of Square Integrable Functions

We start by considering a functional $E\colon [0,T] \times H \to [0,\infty]$ where H is a separable Hilbert space H endowed with a family of scalar products $(\langle \cdot, \cdot \rangle_t)$. We assume that (148) holds for the distances $||x-y||_t := \sqrt{\langle x-y, x-y \rangle_t}$ and that the map $x \mapsto E_t(x)$ is convex and lower semicontinuous. Moreover we require that $t \mapsto E_t(x)$ is Lipschitz continuous in the following way

$$\exists\, C_1 \, \forall x \, \forall s, t \in [0,T]: \quad |E_t(x) - E_s(x)| \leq C_1 E_t(x)|t-s|.$$

In this framework, we will choose a different approach to define a notion of dynamic gradient flows. We will say that an absolutely continuous curve $(x_t)_{0 \leq t \leq T}$ is a *dynamic gradient flow* if

$$\partial_t x_t \in -D_t^- E_t(x_t) \text{ for almost every } t \in (0,T),$$

where $D_t^- E_t(x)$ denotes the $\langle \cdot, \cdot \rangle_t$-subdifferential of E_t at some x in the domain $Dom(E_t)$, which is defined as the set of all $v \in H$ such that

$$E_t(y) - E_t(x) \geq \langle v, y-x \rangle_t \qquad \forall y \in H.$$

We show the following using a time-dependent minimizing movement scheme, see also Theorem 4.38.

Theorem E. *For every* $x \in Dom(E)$ *there exists a unique dynamic gradient flow* $(x_t)_{0 \leq t \leq T}$ *with* $\lim_{t \searrow 0} x_t = x$*.*

Let us remark that although the functional is convex in the space variable and Lipschitz in the time variable we do not have a minimal selection principle, even if we fix the metric. This is explained in Section 4.5. By this we mean that it is not necessary for the norm of the curve's derivative to be the element in the subdifferential with minimal norm, as it is the case in the static setting. But still this type of gradient flow implies dynamic EVI, cf. Proposition 4.35.

We apply the existence result to the framework described in the following. Let (X, d_t, m_t) be a family of complete separable metric measure spaces satisfying (148) and (152). For each $t \in [0,T]$ let us denote by $\mathrm{Ch}_t\colon L^2(X, m_t) \to [0,\infty]$ Cheeger's functional given by

$$\mathrm{Ch}_t(u) = \frac{1}{2} \inf \left\{ \liminf_{n \to \infty} \int_X (\mathrm{lip}_t u_n)^2 dm_t \middle| u_n \in \mathrm{Lip}(X), \int_X |u_n - u|^2 dm_t \to 0 \right\},$$

where $\mathrm{lip}_t u$ denotes the local slope given by

$$\mathrm{lip}_t u(x) = \limsup_{y \to x} \frac{|u(y) - u(x)|}{d_t(x,y)}.$$

By making use of the minimal relaxed gradient $|\nabla_t u|_*$ ([5, Definition 4.2]), this functional admits an integral representation

$$\mathrm{Ch}_t(u) = \frac{1}{2}\int_X |\nabla_t u|_*^2 dm_t,$$

set equal to $+\infty$ if u has no relaxed slope. We obtain the existence of a dynamic gradient flow for the family of convex and lower semicontinuous functional (Ch_t) as a direct consequence of Theorem E, cf. Theorem 4.47. Moreover we identify the gradient flow with the heat flow introduced in Chapter 2 for spaces which satisfy a Riemannian curvature-dimension condition, for the precise statement see Theorem 4.48.

Theorem F. *Let $\bar{u} \in Dom(\mathrm{Ch})$. Then there exists a unique dynamic gradient flow for (Ch_t) starting in $\bar{u}$, i.e. an absolutely continuous curve $(u_t)_{0\leq t\leq T} \subset Dom(\mathrm{Ch})$ solving*

$$\partial_t u_t \in -D_t^- \mathrm{Ch}_t(u_t) \quad \text{for a.e. } t \in (0,T) \tag{155}$$

and $\lim_{t\to 0} u_t = \bar{u}$.

Theorem G. *Let $(\tilde{u}_t)$ be the solution to the heat equation $\partial_t \tilde{u}_t = \Delta_t \tilde{u}_t$ on $(0,T)\times X$ starting in some $\bar{u} \in Dom(\mathrm{Ch})$. Then $(\tilde{u}_t)$ satisfies*

$$\partial_t \tilde{u}_t \in -D_t^- \mathrm{Ch}_t(\tilde{u}_t) \quad \text{for a.e. } t \in (0,T),$$

and can be constructed as the limit of a minimizing movement scheme. Conversely, let (u_t) be the dynamic gradient flow of Cheeger's functional (Ch_t) starting in $\bar{u} \in Dom(\mathrm{Ch})$. Then (u_t) solves the heat equation

$$\partial_t u_t = \Delta_t u_t \ \text{ on } (0,T)\times X.$$

In particular $u_t = \tilde{u}_t$ in $L^2(X)$ for every $t \in [0,T]$.

Structure of the chapter

Let us explain the structure of the chapter in the following. In Section 4.2 we briefly recall the concept of gradient flows in metric spaces. In Section 4.3 we introduce the notion of dynamic EDI-, EDE- and EVI(K,∞)-gradient flows on time-dependent metric spaces $(X,d_t)_{t\in[0,T]}$ satisfying (152) and show that EVI(K,∞) implies EDE. We show existence of dynamic EDI-gradient flows for a class of energy functionals $E\colon [0,T]\times X \to (-\infty,+\infty]$. Moreover we give sufficient conditions for the existence of EDE-gradient flows. In Section 4.4 we apply the results from Section 4.3 and prove existence and uniqueness of dynamic EDI-gradient flows in time-dependent metric measure spaces $(X,d_t,m_t)_{t\in[0,T]}$ for the time-dependent entropy functional $S\colon [0,T]\times\mathcal{P}(X) \to (-\infty,+\infty]$. In Section 4.5 we consider dynamic gradient flows in the form of (155) on time-dependent Hilbert spaces $(H,\langle\cdot,\cdot\rangle_t)_{t\in[0,T]}$. We prove existence and uniqueness of such gradient flows for a class of energy functionals $E\colon [0,T]\times H \to [0,+\infty]$. In Section 4.6 we recall the concept of heat equation on time-dependent metric measure spaces introduced in Chapter 2. We identify the dynamic EDI-gradient flow of the entropy with the forward adjoint heat flow. We apply the results from Section 4.5 and directly obtain existence and uniqueness of a dynamic gradient flow for Cheeger's functional and identify it with the heat flow.

4.2 Gradient Flows in Metric Spaces

We briefly recall the notions of gradient flows on metric spaces (X, d). A curve $x\colon [a, b] \to X$ is said to belong to $\mathrm{AC}^p([a, b]; X)$ for $1 \le p \le \infty$, if there exists $g \in L^p(a, b)$ such that

$$d(x_s, x_t) \le \int_s^t g(r)dr \qquad \text{for every } a \le s \le t \le b. \tag{156}$$

The *metric speed* of x, defined by

$$|\dot{x}_t| := \lim_{h\to 0} \frac{d(x_{t+h}, x_t)}{h},$$

exists for $\mathcal{L}^1$-a.e. $t \in (a, b)$, is of class $L^p(a, b)$ and is the smallest function such that (156) holds, see e.g. [2, Theorem 1.1.2].

Given $E\colon X \to (-\infty, +\infty]$ we define the *slope* $|\nabla E|(x)$ at x by

$$|\nabla E|(x) := \limsup_{y\to x} \frac{(E(x) - E(y))^+}{d(x, y)}.$$

We now are ready to give three possible definitions of gradient flows in a metric framework, cf. [4, 2].

Definition 4.1. *1. An absolutely continuous curve $(x_t) \subset X$ is a EDI-gradient flow if it satisfies the following Energy Dissipation Inequality*

$$E(x_s) + \frac{1}{2}\int_t^s |\dot{x}_r|^2 dr + \frac{1}{2}\int_t^s |\nabla E|^2(x_r)dr \le E(x_t) \quad \forall s \ge t. \tag{157}$$

2. An absolutely continuous curve $(x_t) \subset X$ is a EDE-gradient flow if it satisfies the following Energy Dissipation Equality

$$E(x_s) + \frac{1}{2}\int_t^s |\dot{x}_r|^2 dr + \frac{1}{2}\int_t^s |\nabla E|^2(x_r)dr = E(x_t) \quad \forall s \ge t. \tag{158}$$

3. An absolutely continuous curve $(x_t) \subset X$ is a EVI-gradient flow (with respect to $\lambda \in \mathbb{R}$) if it satisfies the following Evolution Variation Inequality

$$E(x_t) + \frac{1}{2}\frac{d}{dt}d^2(x_t, y) + \frac{\lambda}{2}d^2(x_t, y) \le E(y) \quad \textit{for a.e. } t \in [0, T], \forall y \in X. \tag{159}$$

If the underlying space is a Hilbert space and the energy functional is convex, these formulations are equivalent. Moreover we can characterize the flow in terms of the subdifferential by

$$\dot{x}_t \in -D^- E(x_t), \tag{160}$$

where $D^- E(x)$ consists of all $v \in X$ such that

$$E(x) + \langle v, y - x\rangle \le E(y) \quad \forall y \in X.$$

In this chapter we are interested in finding substitutions for formulations of the form (157) and (158), where the metric as well as the functional varies in time. A formulation in the sense of (159) has already been introduced in Chapter 2. Moreover, in the Hilbert space case, we study the time-dependent counterpart of relations of the form (160).

4.3 Dynamic Gradient Flows in Time-dependent Metric Spaces

In the sequel we fix a one-parameter family of complete geodesic metric spaces $(X, d_t)_t$ indexed by $t \in [0,T]$. We always assume that the map $t \to \log d_t(x,y)$ is Lipschitz continuous, i.e. there exists a constant L such that

$$|\log(d_t(x,y)/d_s(x,y))| \leq L|t-s|. \tag{161}$$

We give a simple example for this setting.

Example 1. *Let M be a smooth closed manifold equipped with a smooth family of Riemannian metrics (g_t) evolving under a Ricci flow, i.e.*

$$\frac{1}{2}\partial_t g_t = -\mathrm{Ric}(g_t),$$

where $\mathrm{Ric}(g)$ denotes the Ricci curvature. At least for short time intervals we have existence and uniqueness of such a flow (see e.g. Theorem 5.2.1 in [61]). Under the assumption that the curvature does not blow up ($|\mathrm{Ric}| \leq L$), we have metric equivalence

$$|\partial_t \log g_t(v,v)| \leq L.$$

This implies that (161) holds for the geodesic distances (d_t).

The Metric Speed

Definition 4.2. *Let $[0,T] \ni t \mapsto x_t \in X$ be a curve. We say that $(x_t) \in \mathrm{AC}^p([0,T];X)$, for $p \in [1,\infty]$, if for any (and thus for all) $t^* \in [0,T]$ there exists a function $g \in L^p(0,T)$ such that*

$$d_{t^*}(x_t, x_s) \leq \int_t^s g(r)dr \quad \forall 0 \leq t \leq s \leq T \in [0,T].$$

We define the length of a curve $x\colon [0,T] \to X$ to be

$$L_x(t) = \lim_{h\to 0} \sup \left\{ \sum_{i=1}^n d_{t_j}(x_{t_j}, x_{t_{j+1}}) : 0 = t_1 < \ldots < t_n = t, t_{j+1} - t_j \leq h \right\}.$$

It is a direct consequence of the definition of $L_x(t)$ that if $x_n \to x$ pointwise as $n \to \infty$ we have $L_x(t) \leq \liminf_{n\to\infty} L_{x_n}(t)$ for every $t \in [0,T]$.

Note that L_x is absolutely continuous as soon as x is and hence we may define the *momentaneous speed* of the curve as the derivative of its length.

$$|\dot{x}|_t := \dot{L}_x(t).$$

Lemma 4.3. *For any curve $x \in \mathrm{AC}^p(0,T)$ the function $t \mapsto |\dot{x}|_t$ is in $L^p(0,T)$, and for almost every $t \in (0,T)$*

$$|\dot{x}|_t = \lim_{s\to t} \frac{d_t(x_s, x_t)}{|s-t|} = |\dot{x}_t|_t.$$

Proof. If we show the second assertion the first assertion is an easy consequence of (161). Let x be an absolutely continuous curve and choose an arbitrary partition $s = t_1 < t_2 < \ldots < t_{N+1} = t$. Then we find

$$\begin{aligned} d_t(x_t, x_s) &\leq \sum_{i=1}^{N} d_{t_i}(x_{t_i}, x_{t_{i+1}}) + C \sum_{i=1}^{N} |t - t_i| d_t(x_{t_i}, x_{t_{i+1}}) \\ &\leq \sum_{i=1}^{N} d_{t_i}(x_{t_i}, x_{t_{i+1}}) + C|t - s| \int_s^t g(r)dr, \end{aligned}$$

where we used (161) and g is some L^p function. Hence we may estimate

$$d_t(x_t, x_s) \leq L(t) - L(s) + C|t - s| \int_s^t g(r)dr.$$

Dividing by $|t - s|$ and letting $s \to t$ we deduce

$$\limsup_{s \to t} \frac{d_t(x_t, x_s)}{|t - s|} \leq \dot{L}_x(t) \text{ for almost every } t.$$

We show the other inequality by contradiction. Fix $\eta > 0$ and consider the set of points

$$F = \left\{ t : \liminf_{s \to t} \left(\frac{d_t(x_s, x_t)}{|s - t|} - \frac{1}{|s - t|} \int_s^t \dot{L}_x(r)dr \right) < -\eta \right\}.$$

We assume that the Lebesgue outer measure $\mathcal{L}^*(F) > 0$. Fix $\delta > 0$ and cover the set F with intervals

$$\mathcal{F} := \bigcup_{t \in F} (t - \delta_t, t + \delta_t), \text{ where } \delta_t < \delta,$$

such that

$$d_t(x_t, x_s) < \int_s^t \dot{L}_x(r)dr - |t - s|\eta/2 \tag{162}$$

for all $t \in F$ and some $s \in (t - \delta_t, t + \delta_t)$. From the Besicovitch covering theorem [17, Theorem 5.8.1] it follows that there exists a constant N and a subcollection $\mathcal{F}_1, \cdots, \mathcal{F}_N$ each consisting of at most countably many disjoint intervals B such that

$$F \subset \bigcup_{i=1}^{N} \bigcup_{B \in \mathcal{F}_i} B.$$

Since the outer measure of F is strictly positive we can find a family $\mathcal{F}_j$ of at most countably many disjoint intervals denoted by $\mathcal{F}_j = \{(t_i - \delta_i, t_i + \delta_i), i \in I\}$ such that $\mathcal{L}^1(\bigcup_{B \in \mathcal{F}_j} B) \geq \frac{1}{N} \mathcal{L}^*(F) > 0$.

We define a curve $x^\delta \colon [0, T] \to X$ in the following way

$$x_t^\delta = \begin{cases} x_{t_i} & \text{if } t \in (t_i, t_i + \delta_i) \\ x_t & \text{else.} \end{cases}$$

Note that this curve is not continuous but still its length is finite. Further we observe that x_t^δ converges to x_t pointwise as δ goes to 0 and hence

$$\liminf_{\delta\to 0} L_{x^\delta}(T) \geq L_x(T).$$

It suffices to show that

$$L_{x^\delta}(T) \leq L_x(T)(1+L\delta) - \frac{\eta}{2}\mathcal{L}^1\left(\bigcup_{i\in I}(t_i, t_i+\delta_i)\right), \tag{163}$$

since then

$$L_x(T) \leq \liminf_{\delta\to 0} L_{x^\delta}(T) \leq L_x(T) - \frac{\eta}{4N}\mathcal{L}^*(F) < L_x(T),$$

which is clearly a contradiction. Hence for the outer measure it must hold $\mathcal{L}^*(F) = 0$ and therefore already $\mathcal{L}^1(F) = 0$. Since L_x is absolutely continuous we conclude

$$\liminf_{s\to t} \frac{d_t(x_s, x_t)}{|s-t|} \geq \dot{L}_x(t) \text{ for almost every } t \in [0,T].$$

It remains to show (163). Take a partition $(p_j)_{j=1}^m$ of $[0,T]$, with $0 < p_{j+1} - p_j \leq h$ and $h << \delta$. Consider the points near the boundary of $(t_i, t_i+\delta_i)$

$$j_i^{\leq} := \max\{j | p_j \leq t_i, 1 \leq j \leq m\}, \quad j_i^{\geq} = \min\{j | p_{j+1} \geq t_i + \delta_i, 1 \leq j \leq m\}.$$

Since x is absolutely continuous we can estimate

$$d_{p_{j_i^{\leq}}}(x^\delta_{p_{j_i^{\leq}}}, x^\delta_{p_{j_i^{\leq}+1}}) = d_{p_{j_i^{\leq}}}(x_{p_{j_i^{\leq}}}, x_{t_i}) \leq \int_{p_{j_i^{\leq}}}^{p_{j_i^{\leq}}+h} g(r)dr,$$

where $g \in L^p(0,T)$. Applying (162), (161) and again the absolute continuity we obtain

$$\begin{aligned}
&d_{p_{j_i^{\geq}}}(x^\delta_{p_{j_i^{\geq}}}, x^\delta_{p_{j_i^{\geq}+1}}) = d_{p_{j_i^{\geq}}}(x_{t_i}, x_{p_{j_i^{\geq}+1}}) \\
&\leq d_{p_{j_i^{\geq}}}(x_{t_i}, x_{t_i+\delta_i}) + d_{p_{j_i^{\geq}}}(x_{t_i+\delta_i}, x_{p_{j_i^{\geq}+1}}) \\
&\leq d_{t_i}(x_{t_i}, x_{t_i+\delta_i})(1+L\delta_i) + \int_{t_i+\delta_i}^{t_i+\delta_i+h} g(r)dr \\
&\leq \int_{t_i}^{t_i+\delta_i} (\dot{L}_x(r) - \eta/2)dr(1+L\delta_i) + \int_{t_i+\delta_i}^{t_i+\delta_i+h} g(r)dr.
\end{aligned}$$

Taking the supremum over all partitions (p_j) and letting $h \to 0$ we can estimate the length of the curve x^δ

$$\begin{aligned}
L_{x^\delta}(T) &\leq \int_{(0,T)\setminus\bigcup_i(t_i,t_i+\delta_i)} \dot{L}_x(r)dr + \sum_i \int_{t_i}^{t_i+\delta_i} (\dot{L}_x(r) - \eta/2)dr(1+C^*\delta_i) \\
&\leq \int_{(0,T)} \dot{L}_x(r)dr(1+L\delta) - \eta/2\mathcal{L}^1\left(\bigcup_i (t_i, t_i+\delta_i)\right),
\end{aligned}$$

which proves (163).

□

The Slope

Definition 4.4. *Let $E\colon [0,T]\times X \to (-\infty,+\infty]$ and $s,t\in[0,T]$, $x\in X$ with $E_t(x)<\infty$. Then the* slope $|\nabla_s E_t|(x)$ *of E_t with respect to d_s is given by*

$$|\nabla_s E_t|(x) = \limsup_{y\to x} \frac{[E_t(x)-E_t(y)]^+}{d_s(x,y)} = \limsup_{y\to x} \max\left\{\frac{E_t(x)-E_t(y)}{d_s(x,y)}, 0\right\}.$$

We mainly deal with the case $t=s$ in the definition of the slope. We estimate the deviation of the d_t slope from the d_s slope in the following lemma.

Lemma 4.5. *Let $s,t\in[0,T]$ and $x\in X$ such that $E_t(x)<\infty$ and $|\nabla_s E_t|(x)<\infty$. Then we have*

$$||\nabla_t E_t|(x) - |\nabla_s E_t|(x)| \le L|t-s||\nabla_s E_t|(x).$$

Proof. This follows from (161) and $\log r \le r-1$ and $\log(r^{-1}) \ge 1-r$ respectively. $\square$

4.3.1 Dynamic EDI- and EDE-Gradient Flows

Let us first motivate the definition of dynamic EDI-gradient flows by considering a Hilbert space X endowed with a family of scalar products $(\langle\cdot,\cdot\rangle_t)$ depending smoothly on t. Let $E_t\colon X\to\mathbb{R}$ be a $\mathcal{C}^1$ functional also smoothly depending on time. In this setting we understand a gradient flow as a curve solving

$$\dot{x}_t = -\nabla_t E_t(x_t). \tag{164}$$

Let us observe that (164) can be rewritten as

$$\frac{d}{dt}E_t(x_t) \le -\frac{1}{2}|\nabla_t E_t|_t^2(x_t) - \frac{1}{2}|\dot{x}_t|_t^2 + (\partial_t E_t)(x_t), \tag{165}$$

where $(\partial_t E_t)(x_t)$ stands for $\frac{d}{ds}E_s(x_t)\Big|_{s=t}$. Indeed, along any differentiable curve it holds

$$\begin{aligned}\frac{d}{dt}E_t(x_t) &= \frac{d}{ds}E_s(x_t)\Big|_{s=t} + \langle\nabla_t E_t(x_t), \dot{x}_t\rangle_t \\ &\ge \frac{d}{ds}E_s(x_t)\Big|_{s=t} - \frac{1}{2}|\nabla_t E_t|_t^2(x_t) - \frac{1}{2}|\dot{x}_t|_t^2,\end{aligned}$$

and we have equality if and only if (164) holds. The functional's dependence on the time variable leads to a "drift" of the gradient flow, i.e. in some sense the gradient flow does not follow the steepest descent. To illustrate this we give an example about the asymptotic behavior.

Example 2. *Let $X=\mathbb{R}$ and $d_t=|x-y|$ for $t\in[0,\infty)$ and $x,y\in\mathbb{R}$. We consider the energy $E_t(x)=(x-t)^2$ and the curve $x_t=\frac{1}{2}e^{-2t}+t-\frac{1}{2}$. Note that*

$$\dot{x}_t = -e^{-2t}+1 = -2(x_t-t) = -\partial_x E_t(x),$$

and hence (x_t) is a gradient flow. A well-known fact in the theory of gradient flows is that for strictly convex functionals the gradient flow converges to the *minimum of the functional as $t \to \infty$, see e.g. [2, Theorem 3.1(v)]. In our case the minima depend on time and are given by $x_t^{\min} = t$. Hence*

$$|x_t - x_t^{\min}| = \frac{|e^{-2t} - 1|}{2},$$

which obviously does not converge to 0 as $t \to \infty$.

Let us now come back to our original family of complete, separable geodesic, metric spaces (X, d_t) such that (161) holds true. We call a measurable functional E on $[0,T] \times X$ *admissible* if it satisfies the following assumptions.

A1 The domain $Dom(E_t) := \{x \in X | E_t(x) < \infty\}$ is time-independent and nonempty.

A2 For each $t \in [0,T]$, $x \mapsto E_t(x)$ is uniformly bounded from below.

A3a For each $t \in [0,T]$, $x \mapsto E_t(x)$ is lower semicontinuous.

A4 The map $t \mapsto E_t(x)$ is uniformly Lipschitz continuous, i.e. there exists a constant L^* such that

$$|E_t(x) - E_s(x)| \leq L^* |t - s| \quad \forall t, s \in [0,T], x \in Dom(E),$$

and the set of differentiability points of the map $t \mapsto E_t(x)$ can be chosen regardless of $x \in X$ as soon as $x \in Dom(E)$.

Note that the Lipschitz continuity of the map $t \mapsto E_t(x)$ provides a.e. differentiability in t for every fixed x. But this is not enough to get a meaningful expression in (165), since we may have that for some absolutely continuous curve (x_t), $t \mapsto E_t$ is not differentiable at x_t for every t. To circumvent this problem we suppose that the set of differentiability points can be chosen independent of x, cf. [25, 53]. To illustrate this we give the following example, which has also been discussed in [53].

Example 3. *Let $X = \mathbb{R}$ and $d_t(x,y) = |x - y|$ for every $t \in [0,T]$, $x, y \in \mathbb{R}$. Consider the following energy functional $E\colon [0,T] \times \mathbb{R} \to [0,\infty)$ given by*

$$E_t(x) = |x - t|.$$

Then the map $t \mapsto E_t(x)$ is clearly Lipschitz continuous with well-defined derivative $\partial_t E_t(x)$ as long as $t \in [0,T] \setminus \{x\}$. If we choose the curve $(x_t)_{t \in [0,T]} \in \mathcal{C}^\infty([0,T])$ by setting $x_t = t$, the map $s \mapsto E_s(x_t)$ is not differentiable at any $t \in [0,T]$. Indeed, for every $t \in [0,T]$ the right derivative $\partial_s E_s(x_t)_{|s=t+}$ equals 1, while the left derivative $\partial_s E_s(x_t)_{|s=t-}$ equals -1.

Definition 4.6. *We call a locally absolutely continuous curve $(x_t)_{0 \leq t \leq T}$ a dynamic EDI-gradient flow for an admissible functional $E\colon [0,T] \times X \to (-\infty, \infty]$, if for every $t \in [0,T]$*

$$E_t(x_t) + \frac{1}{2}\int_0^t |\dot{x}|_r^2 dr + \frac{1}{2}\int_0^t |\nabla_r E_r|^2(x_r) dr \leq E_0(x_0) + \int_0^t (\partial_r E_r)(x_r) dr, \quad (166)$$

where we used the shorthand notation $(\partial_t E_t)(x_t) = \frac{d}{dr}E_r(x_t)|_{r=t}$. *We call a locally absolute continuous curve* $x\colon [0,T] \to X$ *a dynamic EDE-gradient flow for an admissible functional* $E\colon [0,T]\times X \to (-\infty,\infty]$*, if for every* $t\in[0,T]$

$$E_t(x_t) + \frac{1}{2}\int_0^t |\dot{x}|_r^2 dr + \frac{1}{2}\int_0^t |\nabla_r E_r|^2(x_r) dr = E_0(x_0) + \int_0^t (\partial_r E_r)(x_r) dr, \quad (167)$$

Clearly, (167) implies (166). In the following we want to give sufficient conditions for the other implication.

Definition 4.7. *We say that the above mentioned functional* E *is* K*-convex for* $K\in\mathbb{R}$*, if for every* $t\in[0,T]$ *and for any pair of points* $x,y\in Dom(E)$ *there exists a* d_t*-geodesic* $(\gamma_a)_{a\in[0,1]}$ *connecting* x *and* y *such that for all* $a\in[0,1]$

$$E_t(\gamma_a) \le (1-a)E_t(\gamma_0) + aE_t(\gamma_1) - K\frac{a(1-a)}{2}d_t^2(\gamma_0,\gamma_1). \qquad (168)$$

The convexity assumption allows us to reformulate the slope

$$|\nabla_t E_t|(x) = \sup_{y\neq x}\left[\frac{E_t(x)-E_t(y)}{d_t(x,y)} + \frac{K^-}{2}d_t(x,y)\right]^+, \qquad (169)$$

with $K^- := \max\{0,-K\}$, cf. [4, Theorem 2.4.9].

The next proposition can be thought of as a weak chain rule in the sense of [2, Proposition 3.19]. The convexity of the functional plays an important role in the proof of this result. Unlike in the static case we additionally have to impose a condition on the difference quotients of the functionals, cf. [25, Theorem 5.4].

Proposition 4.8. *Let* $E\colon [0,T]\times X \to (-\infty,+\infty]$ *be a* K*-convex admissible functional. Moreover assume that for almost every* $t\in[0,T]$

$$\liminf_{n\to\infty}\frac{E_{t_n}(x_n)-E_t(x_n)}{t_n - t} \ge \partial_t E_t(x), \text{ if } t_n \searrow t,\ x_n \xrightarrow{d} x \text{ as } n\to\infty. \qquad (170)$$

Then for every locally absolutely continuous curve $(x_t)\subset Dom(E)$*, the function* $t\mapsto E_t(x_t)$ *is absolutely continuous and it holds*

$$E_t(x_t) - E_s(x_s) \ge \int_s^t (\partial_r E_r)(x_r)\,dr - \int_s^t |\dot{x}|_r |\nabla_r E_r|(x_r)\,dr, \quad s<t. \qquad (171)$$

In particular, if (x_t) *is a dynamic EDI-gradient flow, it is a dynamic EDE-gradient flow as well.*

Proof. In view of [4, Lemma 1.1.4(a)] we can find an increasing and absolutely continuous map $\boldsymbol{s}\colon [0,T]\to[0,L]$, whose inverse $\boldsymbol{t}$ is Lipschitz. The reparametrization $\hat{x}_{\boldsymbol{s}(t)} := x(t)$ satisfies $|\dot{\hat{x}}_s|_{t^*} \le 1$ for almost every $s\in[0,L]$ with respect to some fixed metric d_{t^*}. Notice that it is sufficient to prove that $s\mapsto E_{\boldsymbol{t}(s)}(\hat{x}_s) =: \varphi(s)$ is absolutely continuous, as then $E_t(x_t) = E_t(\hat{x}_{\boldsymbol{s}(t)})$ is absolutely continuous and for almost every $t\in[0,T]$

$$\begin{aligned}
\frac{d}{dt}E_t(x_t) &= \lim_{h\to 0}\frac{E_{t+h}(x_{t+h})-E_t(x_t)}{h}\\
&\ge \liminf_{h\to 0}\frac{E_{t+h}(x_{t+h})-E_t(x_{t+h})}{h} + \liminf_{h\to 0}\frac{E_t(x_{t+h})-E_t(x_t)}{d_t(x_{t+h},x_t)}\frac{d_t(x_{t+h},x_t)}{h}\\
&\ge \partial_t E_t(x_t) - \limsup_{h\to 0}\frac{[E_t(x_t)-E_t(x_{t+h})]^+}{d_t(x_{t+h},x_t)}\frac{d_t(x_{t+h},x_t)}{h}\\
&\ge \partial_t E_t(x_t) - |\nabla_t E_t|(x_t)|\dot{x}|_t,
\end{aligned}$$

where we used (170) in the third inequality. After integration we obtain (171).

In view of the convexity of E we may use the representation formula of the slope (169) and write using $a^+ \le (a+b)^+ + b^-$ and the Lipschitz property of the functional

$$\begin{aligned}\varphi(s_1) - \varphi(s_0) &\le |\nabla_{\boldsymbol{t}(s_1)} E_{\boldsymbol{t}(s_1)}|(\hat{x}_{s_1}) d_{\boldsymbol{t}(s_1)}(\hat{x}_{s_1}, \hat{x}_{s_0}) \\ &\quad + \frac{K^-}{2} d^2_{\boldsymbol{t}(s_1)}(\bar{x}_{s_1}, \hat{x}_{s_0}) + L^*|s_1 - s_2| \\ &\le \Big(|\nabla_{\boldsymbol{t}(s_1)} E_{\boldsymbol{t}(s_1)}|(\hat{x}_{s_1}) + \frac{K^-}{2} D \Big) e^C |s_1 - s_0| + L^*|s_1 - s_0|,\end{aligned} \tag{172}$$

where D is the finite diameter of the image $\{\hat{x}_s\}_s$ with respect to $d_{\boldsymbol{t}}$. Changing the roles of s_0 and s_1 yields

$$\begin{aligned}&|\varphi(s_1) - \varphi(s_0)| \\ &\le \Big(|\nabla_{\boldsymbol{t}(s_1)} E_{\boldsymbol{t}(s_1)}|(\hat{x}_{s_1}) + |\nabla_{\boldsymbol{t}(s_0)} E_{\boldsymbol{t}(s_0)}|(\hat{x}_{s_0}) + \frac{K^-}{2} D \Big) e^C |t - s| + L^*|t - s|.\end{aligned}$$

Applying [4, Lemma 1.2.6] we conclude that the map $s \mapsto \varphi(s)$ is in the Sobolev space $W^{1,1}(0, L)$. To prove absolute continuity we simply check that it coincides with its continuous representative. We already know that $s \mapsto \varphi(s)$ is lower semicontinuous and therefore continuity follows if we show

$$\limsup_{\varepsilon \searrow 0} \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} \varphi(s + r) dr \le \varphi(s) \quad \forall s \in (0, L).$$

This can be seen by applying (172) once more and we get

$$\begin{aligned}&\limsup_{\varepsilon \searrow 0} \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} \varphi(s + r) - \varphi(s)\, dr \\ &\le \limsup_{\varepsilon \searrow 0} \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} \Big(|\nabla_{\boldsymbol{t}(s+r)} E_{\boldsymbol{t}(s+r)}|(\hat{x}_{s+r}) + \frac{K^-}{2} D \Big) e^C |r| + L^*|r| dr \\ &\le \limsup_{\varepsilon \searrow 0} \frac{1}{2} \int_{-\varepsilon}^{\varepsilon} \Big(|\nabla_{\boldsymbol{t}(s+r)} E_{\boldsymbol{t}(s+r)}|(\hat{x}_{s+r}) + \frac{K^-}{2} D) e^C + L^* dr = 0.\end{aligned}$$

□

4.3.2 Dynamic EVI(K, ∞)-Gradient Flows

Let us recall the dynamic version of EVI(K,∞)-gradient flows introduced in Chapter 2.

Definition 4.9. *For $s, t \in [0, T]$ and an absolutely continuous curve $(x_a)_{a \in [0,1]}$, we define the action*

$$\begin{aligned}\mathcal{A}_{s,t}(x) = \limsup_{h \to 0} \Big\{ &\sum_{i=1}^{n} (a_i - a_{i-1})^{-1} d^2_{\vartheta(a_{i-1})}(x_{a_{i-1}}, x_{a_i}) | \\ &0 = a_0 < \cdots < a_n = 1, a_i - a_{i-1} \le h \Big\},\end{aligned}$$

where $\vartheta \colon [0, 1] \to [0, \infty)$ denotes the linear interpolation with $\vartheta(0) = s$ and $\vartheta(1) = t$. For two points $x^0, x^1 \in X$ we define

$$d^2_{s,t}(x^0, x^1) = \inf\{\mathcal{A}_{s,t}(x) | x \colon [0, 1] \to X \textit{ absolutely continuous}, x_0 = x^0, x_1 = x^1\}.$$

Note that using the definition of the metric speed we obtain for the action the more intuitive expression, cf. Proposition 2.72,

$$\mathcal{A}_{s,t}(x) = \int_0^1 |\dot{x}_a|^2_{\vartheta(a)}\, da.$$

We understand $d_{s,t}(x,y)$ as "dynamic distance" between the points x and y but, of course, strictly speaking it does not define a metric, since e.g. $d_{s,t}(x,y) \neq d_{s,t}(y,x)$ as soon as $x \neq y$. However, it clearly holds $d_{s,t}(x,y) = d_{t,s}(y,x)$, $d_{t,t}(x,y) = d_t(x,y)$ and $d_{s,t}(x,x) = 0$.

We will use the following notation: $\partial_t^+ u(t) := \limsup_{s\to t} \frac{u(t)-u(s)}{t-s}$.

Definition 4.10. *Let $E : [0,T] \times X \to (-\infty, \infty]$ be a lower semicontinuous functional in X. An absolutely continuous curve $(x_t)_{0\le t\le T}$ will be called* dynamic EVI(K,∞)-gradient flow *for E if for all $t \in (0,T)$ and all $y \in Dom(E_t)$*

$$\frac{1}{2}\partial_s^+ d^2_{s,t}(x_s,y)\Big|_{s=t} + \frac{K}{2} d_t^2(x_t,y) \ \le E_t(y) - E_t(x_t).$$

We say that the gradient flow $(x_t)_{0\le t\le T}$ starts *in $x' \in X$ if $\lim_{t\searrow 0} x_t = x'$.*

We show uniqueness of dynamic EVI(K,∞) flows by proving a contraction type estimate. This estimate involves the logarithmic Lipschitz control L from (161). For an estimate without this control see Theorem 2.77.

Lemma 4.11. *The following holds true.*

1. *Suppose that (x_t) is a EVI(K,∞)-gradient flow. Then for every $t \in (0,T)$*
$$\frac{1}{2}\partial_s^+ d_t^2(x_s,y)_{|s=t} \le E_t(y) - E_t(x_t) + (L - \frac{K}{2}) d_t^2(x_t,y). \tag{173}$$

2. *There exists at most one EVI(K,∞)-gradient flow starting in x'. More precisely the following holds: Let (x_t) and (y_t) be two EVI(K,∞)-gradient flows. Then for all $s < t$*
$$d_t(x_t,y_t) \le e^{(3L-K)(t-s)} d_s(x_s,y_s). \tag{174}$$

Proof. To show the first assertion note that with $d_{t,s}(y,x_s) = d_{s,t}(x_s,y)$

$$\begin{aligned}
\partial_s^+ d^2_{t,s}(y,x_s)_{s=t+} &:= \limsup_{s\searrow t} \frac{d^2_{t,s}(y,x_s) - d_t^2(y,x_t)}{s-t} \\
&\ge \limsup_{s\searrow t} \frac{e^{-2L(s-t)} d_t^2(y,x_s) - d_t^2(y,x_t)}{s-t} \\
&\ge \limsup_{s\searrow t} \Big\{ \frac{d_t^2(y,x_s) - d_t^2(y,x_t)}{s-t} + \frac{(e^{-2L(s-t)} - 1)}{s-t} d_t^2(y,x_s) \Big\} \\
&= \partial_s^+ d_t^2(y,x_s)_{|s=t+} + \lim_{s\searrow t} \frac{(e^{-2L(s-t)}-1)}{s-t} d_t^2(y,x_s) \\
&= \partial_s^+ d_t^2(y,x_s)_{|s=t+} - 2L d_t^2(y,x_t),
\end{aligned}$$

where the first inequality is due to the logarithmic Lipschitz continuity (161), and the second equality follows from the absolute continuity of (x_t). The same argument holds for $\partial_s^+ d_{s,t}^2(x_s, y)_{s=t-} := \limsup_{s\nearrow t} \frac{d_{s,t}^2(x_s,y)-d_t^2(x_t,y)}{s-t}$ replacing $\partial_s^+ d_t^2(y, x_s)_{|s=t+}$ by $\partial_s^+ d_t^2(y, x_s)_{|s=t-}$, and hence from the EVI$(K,\infty)$ inequality we deduce

$$\frac{1}{2}\partial_s^+ d_t^2(x_s, y)_{|s=t} \leq E_t(y) - E_t(x_t) + (L - \frac{K}{2})d_t^2(x_t, y).$$

In order to show the second assertion, let $(x_t), (y_t)$ be two EVI(K,∞) gradient flows. Observe that from the absolute continuity of (x_t) and (y_t) it follows that the map $t \mapsto d_t^2(x_t, y_t)$ is absolutely continuous as well. This can be seen by applying triangle inequality and (161). Hence we may write for a.e. $t \in (0, T)$

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}d_t^2(x_t, y_t) \leq& \frac{1}{2}\limsup_{s\nearrow t} \frac{d_t^2(x_t, y_t) - d_t^2(x_s, y_t)}{t-s} \\ &+\frac{1}{2}\limsup_{s\nearrow t} \frac{d_t^2(x_s, y_t) - d_s^2(x_s, y_t)}{t-s} \\ &+\frac{1}{2}\limsup_{s\searrow t} \frac{d_t^2(x_t, y_s) - d_t^2(x_t, y_t)}{s-t},\end{aligned} \tag{175}$$

where we used an adaption of [4, Lemma 4.3.4]. Applying (173) and (161) we obtain for a.e. $t \in (0, T)$

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}d_t^2(x_t, y_t) \leq& E_t(y_t) - E_t(x_t) + (L - \frac{K}{2})d_t^2(x_t, y_t) \\ &+ Ld_t^2(x_t, y_t) \\ &+ E_t(x_t) - E_t(y_t) + (L - \frac{K}{2})d_t^2(x_t, y_t) \\ =&(3L - K)d_t^2(x_t, y_t).\end{aligned}$$

We conclude from Gronwall's inequality for a.e. $t > s$

$$d_t^2(x_t, y_t) \leq e^{(6L-2K)(t-s)}d_s^2(x_s, y_s).$$

From the continuity of $t \mapsto d_t(x_t, y_t)$ we obtain that the estimate holds for every $t > s$ and in particular we have uniqueness.

□

In this general framework it is possible to produce dynamic EVI-gradient flows which are not dynamic EDI-gradient flows as we see in the next example.

Example 4. *Let $X = \mathbb{R}$ and $d_t(x, y) = |x - y|$ for every $t \in [0, T]$, $x, y \in X$. As already seen in Example 3, the energy functional $E_t(x) = |x - t|$ is not differentiable at $x_t = t$ for any $t \in [0, T]$. Hence it is not a EDI-gradient flow in the sense of Definition 4.6. But it immediately follows from*

$$\frac{1}{2}\partial_t|x_t - y|^2 = (t - y) \leq |y - t| = E_t(y) - E_t(x_t), \qquad \forall y \in X,$$

that (x_t) is a EVI$(0,\infty)$-gradient flow.

We can exclude such behavior if we restrict ourselves to admissible functionals.

Proposition 4.12. *Let $E\colon [0,T]\times X \to \mathbb{R}$ be an admissible functional, i.e. satisfying the assumptions **A1, A2, A3** and **A4** from the previous section. Let (x_t) be a dynamic EVI(K,∞)-gradient flow for E such that $(x_t) \in AC^2_{loc}([0,T];X)$ and $t \mapsto E_t(x_t)$ is absolutely continuous. Then it is a dynamic EDE-gradient flow as well.*

Proof. First note that for a.e. t

$$\frac{1}{2}\partial_s^+ d_t^2(x_s,y)_{|s=t} \geq -|\dot{x}_t|_t d_t(x_t,y). \tag{176}$$

Since E is admissible and $t \to E_t(x_t)$ is supposed to be absolutely continuous it holds for a.e. t

$$\begin{aligned}
\frac{d}{dt}E_t(x_t) =&(\partial_t E_t)(x_t) + \liminf_{h\to 0}\frac{E_t(x_{t+h})-E_t(x_t)}{h}\\
=&(\partial_t E_t)(x_t) + \liminf_{h\to 0}\frac{E_t(x_{t+h})-E_t(x_t)}{d_t(x_{t+h},x_t)}\frac{d_t(x_{t+h},x_t)}{h}\\
\geq&(\partial_t E_t)(x_t) - \limsup_{h\to 0}\frac{E_t(x_t)-E_t(x_{t+h})}{d_t(x_{t+h},x_t)}\frac{d_t(x_{t+h},x_t)}{h}\\
\geq&(\partial_t E_t)(x_t) - \frac{1}{2}|\nabla_t E_t|^2(x_t) - \frac{1}{2}|\dot{x}_t|_t^2.
\end{aligned} \tag{177}$$

To show the converse inequality recall that $t \mapsto d_t^2(x_t,y)$ is absolutely continuous. Hence, applying the same calculation as in (175) to the constant curve $y_t \equiv y$, we can write for every $t \in [0,T-h]$ and every y

$$\begin{aligned}
\frac{1}{2}d_{t+h}^2(x_{t+h},y) - \frac{1}{2}d_t^2(x_t,y) &= \frac{1}{2}\int_t^{t+h}\frac{d}{ds}d_s^2(x_s,y)ds\\
&\leq \int_t^{t+h} E_s(y) - E_s(x_s) + (2L - \frac{K}{2})d_s^2(x_s,y)ds.
\end{aligned}$$

We set $y = x_t$ and find

$$\begin{aligned}
\frac{1}{2}d_{t+h}^2(x_{t+h},x_t) \leq& h\int_0^1 E_{t+hr}(x_t) - E_{t+hr}(x_{t+hr})dr\\
&+(2L-\frac{K}{2})\int_t^{t+h} d_r^2(x_r,x_t)dr.
\end{aligned}$$

Again by (161) and the 2-absolute continuity of (x_t) we obtain for some function $g \in L^2_{loc}[0,T]$

$$\frac{1}{2}d_t^2(x_{t+h},x_t) \leq e^{2Lh}\left[h\int_0^1 E_{t+hr}(x_t) - E_{t+hr}(x_{t+hr})dr + |2L-\frac{K}{2}|h^2\int_t^{t+h} g_u^2\,du\right].$$

Dividing by h^2 and letting $h \searrow 0$, dominated convergence yields

$$\begin{aligned}
\frac{1}{2}|\dot{x}_t|_t^2 &\leq \int_0^1 \lim_{h\searrow 0}\frac{E_t(x_t)-E_{t+hr}(x_{t+hr})}{h} + \frac{E_{t+hr}(x_t)-E_t(x_t)}{h}dr\\
&= -\frac{1}{2}\frac{d}{dt}E_t(x_t) + \frac{1}{2}(\partial_t E_t)(x_t),
\end{aligned} \tag{178}$$

for a.e. $t \in (0,T)$. Concerning the slope of E we find that using (173) and (176)

$$\begin{aligned} |\nabla_t E_t|(x_t) &= \limsup_{y \to x_t} \frac{[E_t(x_t) - E_t(y)]^+}{d_t(x,y)} \\ &\leq \limsup_{y \to x_t} \frac{\left[-\partial_s^+ d_t^2(x_s,y)_{s=t} + (2L-K) d_t^2(x_t,y)\right]^+}{2d_t(x_t,y)} \\ &\leq \limsup_{y \to x_t} \frac{\left[2|\dot{x}_t|_t d_t(x_t,y) + (2L-K) d_t^2(x_t,y)\right]^+}{2d_t(x_t,y)} \leq |\dot{x}_t|_t, \end{aligned} \tag{179}$$

for almost every t. Combining (178) and (179) we conclude

$$\begin{aligned} \frac{d}{dt} E_t(x_t) &\leq (\partial_t E_t)(x_t) - |\dot{x}_t|_t^2 \\ &\leq (\partial_t E_t)(x_t) - \frac{|\dot{x}_t|_t^2}{2} - \frac{|\nabla_t E_t|^2(x_t)}{2}. \end{aligned} \tag{180}$$

We obtain (167) from (177) and (180) after integrating on the interval $(0,t)$. $\square$

4.3.3 Existence of Dynamic EDI-Gradient Flows

We are interested in the following problem.

Problem 1. *Given a function $E\colon [0,T] \times X \to (-\infty, +\infty]$, and an initial value $\bar{x} \in Dom(E)$, find an EDI-gradient flow (x_t) for E.*

Under suitable topological assumptions we will find a gradient flow for a certain class of energy functionals using the minimizing movement scheme, which we describe in the subsequent sections, cf. [4].

Topological assumptions

We additionally impose a topology σ on X such that σ is weaker than the topology induced by (d_t) and d_t is sequentially σ-lower semicontinuous, i.e.

$$\text{if } x_n \overset{\sigma}{\rightharpoonup} x \text{ and } y_n \overset{\sigma}{\rightharpoonup} y, \text{ then } \liminf_{n\to\infty} d_t(x_n,y_n) \geq d_t(x,y) \text{ for every } t \in [0,T].$$

Let $E\colon [0,T] \times X \to (-\infty, \infty]$ be a functional satisfying **A1**, **A2**, and **A4**. We will extend our assumptions by the following.

A5 If $(x_n) \subset X$ with $\sup_{n,m} d_t(x_n,x_m) < \infty$, then (x_n) admits a σ-convergent subsequence.

A3b For each $t \in [0,T]$, $x \mapsto E_t(x)$ is sequentially σ-lower semicontinuous.

Approximation

We fix a time step $h > 0$ and subdivide the interval $[0,T]$ into the partition

$$\mathcal{P}_h := \{t_0 = 0 < t_1 < \cdots < t_{N-1} < T \leq t_N\}, \qquad t_n = nh, N \in \mathbb{N}.$$

For $0 \leq t \leq T$ we define the piecewise constant interpolants $\overline{h}(t)$ and $\underline{h}(t)$ associated with the partition $\mathcal{P}_h$ in the following way;

$$\overline{h}(0) = 0 = \underline{h}(0), \text{ and for } t \in (t_{n-1}, t_n] \quad \overline{h}(t) = t_n, \quad \underline{h}(t) = t_{n-1}. \tag{181}$$

The definition implies that $\overline{h}(t) \searrow t$ and $\underline{h}(t) \nearrow t$ if $h \searrow 0$.

For a given initial value $\bar{x}$ we recursively define a sequence (x_n^h) of minimizers by

$$x_0^h := \bar{x}, \qquad x_n^h := \arg\min_x \left\{ E_{t_n}(x) + \frac{1}{2h} d_{t_n}^2(x, x_{n-1}^h) \right\} \tag{182}$$

Proposition 4.13. *For every $\bar{x} \in Dom(E)$ and $h > 0$ there exists a solution to the minimization problem* (182)*.*

Proof. Existence follows by the direct method of calculus. Define

$$\phi(h, \bar{x}, t; \cdot) := E_t(\cdot) + \frac{1}{2h} d_t^2(\bar{x}, \cdot).$$

Since E is uniformly bounded from below we may take a minimizing sequence $(x_n)_{n\in\mathbb{N}}$ such that $d_t^2(x_n, \bar{x})$ remains bounded uniformly in n. From the triangle inequality we deduce that $\sup_{n,m} d_t(x_n, x_m) < \infty$. Hence **A5** guarantees existence of a σ-convergent subsequence x_{n_k}. The weak limit point $x \in Dom(E)$ is a minimizer of $\phi(h, \bar{x}, t; \cdot)$, which is due to the σ-lower semicontinuity of the distance and the functional. □

Definition 4.14. *Fix $h > 0$ and let $s \in [0, T-h]$. For $0 < r < T - s$ define*

$$J_{s,r}(y) := \min_x \left\{ E_{s+r}(x) + \frac{1}{2r} d_{s+h}^2(x, y) \right\}, \tag{183}$$

$$A_{s,r}(y) := \arg\min_x \left\{ E_{s+r}(x) + \frac{1}{2r} d_{s+h}^2(x, y) \right\}. \tag{184}$$

Lemma 4.15. *For $x_r \in A_{s,r}(y)$ we have*

$$|\nabla_{s+h} E_{s+r}|(x_r) \le \frac{1}{r} d_{s+h}(x_r, y)$$

and for $0 < r_1 < r_2 < T - s$

$$d_{s+h}^2(x_{r_1}, y) \le d_{s+h}^2(x_{r_2}, y) + 4 r_1 r_2 L^*. \tag{185}$$

Proof. By optimality of x_r we have for every $x \in X$

$$\begin{aligned} \frac{E_{s+r}(x_r) - E_{s+r}(x)}{d_{s+h}(x_r, x)} &\le \frac{d_{s+h}^2(x, y) - d_{s+h}^2(x_r, y)}{2r d_{s+h}(x_r, x)} \\ =&\frac{(d_{s+h}(x, y) - d_{s+h}(x_r, y))(d_{s+h}(x, y) + d_{s+h}(x_r, y))}{2r d_{s+h}(x_r, x)} \\ \le&\frac{(d_{s+h}(x, y) + d_{s+h}(x_r, y))}{2r}. \end{aligned}$$

Taking the limsup as $x \to x_r$ we get the assertion. To show the second assertion note that on the one hand we have

$$E_{s+r_1}(x_{r_1}) + \frac{1}{2r_1} d_{s+h}^2(x_{r_1}, y) \le E_{s+r_1}(x_{r_2}) + \frac{1}{2r_1} d_{s+h}^2(x_{r_2}, y),$$

and on the other

$$E_{s+r_2}(x_{r_2}) + \frac{1}{2r_2} d_{s+h}^2(x_{r_2}, y) \le E_{s+r_2}(x_{r_1}) + \frac{1}{2r_2} d_{s+h}^2(x_{r_1}, y).$$

Adding these two inequalities, using the Lipschitz property of $t \mapsto E_t(x)$ and dividing by $\frac{1}{2r_1} - \frac{1}{2r_2}$ yields (185). □

Lemma 4.16. *The map $r \mapsto J_{s,r}(y)$ is locally Lipschitz and for almost every $r \in (0, T-s)$ we have for $x_r \in A_{s,r}(y)$*

$$\frac{d}{dr} J_{s,r}(y) = -\frac{1}{2r^2} d^2_{s+h}(x_r, y) + (\partial_r E_{s+r})(x_r). \tag{186}$$

Proof. Fix $0 < r_1 < r_2 < T-s$. Then

$$\begin{aligned} J_{s,r_2}(y) - J_{s,r_1}(y) =& E_{s+r_2}(x_{r_2}) - E_{s+r_1}(x_{r_1}) \\ &+ \frac{1}{2r_2} d^2_{s+h}(x_{r_2}, y) - \frac{1}{2r_1} d^2_{s+h}(x_{r_1}, y) \\ \leq& E_{s+r_2}(x_{r_1}) - E_{s+r_1}(x_{r_1}) + \frac{r_1 - r_2}{2r_2 r_1} d^2_{s+h}(x_{r_1}, y) \\ \leq& L^*(r_2 - r_1) - \frac{r_2 - r_1}{2r_2 r_1} d^2_{s+h}(x_{r_1}, y), \end{aligned} \tag{187}$$

where L^* denotes the Lipschitz constant from **A4**. Conversely, changing the roles of x_{r_1} and x_{r_2}, we obtain

$$J_{s,r_2}(y) - J_{s,r_1}(y) \geq -L^*(r_2 - r_1) - \frac{r_2 - r_1}{2r_2 r_1} d^2_{s+h}(x_{r_2}, y).$$

Combining these two inequalities yields

$$|J_{s,r_2}(y) - J_{s,r_1}(y)| \leq L^*|r_2 - r_1| + \frac{|r_2 - r_1|}{2r_1 r_2} d^2_{s+h}(x_{r_2}, y),$$

which means $r \mapsto J_{s,r}(y)$ is locally Lipschitz. Dividing by $r_2 - r_1$ and letting $r_1 \to r_2$ in (187) yields on the one hand for the left derivative

$$\frac{d^-}{dr} J_{s,r}(y) \leq -\frac{1}{2r^2} d^2_{s+h}(x_r, y) + (\partial_r E_{s+r})(x_r),$$

for every differentiability point r of $r \mapsto E_{t+r}$. On the other hand we obtain similarly for the right derivative

$$\frac{d^+}{dr} J_{s,r}(y) \geq -\frac{1}{2r^2} d^2_{s+h}(x_r, y) + (\partial_r E_{s+r})(x_r),$$

for every differentiability point of $r \mapsto E_{t+r}$. By local Lipschitz continuity we have for a.e. $0 < r < T-s$

$$\frac{d}{dr} J_{s,r}(\nu) = -\frac{1}{2r^2} d^2_{s+h}(x_r, y) + (\partial_r E_{s+r})(x_r).$$

□

Lemma 4.17. *For $s \in [0, T]$ and $0 < r_1 < r_2 < T-s$*

$$E_s(y) \geq J_{s,r_1}(y) - Cr_1 \geq J_{s,r_2}(y) - Cr_2 \tag{188}$$

$$\lim_{r \to 0} d_{s+h}(y, x_r) = 0 \text{ if } y \in Dom(E). \tag{189}$$

In particular $\lim_{r\to 0} J_{s,r}(y) = E_s(y)$.

Proof. The first inequality in (188) directly follows from

$$E_{s+r_1}(x_{r_1}) + \frac{1}{2r_1} d^2_{s+h}(x_{r_1}, y) \le E_{s+r_1}(y) \le E_s(y) + L^* r_1.$$

The second one follows by

$$\begin{aligned} E_{s+r_1}(x) + \frac{1}{2r_1} d^2_{s+h}(x,y) &\ge E_{s+r_1}(x) + \frac{1}{2r_2} d^2_{s+h}(x,y) \\ &\ge E_{s+r_2}(x) + \frac{1}{2r_2} d^2_{s+h}(x,y) - L^*(r_2 - r_1), \end{aligned}$$

and minimizing over all x. Since for every $x \in Dom(E)$

$$\begin{aligned} 0 \le d^2_{s+h}(y, x_r) &\le -2r E_{s+r}(x_r) + d^2_{s+h}(y,x) + 2r E_{s+r}(x) \\ &\le -2r \inf E + d^2_{s+h}(y,x) + 2r E_{s+r}(x). \end{aligned}$$

Passing to the limit $r \to 0$

$$\lim_{r\to 0} d^2_{s+h}(x_r, y) \le d^2_{s+h}(x,y) \text{ for every } x \in Dom(E).$$

Since $y \in Dom(E)$ we conclude (189). To check the last one we combine (188) with the lower semicontinuity of $x \mapsto E_t(x)$,

$$E_t(y) \ge \limsup_{r\to 0} J_{t,r}(y) \ge \liminf_{r \to 0} E_{t+r}(x_r) \ge E_t(y).$$

□

Corollary 4.18. *For every $0 < r_0 < T - s$ we have*

$$\begin{aligned} &E_{s+r_0}(x_{r_0}) + \frac{1}{2r_0} d^2_{s+h}(x_{r_0}, y) \\ &= E_s(y) - \int_0^{r_0} \frac{1}{2r^2} d^2_{s+h}(x_r, y) dr + \int_0^{r_0} (\partial_r E_{s+r})(x_r) dr. \end{aligned} \tag{190}$$

Proof. Integrate (186) from 0 to r_0 and use that $\lim_{r\to 0} J_{s,r}(y) = E_s(y)$. □

In the following we introduce dynamic counterparts for the variational interpolation, the discrete speed and the discrete slope, cf. [4, 53].

Definition 4.19. *Let $\bar{x} \in Dom(E)$ be the initial value and x_n^h be a sequence defined by the minimization problem* (182)*. A* discrete solution *is a curve $t \mapsto \bar{x}_t^h$ defined by*

$$\bar{x}_t^h = x_n^h, \text{ for } t \in (t_{n-1}, t_n],$$

and $\bar{x}_0^h = \bar{x}$.
A variational interpolation *is a map $t \to \tilde{x}_t^h$ defined by*

$$\begin{aligned} \tilde{x}_t^h = \arg\min \Big\{ E_t(x) + \frac{1}{2r} d^2_{t_n}(x, x^h_{n-1}) \Big\}, \\ \text{for } t = t_{n-1} + r \in (t_{n-1}, t_n], \end{aligned}$$

and $\tilde{x}_0^h = \bar{x}$.
We define the discrete speed $Dsp^h \colon [0,T] \to [0,\infty)$ *and the* discrete slope $Dsl^h \colon [0,T] \to [0,\infty)$ *in the following way*

$$\begin{aligned} Dsp_r^h &= \frac{1}{h} d_{t_n}(\bar{x}_{t_n}^h, \bar{x}_{t_{n-1}}^h), \quad r \in (t_{n-1}, t_n], \\ Dsl_r^h &= \frac{1}{(r - t_{n-1})} d_{t_n}(\bar{x}_{t_{n-1}}^h, \tilde{x}_r^h), \quad r \in (t_{n-1}, t_n]. \end{aligned}$$

Note that $\tilde{x}_{t_n}^h = x_n^h = \bar{x}_{t_n}^h$.

Proposition 4.20. *We have for* $0 \le s \le t \le T$

$$\begin{aligned} E_{\overline{h}(t)}(\bar{x}_t^h) &+ \frac{1}{2}\int_{\overline{h}(s)}^{\overline{h}(t)} (Dsp_r^h)^2 dr + \frac{1}{2}\int_{\overline{h}(s)}^{\overline{h}(t)} (Dsl_r^h)^2 dr \\ &= E_{\overline{h}(s)}(\bar{x}_s^h) + \int_{\overline{h}(s)}^{\overline{h}(t)} (\partial_r E_r)(\tilde{x}_r^h) dr. \end{aligned} \tag{191}$$

Proof. Let $t \in (t_{n-1}, t_n]$. We want to apply equation (190) with $s = t_{n-1}$, $r_0 = t - s$, $y = x_{n-1}^h$. Then with $x_{r_0} = \tilde{x}_t^h$ and $x_r = \tilde{x}_{t_{n-1}+r}^h$ we find

$$\begin{aligned} E_t(\tilde{x}_t^h) &+ \frac{1}{2(t - t_{n-1})} d_{t_n}^2(\tilde{x}_t^h, x_{n-1}^h) + \int_{t_{n-1}}^{t} \frac{1}{2(r - t_{n-1})^2} d_{t_n}^2(x_{n-1}^h, \tilde{x}_r^h) dr \\ &= E_{t_{n-1}}(x_{n-1}^h) + \int_{t_{n-1}}^{t} (\partial_r E_r)(\tilde{x}_r^h) dr. \end{aligned}$$

For $t = t_n$ we obtain

$$\begin{aligned} E_{t_n}(\bar{x}_{t_n}^h) &+ \frac{1}{2h^2}\int_{t_{n-1}}^{t_n} d_{t_n}^2(\bar{x}_{t_n}^h, \bar{x}_{t_{n-1}}^h) dr + \int_{t_{n-1}}^{t_n} \frac{1}{2(r - t_{n-1})^2} d_{t_n}^2(\bar{x}_{t_{n-1}}^h, \tilde{x}_r^h) dr \\ &= E_{t_{n-1}}(\bar{x}_{t_{n-1}}^h) + \int_{t_{n-1}}^{t_n} (\partial_r E_r)(\tilde{x}_r^h) dr. \end{aligned} \tag{192}$$

Summing up from $n+1$ to m yields

$$\begin{aligned} &E_{t_m}(\bar{x}_{t_m}^h) + \frac{1}{2h^2}\sum_{j=n+1}^{m} \int_{t_{j-1}}^{t_j} d_{t_j}^2(\bar{x}_{t_j}^h, \bar{x}_{t_{j-1}}^h) dr \\ &+ \sum_{j=n+1}^{m} \int_{t_{j-1}}^{t_j} \frac{1}{2(r - t_{j-1})^2} d_{t_j}^2(\bar{x}_{t_{j-1}}^h, \tilde{x}_r^h) dr = E_{t_n}(\bar{x}_{t_n}^h) + \int_{t_n}^{t_m} (\partial_r E_r)(\tilde{x}_r^h) dr. \end{aligned}$$

Now plugging in the definitions of the discrete slope and the discrete speed respectively

$$\begin{aligned} E_{t_m}(\bar{x}_{t_m}^h) &+ \frac{1}{2}\int_{t_n}^{t_m} (Dsp_r^h)^2 dr + \frac{1}{2}\int_{t_n}^{t_m} (Dsl_r^h)^2 dr \\ &= E_{t_n}(\bar{x}_{t_n}^h) + \int_{t_n}^{t_m} (\partial E_r)(\tilde{x}_r^h) dr, \end{aligned}$$

which shows (191).

□

Remark 4.21. *Alternatively, for $t \in (t_{n-1}, t_n]$ we can write*

$$E_t(\tilde{x}_t^h) + \frac{1}{2(t-t_{n-1})} d_{t_n}^2(\tilde{x}_t^h, \tilde{x}_{t_{n-1}}^h) + \frac{1}{2}\int_0^{t_{n-1}} (Dsp_r^h)^2 dr + \frac{1}{2}\int_0^t (Dsl_r^h)^2 dr$$
$$= E_0(\bar{x}) + \int_0^t (\partial_r E_r)(\tilde{x}_r^h) dr.$$

The following proposition provides essential a priori bounds.

Proposition 4.22. *There exist constants C_1, C_2, C_3 such that for all $0 \le t, mh \le T$*

$$E_t(\tilde{x}_t^h) \le C_1, \tag{193}$$

$$\frac{1}{2h}\sum_{n=1}^{m} d_{t_n}^2(\bar{x}_{t_n}^h, \bar{x}_{t_{n-1}}^h) \le C_2, \tag{194}$$

$$d_{t^*}^2(\tilde{x}_t^h, \bar{x}_t^h) \le C_3 h, \text{ for some fixed } t^*. \tag{195}$$

Proof. From Remark 4.21 we deduce

$$E_t(\tilde{x}_t^h) \le E_0(\bar{x}) + L^* T,$$

which shows (193).

We drop the nonnegative slope term in equation (192) to obtain

$$\frac{1}{2h} d_{t_n}^2(\bar{x}_{t_n}^h, \bar{x}_{t_{n-1}}^h) \le E_{t_{n-1}}(\bar{x}_{t_{n-1}}^h) - E_{t_n}(\bar{x}_{t_n}^h) + \int_{t_{n-1}}^{t_n} (\partial_r E_r)(\tilde{x}_r^h) dr.$$

Summing up to m and applying the Lipschitz property of $t \mapsto E_t$

$$\frac{1}{2h}\sum_{n=1}^{m} d_{t_n}^2(\bar{x}_{t_n}^h, \bar{x}_{t_{n-1}}^h) \le E_{t_0}(\bar{x}_{t_0}^h) - E_{t_m}(\bar{x}_{t_m}^h) + \int_{t_0}^{t_{m-1}} (\partial_r E_r)(\tilde{x}_r^h) dr$$
$$\le E_{t_0}(\bar{x}_{t_0}^h) - E_{t_m}(\bar{x}_{t_m}^h) + TL^*,$$

we obtain on the one hand

$$E_{t_m}(\bar{x}_{t_m}^h) \le E_{t_0}(\bar{x}_{t_0}^h) + TL^*,$$

and since $\inf E(x) > -\infty$

$$\frac{1}{2h}\sum_{n=1}^{m} d_{t_n}^2(\bar{x}_{t_n}^h, \bar{x}_{t_{n-1}}^h) \le C_2.$$

To show (195) note that for $t \in (t_{n-1}, t_n]$

$$d_{t_n}^2(\tilde{x}_t^h, \bar{x}_t^h) = d_{t_n}^2(\tilde{x}_t^h, x_n^h) \le 2d_{t_n}^2(\tilde{x}_t^h, x_{n-1}^h) + 2d_{t_n}^2(x_{n-1}^h, x_n^h)$$
$$\le 4d_{t_n}^2(x_n^h, x_{n-1}^h) + 8(t-t_{n-1})hC,$$

where the third inequality is a consequence of (185). Applying (194) and (161) we conclude (195). ☐

Proposition 4.23. *There exists a subsequence h_n with $\lim_n h_n = 0$, a curve $(x_t) \in \mathrm{AC}^2([0,T];X)$ and a function $A \in L^2(0,T)$ such that*

$$\bar{x}_t^{h_n} \overset{\sigma}{\rightharpoonup} x_t, \quad \tilde{x}_t^{h_n} \overset{\sigma}{\rightharpoonup} x_t \text{ for all } t,$$

$$\text{and } |Dsp^{h_n}| \rightharpoonup A \text{ weakly in } L^2(0,T).$$

Further $|\dot{x}|_t \leq A(t)$ holds almost everywhere.

Proof. We want to apply a refined version of Arzelà-Ascoli [4, Proposition 3.3.1] to the family $(\bar{x}^h)_{h>0}$. Owing to the estimates (161) and (194) we have

$$d_{t^*}(\bar{x}_t^h, \bar{x}) \leq tC_2 e^C,$$

and together with **A5** this yields that the curves $\bar{x}^h : [0,T] \to X$ take values in a σ-sequentially compact set. From the estimate (194) we further deduce

$$\int_t^s |Dsp_r^h|^2 dr \leq \sum_{j=t_n}^{t_m} \frac{1}{h} d_{t_j}^2(\bar{x}_{t_j}^h, \bar{x}_{t_{j-1}}^h) \leq 2C_2,$$

for $\overline{h}(s) = t_m$, $\underline{h}(t) = t_n$. Applying the Banach Alaoglu Theorem we can extract a subsequence h_n and a function $A \in L^2([0,T])$ such that $|Dsp^{h_n}| \rightharpoonup A$ weakly in $L^2([0,T])$. For fixed t^* and $s < t$ we deduce from the log-Lipschitz property (161)

$$\begin{aligned} d_{t^*}(\bar{x}_t^h, \bar{x}_s^h) &\leq \int_{\underline{h}(s)}^{\overline{h}(t)} \frac{1}{h} d_{t^*}(\bar{x}_r^h, \bar{x}_{r-h}^h) dr \\ &\leq \int_{\underline{h}(s)}^{\overline{h}(t)} \frac{1}{h} d_{\overline{h}(r)}(\bar{x}_r^h, \bar{x}_{r-h}^h) e^{L|\overline{h}(r)-t^*|} dr, \end{aligned}$$

and hence

$$\limsup_{n\to\infty} d_{t^*}(\bar{x}_t^{h_n}, \bar{x}_s^{h_n}) \leq \int_t^s A(r) e^{L|r-t^*|} dr.$$

Propostion 3.3.1 in [4] and (195) imply that there exists a further subsequence, not relabeled, and a limit curve $x : [0,T] \to X$ such that

$$\bar{x}_t^{h_n} \overset{\sigma}{\rightharpoonup} x_t, \quad \tilde{x}_t^{h_n} \overset{\sigma}{\rightharpoonup} x_t \quad \forall t \in [0,T].$$

This curve is absolutely continuous since

$$d_{t^*}(x_t, x_s) \leq \liminf_{n\to\infty} d_{t^*}(\bar{x}_t^{h_n}, \bar{x}_s^{h_n}) \leq \int_s^t A(r) e^{L|r-t^*|} dr,$$

In particular if we take $t^* = t$ in the argumentation above the Lebesgue differentiation theorem implies that

$$|\dot{x}|_t \leq \limsup_{s\nearrow t} \frac{1}{t-s} \int_s^t A(r) e^{L|r-t|} dr \leq A(t)$$

holds true for almost every t.

□

Proposition 4.24. *Suppose additionally to our standing assumptions* ***A1**, **A2**, **A3b**, **A4*** *and* ***A5*** *that*

- *If* $x_n \overset{\sigma}{\rightharpoonup} x$ *as* $n \to \infty$ *then*
$$\limsup_{n\to\infty} \partial_t E_t(x_n) \leq \partial_t E_t(x), \tag{196}$$
- *if* $t_n \to t$ *and* $x_n \overset{\sigma}{\rightharpoonup} x$, *then*
$$|\nabla_t E_t|^2(x) \leq \liminf |\nabla_{t_n} E_t|^2(x_n).$$

Then every limit curve x_t *from Proposition 4.23 satisfies the EDI formula*

$$E_t(x_t) + \frac{1}{2}\int_0^t |\dot{x}|_r^2 dr + \frac{1}{2}\int_0^t |\nabla_r E_r|^2(x_r)dr \leq E_0(\bar{x}) + \int_0^t (\partial_r E_r)(x_r)dr, \tag{197}$$

for every $t \in [0,T]$.

Proof. Recall that Proposition 4.20 states for $s = 0$

$$E_{\overline{h_n}(t)}(\bar{x}_t^{h_n}) + \frac{1}{2}\int_0^{\overline{h_n}(t)} (Dsp_r^{h_n})^2 dr + \frac{1}{2}\int_0^{\overline{h_n}(t)} (Dsl_r^{h_n})^2 dr$$
$$= E_0(\bar{x}) + \int_0^{\overline{h_n}(t)} (\partial_r E_r)(\tilde{x}_r^{h_n})dr.$$

Since both $\bar{x}_t^{h_n}$, $\tilde{x}_t^{h_n}$ σ-converges to x_t for every t, $x \mapsto E_t(x)$ and $x \mapsto \partial_t E_t(x)$ is σ-upper semicontinuous and $t \to E_t(x)$ is Lipschitz continuous uniformly in x, we know

$$\liminf_{n\to\infty} E_{\overline{h_n}(t)}(\bar{x}_t^{h_n}) \geq E_t(x_t),$$

and

$$\int_0^t (\partial_r E_r)(x_r)dr \geq \int_0^t \limsup (\partial_r E_r)(\tilde{x}_r^{h_n})dr \geq \limsup \int_0^t (\partial_r E_r)(\tilde{x}_r^{h_n})dr,$$

where the last inequality follows from Fatou's Lemma. From Proposition 4.23 and Lemma 4.15 we deduce

$$\int_0^t |\dot{x}|_r^2 dr \leq \int_0^t A(r)^2 dr \leq \liminf_{n\to\infty} \int_0^t (Dsp_r^{h_n})^2 dr,$$

and

$$\int_0^t |\nabla_r E_r|^2(x_r)dr \leq \liminf \int_0^t |\nabla_{\overline{h_n}(r)} E_r|^2(\tilde{x}_r^{h_n})dr \leq \liminf \int_0^t (Dsl_r^{h_n})^2 dr$$

Combining these inequalities with (191) we conclude

$$\begin{aligned}
&E_t(x_t) + \frac{1}{2}\int_0^t |\dot{x}|_r^2 dr + \frac{1}{2}\int_0^t |\nabla_r E_r|^2(x_r)dr \\
\leq &\liminf \left[E_{\overline{h_n}(t)}(\bar{x}_t^{h_n}) + \frac{1}{2}\int_0^{\overline{h_n}(t)} (Dsp_r^{h_n})^2 dr + \frac{1}{2}\int_0^{\overline{h_n}(t)} (Dsl_r^{h_n})^2 dr \right] \\
\leq &\liminf \left[E_0(\bar{x}) + \int_0^{\overline{h_n}(t)} (\partial_r E_r)(\tilde{x}_r^{h_n})dr \right] \\
\leq &E_0(\bar{x}) + \int_0^t (\partial_r E_r)(x_r)dr,
\end{aligned}$$

which is the assertion.

□

4.4 Dynamic Gradient Flow of the Entropy

In this section we want to study gradient flows for the Boltzmann entropy on space of probability measures, where the metric of the space and the reference measure of the entropy varies in time. To show existence we apply the results from Section 4.3.3. We then go on to show also uniqueness.

Let X be a topological space equipped with a family of complete separable geodesic metrics $(d_t)_{t\in[0,T]}$ satisfying (161) and a Borel probability measure m. We define $\mathcal{P}(X)$ to be the space of Borel probability measures on X and we denote the subspace of probability measures absolutely continuous to the measure m by $\mathcal{P}^{ac}(X)$. Further let $\mathcal{P}_2(X)$ be the space of probability measures with bounded second moments on X

$$\mathcal{P}_2(X) := \Big\{\mu \in \mathcal{P}(X)\Big| \int d_t^2(x,x_0)d\mu(x) < \infty \\ \text{for some, and thus any, } x_0 \in X, t \in [0,T]\Big\}.$$

We say that a sequence $\mu_n \subset \mathcal{P}(X)$ converges weakly to μ if $\lim \int_X f d\mu_n = \int_X f d\mu$ for every continuous bounded function $f \in \mathcal{C}_b^0(X)$. We say that a sequence $\rho_n \subset L^1(X,m)$ converges weakly to ρ if $\lim \int_X f\rho_n dm = \int_X f\rho dm$ for every bounded function $f \in L^\infty(X,m)$. Note that if ρ_n converges weakly to ρ in $L^1(X,m)$ then $\mu_n = \rho_n m$ converges weakly to $\mu = \rho m$ in $\mathcal{P}(X)$.

4.4.1 Time-dependent Kantorovich Metrics

For every metric d_t we define the L^2-Kantorovich distance W_t on the space $\mathcal{P}_2(X)$:

$$W_t(\mu,\nu) = \inf\{C_t(\gamma) : \pi^1_\#\gamma = \mu, \pi^2_\#\gamma = \nu\}^{1/2},$$

where $C_t(\gamma)$ is the cost of the plan $\gamma \in \mathcal{P}(X \times X)$

$$C_t(\gamma) = \int d_t^2(x,y)d\gamma(x,y),$$

and $\pi^i_\#\gamma$ denote the first and second marginals of γ.

For each $t \in [0,T]$, $(\mathcal{P}_2(X), W_t)$ is a geodesic Polish space, see e.g. [62, 2]. It is well-known that convergence in the L^2-Kantorovich distance W_t implies weak convergence in $\mathcal{P}(X)$ and that W_t is lower semicontinuous on $\mathcal{P}(X)$ (cf. [62, Theorem 6.8] and [62, Remark 6.10]). The bound (161) is equivalent to

$$|\log W_t(\mu,\nu)/W_s(\mu,\nu)| \le L|t-s|, \tag{198}$$

for all s,t and all probability measures on X, see Lemma 2.1 in [59].

The convexity of the squared metric speed is crucial for showing uniqueness of the gradient flow. More precisely we have the following result [26, Lemma 14].

Lemma 4.25. *Let* $(\mu_t^1), (\mu_t^2) \in \mathrm{AC}^2([0,T]; \mathcal{P}(X))$ *be two absolutely continuous curves. Define* $\mu_t^{1,2} = (\mu_t^1 + \mu_t^2)/2$. *Then* $(\mu_t^{1,2})$ *is absolutely continuous and the following bound on its metric derivative holds*

$$|\dot{\mu}^{1,2}|_t^2 \leq \frac{|\dot{\mu}^1|_t^2 + |\dot{\mu}^2|_t^2}{2}.$$

Proof. Fix $s, t \in [0,T]$. Pick optimal plans γ^1, γ^2, which minimize $W_t(\mu_t^1, \mu_s^1)$ and $W_t(\mu_t^2, \mu_s^2)$ respectively. Then the plan $(\gamma^1 + \gamma^2)/2$ has marginals $\mu_t^{1,2}$ and $\mu_s^{1,2}$ and therefore it holds

$$\begin{aligned} W_t^2(\mu_t^{1,2}, \mu_s^{1,2}) \leq \int d_t^2(x,y)\, d\frac{\gamma^1+\gamma^2}{2}(x,y) \\ = \frac{1}{2}\int d_t^2(x,y)\, d\gamma^1(x,y) + \frac{1}{2}\int d_t^2(x,y)\, d\gamma^2(x,y) \\ = \frac{1}{2}W_t^2(\mu_t^1, \mu_s^1) + \frac{1}{2}W_t^2(\mu_t^2, \mu_s^2). \end{aligned}$$

Thus the curve $(\mu_t^{1,2})$ is absolutely continuous. Dividing by $(s-t)^2$ and taking the superior limit as s goes to t we get for its speed

$$|\dot{\mu}^{1,2}|_t^2 \leq \frac{|\dot{\mu}^1|_t^2 + |\dot{\mu}^2|_t^2}{2}.$$

□

4.4.2 Time-dependent Boltzmann Entropy

We consider a family of measures $(m_t)_{t\in[0,T]}$ on X. We suppose that for every $t \in [0,T]$ there exists a function $f_t \in L^\infty(X, m)$ such that $m_t = e^{-f_t}m$. Moreover let us always assume that there exists a constant L^* such that

$$|f_t(x) - f_s(x)| \leq L^*|t-s| \tag{199}$$

for all s, t and all x.

We denote by S_t the relative Boltzmann entropy with respect to m_t,

$$\begin{aligned} S\colon [0,T] \times \mathcal{P}_2(X) \to [-\infty, \infty], \\ (t,\mu) \mapsto S_t(\mu) = \mathrm{Ent}(\mu|m_t) = \int \rho \log \rho \, dm_t, \end{aligned}$$

where $\rho = d\mu/\,dm_t$ provided that $\mu \ll m_t$. Otherwise we set $S_t(\mu) = \infty$. It follows directly from the representation of the measures m_t that

$$S_t(\mu) = \mathrm{Ent}(\mu) + \int f_t(x) d\mu(x),$$

where $\mathrm{Ent}(\mu) = \mathrm{Ent}(\mu|m)$.

In the next lemma we list the crucial properties of the relative entropy functional.

Lemma 4.26. *The entropy* $S\colon [0,T] \times \mathcal{P}_2(X) \to [-\infty, \infty]$ *satisfies* ***A1, A2, A3b*** *and* ***A4****, i.e.*

1. *The domain $Dom(S_t)$ is time-independent.*
2. *$S_t(\mu)$ is uniformly bounded from below.*
3. *For each $t \in [0,T]$, $\mu \mapsto S_t(\mu)$ is lower semicontinuous with respect to weak convergence on th space of probability measures.*
4. *For every $\mu \in Dom(S)$ the map $t \mapsto S_t(\mu)$ is Lipschitz continuous with Lipschitz constant L^* and for the derivative it holds*
$$\partial_t S_t(\mu) = \int_X \partial_t f_t(x) d\mu(x) \text{ for a.e. } t \in [0,T].$$
Moreover the set of differentiability points of $t \mapsto S_t(\mu)$ can be chosen independent of μ.

Proof. The domain is time-independent by virtue of (199). Since $m(x) = 1$ we can estimate $\mathrm{Ent}(\mu) \geq 0$ and hence for $\mu \in \mathcal{P}^{ac}(X)$
$$S_t(\mu) \geq \int f_t(x) d\mu(x) \geq -||f_t||_{L^\infty} \geq -||f_s||_{L^\infty} - L^* T.$$
If $\mu \notin \mathcal{P}^{ac}(X)$ we know that $S_t(\mu) = \infty$ and we conclude $\inf_{t,\mu} S_t(\mu) > -\infty$.

For every t the measure $m_t(X)$ is finite and thus $\mu \to S_t(\mu)$ is lower semicontinuous with respect to weak convergence (Lemma 4.1 in [57]).

Fix $\mu \in Dom(S)$. The Lipschitz continuity of $t \mapsto f_t(x)$ ensures $|\partial_t S_t(\mu)| \leq C$. It is clear that for every $x \in X$ the map $t \mapsto f_t(x)$ is differentiable for a.e. $t \in [0,T]$. Hence the integral $\int_X \int_{t_1}^{t_2} |\partial_t f_t(x)| dt d\mu(x)$ exists and the Fubini-Tonelli theorem states
$$\int_X \int_{t_1}^{t_2} \partial_t f_t(x) dt d\mu(x) = \int_{t_1}^{t_2} \int_X \partial_t f_t(x) d\mu(x) dt.$$
The Fubini theorem again yields that for a.e. t the map $x \mapsto \partial_t f_t(x)$ is μ-integrable and so for a.e. t the integral
$$\int_X \partial_t f_t(x) d\mu(x)$$
exists. Take a differentiability point t of $f_t(x)$. Then for μ-a.e. $x \in X$
$$\lim_{h\to 0} \frac{1}{h}(f_{t+h} - f_t)(x) = \partial_t f_t(x), \text{ and } \frac{1}{h}|(f_{t+h} - f_t)(x)| \leq L^*.$$
Hence we conclude that for a.e. $t \in [0,T]$
$$\lim_{h\to 0} \frac{1}{h}[S_{t+h}(\mu) - S_t(\mu)] = \lim_{h\to 0} \int \frac{1}{h}[f_{t+h}(x) - f_t(x)] d\mu(x) = \int \partial_t f_t(x) d\mu(x),$$
where the last equality is due to the dominated convergence theorem. Finally, for $\mu \ll m$, the inclusions
$$\begin{aligned} &\left\{ t \in [0,T] \middle| \lim \frac{1}{h}[S_{t+h}(\mu) - S_t(\mu)] \text{ exists} \right\} \\ \subset &\left\{ t \in [0,T] \middle| \lim \frac{1}{h}[f_{t+h}(x) - f_t(x)] \text{ exists for } \mu \text{ a.e. } x \right\} \\ \subset &\left\{ t \in [0,T] \middle| \lim \frac{1}{h}[f_{t+h}(x) - f_t(x)] \text{ exists for } m \text{ a.e. } x \right\} \end{aligned}$$

show that the set of differentiability points of $t \mapsto S_t(\mu)$ does not depend on μ, since the complement $\{t \in [0,T] \big| \lim \frac{1}{h}[f_{t+h}(x) - f_t(x)]$ exists for m a.e. $x\}^C$ is negligible. $\square$

Since we want to apply the results from Section 4.3.3, we still need to check the assumptions in Proposition 4.24. To show the lower semicontinuity of the squared slope, it is essential to assume that the entropy is K-convex for some $K \in \mathbb{R}$, i.e. we have a uniform bound on the Ricci curvature of (X, d_t, m_t). We briefly recall the arguments in [26].

Definition 4.27. *The set $GP \subset \mathcal{P}(X^2)$ is the set of plans γ such that*

1. *the marginals $\pi^i_\# \gamma$, $i = 1, 2$ are absolutely continuous with densities bounded away from 0 and ∞,*
2. $\sup_{(x,y) \in \operatorname{supp}(\gamma)} d_t(x,y) < \infty$ *for some $t \in [0,T]$, and thus for any.*

Given $\gamma \in GP$ and $\mu \in \mathcal{P}_2^{ac}(X)$, we define the plan $\gamma_\mu \in \mathcal{P}(X^2)$ and the measure $\nu_{\gamma,\mu} \in (X)$ as

$$d\gamma_\mu(x,y) = \frac{d\mu(x)}{d\pi^1_\#\gamma(x)} d\gamma(x,y), \quad \nu_{\gamma,\mu} = \pi^2_\# \gamma_\mu.$$

Note that since $\gamma_\mu \ll \gamma$, we have $\nu_{\gamma,\mu} \ll m$ with density

$$g_{\gamma,\mu}(y) = \frac{d\pi^2_\#\gamma(y)}{dm(y)} \int \frac{d\mu(x)}{d\pi^1_\#\gamma(x)} d\gamma_y(x),$$

where $(\gamma_y)_y \subset \mathcal{P}(X)$ is the disintegration of γ with respect to its second marginal.

Observe that from 2. of the definition of the set GP we have that the cost $C_t(\gamma)$ of a plan $\gamma \in GP$ is always finite and $\nu_{\gamma,\mu} \in \mathcal{P}_2(X)$ since $\mu \in \mathcal{P}_2(X)$.

The next Proposition gives an alternative representation formula for the slope in terms of good plans, cf. [26, Theorem 12].

Proposition 4.28. *For every $t \in [0,T]$ and every $\mu \in Dom(S)$ it holds*

$$\begin{aligned}
&\sup_{\substack{\nu \in \mathcal{P}_2(X) \\ \nu \neq \mu}} \frac{(S_t(\mu) - S_t(\nu) - \frac{K^-}{2} W_t^2(\mu,\nu))^+}{W_t(\mu,\nu)} \\
&= \sup_{\gamma \in GP} \frac{(S_t(\mu) - S_t(\nu_{\gamma,\mu}) - \frac{K^-}{2} C_t(\gamma_\mu))^+}{\sqrt{C_t(\gamma_\mu)}},
\end{aligned} \tag{200}$$

where the value of the second expression is taken by definition as 0 if $C_t(\gamma_\mu) = 0$.

Proof. We start with proving $\geq$. For this fix a plan $\gamma \in GP$ such that $\nu_{\gamma,\mu} \neq \mu$. From $C_t(\gamma_\mu) \geq W_t^2(\mu, \nu_{\gamma,\mu}) > 0$ we obtain

$$\begin{aligned}
&\frac{(S_t(\mu) - S_t(\nu_{\gamma,\mu}) - \frac{K^-}{2} W_t^2(\mu,\nu_{\gamma,\mu}))^+}{W_t(\mu,\nu_{\gamma,\mu})} \\
&\geq \frac{(S_t(\mu) - S_t(\nu_{\gamma,\mu}) - \frac{K^-}{2} C_t(\gamma_\mu))^+}{\sqrt{C_t(\gamma_\mu)}}.
\end{aligned}$$

To show the reverse inequality take $\nu \in \mathcal{P}_2^{ac}(X)$ different from μ. Lemma 10 in [26] provides a sequence $(\gamma^n) \subset GP$ such that $S_t(\nu_{\gamma^n,\mu}) \to S_t(\nu)$ and $C_t(\gamma^n_\mu) \to W_t^2(\mu,\nu)$ as $n \to \infty$ and hence

$$\begin{aligned} &\frac{(S_t(\mu) - S_t(\nu) - \frac{C}{2} W_t^2(\mu,\nu))^+}{W_t(\mu,\nu)} \\ = \lim_{n\to\infty} &\frac{(S_t(\mu) - S_t(\nu_{\gamma^n,\mu}) - \frac{C}{2} C_t(\gamma^n_\mu))^+}{\sqrt{C_t(\gamma^n_\mu)}}, \end{aligned}$$

which shows $\leq$.

□

We get the following as consequence of formula (200), cf. [26, Corollary 13].

Corollary 4.29. *Suppose that S is K-convex. Then for every $t \in [0,T]$*

$$|\nabla_t S_t|^2(\mu) \leq \liminf |\nabla_t S_t|^2(\mu_n),$$

whenever $\mu_n \rightharpoonup \mu$ as $n \to \infty$ such that $\sup_n S_t(\mu_n) < \infty$. Further $\mu \mapsto |\nabla_t S_t|^2$ is convex with respect to linear interpolation on the sublevels of S.

Proof. Consider the map $\mu \mapsto C_t(\gamma_\mu)$. It is clearly linear. Also, one can show that it is weakly continuous on sublevels of the entropy. From [26, Proposition 11] we further know that $\mu \mapsto S_t(\mu) - S_t(\nu_{\gamma,\mu})$ is lower semicontinuous with respect to weak convergence on sublevels of the entropy and convex with respect to linear interpolation. Hence

$$\mu \mapsto S_t(\mu) - S_t(\nu_{\gamma,\nu}) - \frac{K^-}{2} C_t(\gamma_\mu)$$

is lower semicontinuous with respect to weak convergence on the sublevels of the entropy. The same holds true for its positive part. Now apply that the function $\Psi\colon \mathbb{R}^2 \to \mathbb{R}$ defined by

$$\Psi(a,b) = \begin{cases} \frac{a^2}{b} & \text{if } b > 0, \\ 0 & \text{if } a = b = 0, \\ +\infty & \text{if } a \neq 0, b = 0 \textit{ or } b < 0, \end{cases}$$

is convex, continuous on $[0,\infty)^2 \setminus \{(0,0)\}$ and increasing in a, and the conclusion follows. From formula (200) the assertion follows.

□

4.4.3 Existence and Uniqueness of EDE-Gradient Flow for the Entropy

In this section we want to show existence and uniqueness of the dynamic EDI-gradient flow with respect to the functional S on the complete geodesic space $(\mathcal{P}_2(X), W_t)$. For this we additionally have to assume that X is boundedly compact, i.e. closed balls are compact. For this reason we can take **A5** for granted, as shown in the next lemma.

Lemma 4.30. *Assume that X is boundedly compact. Then the following holds true. If $(\mu_n) \subset \mathcal{P}_2(X)$ with $\sup_{n,m} W_t(\mu_n, \mu_m) < \infty$, then $(\mu_n)_n$ is sequentially precompact with respect to weak convergence.*

Proof. If $\sup_{n,m} W_t(\mu_n, \mu_m) < \infty$ for a sequence $(\mu_n) \subset \mathcal{P}_2(X)$ the second moments are uniformly bounded. Then Lemma 16 in [26] implies that (μ_n) is tight. Applying Prokhorov's theorem we infer that (μ_n) is weakly sequentially precompact.

□

Theorem 4.31. *Assume additionally that X is boundedly compact. Suppose that S is K-convex for some $K \in \mathbb{R}$. Then for every $\bar{\mu} \in Dom(S)$ there exists a curve $(\mu_t) \in \mathrm{AC}^2([0,T], \mathcal{P}_2(X))$ starting in $\bar{\mu}$ and satisfying*

$$S_t(\mu_t) + \frac{1}{2}\int_0^t |\dot{\mu}_r|_r^2 dr + \frac{1}{2}\int_0^t |\nabla_r S_r|^2(\mu_r)dr \le S_0(\mu_0) + \int_0^t (\partial_r S_r)(\mu_r)dr, \quad (201)$$

for every $t \in [0,T]$.

Proof. We may apply Proposition 4.23 and obtain a limit curve $\mu \in \mathrm{AC}^2([0,T]; \mathcal{P}_2(X))$ starting in $\bar{\mu}$ such that

$$\bar{\mu}_t^{h_n} \rightharpoonup \mu_t, \text{ and } \tilde{\mu}_t^{h_n} \rightharpoonup \mu_t \quad \forall t \in [0,T],$$

where $\bar{\mu}^h$ and $\tilde{\mu}^h$ are defined as in Definition 4.19. and satisfy by Proposition 4.20

$$\begin{aligned} S_{\overline{h_n}(t)}(\bar{\mu}_t^{h_n}) + \frac{1}{2}\int_0^{\overline{h_n}(t)} (Dsp_r^{h_n})^2 dr + \frac{1}{2}\int_0^{\overline{h_n}(t)} (Dsl_r^{h_n})^2 dr \\ = S_0(\bar{\mu}) + \int_0^{\overline{h_n}(t)} (\partial_r S_r)(\tilde{\mu}_r^{h_n})dr. \end{aligned}$$

From Corollary 4.29, Lemma 4.15 and Lemma 4.5 together with (161), applying Fatou's Lemma we obtain

$$\begin{aligned} &\int_0^t |\nabla_r S_r|^2(\mu_r)dr \le \liminf_{n\to\infty} \int_0^t |\nabla_r S_r|^2(\tilde{\mu}_r^{h_n})dr \\ &\le \liminf_{n\to\infty} \int_0^t (|\nabla_{\overline{h_n}(r)} S_r|(\tilde{\mu}_r^{h_n}) + |\nabla_r S_r|(\tilde{\mu}_r^{h_n}) - |\nabla_{\overline{h_n}(r)} S_r|(\tilde{\mu}_r^{h_n}))^2 dr \\ &\le \liminf_{n\to\infty} \left[\int_0^t (Dsl_r^{h_n})^2 dr + 2Ch_n \int (Dsl_r^{h_n})^2 dr + Ch_n^2 \int (Dsl_r^{h_n})^2 dr \right]. \end{aligned}$$

We deduce

$$\int_0^t |\nabla_r S_r|^2(\mu_r)dr \le \liminf \int_0^t (Dsl_r^{h_n})^2 dr$$

from the estimate $\int_0^t (Dsl_r^{h_n})^2 dr \le S_0(\mu) + L^*T - \inf_{t,\mu} S_t(\mu)$

To show that (201) is valid, it is left to show that

$$\lim \int_0^t (\partial_r S_r)(\tilde{\mu}_r^{h_n})dr = \int_0^t (\partial_r S_r)(\mu_r)\, dr.$$

This already follows if we prove that a stronger convergence than weak convergence of measures holds true. In fact, from (193) we know that there exists a density $\tilde{\rho}_t^{h_n} = d\tilde{\mu}_t^{h_n}/dm \in L^1(X,m)$ for every $t \in [0,T]$ and $n \in \mathbb{N}$. The lower semicontinuity of the entropy implies that $\sup_t S_t(\mu_t) < \infty$ and thus $\mu_t = \rho_t m$, for some $\rho_t \in L^1(X,m)$. Choose an arbitrary subsequence h_{n_k}. Then since the family of densities $(\tilde{\rho}_t^{h_{n_k}})_k$ is equiintegrable, i.e.

$$\sup_k \int_X \max\{0, \rho_t^{h_{n_k}} \log \rho_t^{h_{n_k}}\} dm < \infty,$$

(cf. [17, Theorem 4.5.9]), the Dunford-Pettis Theorem ([17, Corollary 4.7.19]) ensures that there exists a subsubsequence $\tilde{\rho}_t^{h_{n_{k_l}}}$ that converges in the weak topology of $L^1(X,m)$ to the function $\rho_t \in L^1(X,m)$. Hence for the original subsequence we already have

$$\tilde{\rho}_t^{h_n} \rightharpoonup \rho_t \text{ in } L^1(X,m) \quad \forall t \in [0,T].$$

As a direct consequence we obtain (201), since similar as in Proposition 4.24

$$\begin{aligned}
&S_t(\mu_t) + \frac{1}{2}\int_0^t |\dot{\mu}|_r^2 dr + \frac{1}{2}\int_0^t |\nabla_r S_r|^2(\mu_r) dr \\
\leq &\liminf_{n\to\infty} \left[S_{\overline{h_n}(t)}(\bar{\mu}_t^{h_n}) + \frac{1}{2}\int_0^{\overline{h_n}(t)} (Dsp_r^{h_n})^2 dr + \frac{1}{2}\int_0^{\overline{h_n}(t)} (Dsl_r^{h_n})^2 dr \right] \\
\leq &\liminf_{n\to\infty} \left[S_0(\bar{\mu}) + \int_0^{\overline{h_n}(t)} (\partial_r S_r)(\tilde{\mu}_r^{h_n}) dr \right] \\
\leq &S_0(\bar{\mu}) + \int_0^t (\partial_r S_r)(\mu_r) dr.
\end{aligned}$$

□

Remark 4.32. *Actually, the statement of Theorem 4.31 holds true without assuming that X is boundedly compact, since m is assumed to be finite. If (X,d) is a Polish space and $m \in \mathcal{P}(X)$ we may apply $z \log z \geq -1/e$ and Jensen's inequality to obtain*

$$\mathrm{Ent}_m(\mu) \geq \mu(E) \log\left(\frac{\mu(E)}{m(E)}\right) - \frac{1}{e} \quad \forall E \in \mathcal{B}(X).$$

*Taking into account that the singleton $\{m\}$ is tight, this shows tightness of the sublevels of the entropy since $\mu(E) \to 0$ as $m(E) \to 0$. Hence we could replace our assumption **A5** in section 4.3.3 by assuming that for each $t \in [0,T]$ the sublevels of the functional are sequentially σ-compact. See also [5, Remark 7.3].*

Theorem 4.33. *Assume S is K-convex and $\bar{\mu} \in Dom(S)$. Then there exists at most one dynamic EDI-gradient flow (μ_t) starting in $\bar{\mu}$. Moreover we have equality in* (201)*, i.e. for every $t \in [0,T]$ it holds the following dynamic EDE*

$$S_t(\mu_t) + \frac{1}{2}\int_0^t |\dot{\mu}_r|_r^2 dr + \frac{1}{2}\int_0^t |\nabla_r S_r|^2(\mu_r) dr = S_0(\mu_0) + \int_0^t (\partial_r S_r)(\mu_r) dr.$$

Proof. Let us first observe that a weak chain rule for gradient flows is applicable. For this we prove that a variant of the assumption in Proposition 4.8 concerning the time derivative is satisfied by the entropy S_t. We choose a sequence $\mu_n = \rho_n m$ converging to $\mu = \rho m$ such that $\sup_n S(\mu_n) < \infty$. We need to show that for almost every t

$$\lim_{n\to\infty} \frac{S_{t_n}(\mu_n) - S_t(\mu_n)}{t_n - t} = \lim_{n\to\infty} \int_X \frac{f_{t_n}(x) - f_t(x)}{t_n - t} \rho_n(x) dm(x) = \partial_t S_t(\mu), \tag{202}$$

if $t_n \searrow t$ as $n \to \infty$. This would imply the weak chain rule (171) in Proposition 4.8 restricted to curves which are contained in the sublevels of the functional. In order to show (202) note that as in the proof of Theorem 4.31 the sequence (ρ_n) is equi-integrable and thus ρ_n converges to ρ in duality with L^∞ functions. Then we decompose

$$\begin{aligned}
&\int_X \frac{f_{t_n} - f_t}{t_n - t} \rho_n dm = \int_X \Big(\frac{f_{t_n} - f_t}{t_n - t} - \partial_t f_t\Big) \rho_n dm + \int \partial_t f_t \rho_n \, dm \\
&= \int_{|\rho_n| < M} \Big(\frac{f_{t_n} - f_t}{t_n - t} - \partial_t f_t\Big) \rho_n dm + \int_{|\rho_n| \geq M} \Big(\frac{f_{t_n} - f_t}{t_n - t} - \partial_t f_t\Big) \rho_n dm \\
&\quad + \int \partial_t f_t \rho_n \, dm.
\end{aligned}$$

The third integral clearly converges to $\int \partial_t f_t \rho \, dm = \partial_t S_t(\mu)$ by Lemma 4.26, while the first vanishes by dominated convergence. The second vanishes after letting $n \to \infty$ and then $M \to \infty$ by equi-integrability of (ρ_n).

Let us assume that there exist two dynamic EDI-gradient flows (μ_t^1), (μ_t^2) starting from $\bar{\mu} \in Dom(S)$. As seen in the proof of Theorem 4.31 we know that these curves are contained in the sublevels of S and hence together with the weak chain rule it follows

$$S_0(\bar{\mu}) = S_t(\mu_t^1) + \frac{1}{2} \int_0^t |\dot{\mu}^1|_r^2 dr + \frac{1}{2} \int_0^t |\nabla_r S_r|^2(\mu_r^1) dr - \int_0^t \partial_r S_r(\mu_r^1) dr,$$

$$S_0(\bar{\mu}) = S_t(\mu_t^2) + \frac{1}{2} \int_0^t |\dot{\mu}^2|_r^2 dr + \frac{1}{2} \int_0^t |\nabla_r S_r|^2(\mu_r^2) dr - \int_0^t \partial_r S_r(\mu_r^2) dr.$$

Now define

$$\mu_t^{1,2} = \frac{\mu_t^1 + \mu_t^2}{2} \quad t \geq 0.$$

Then $\mu_0^{1,2} = \bar{\mu}$ and from the strict convexity of the entropy, the convexity of the squared slope (Corollary 4.29) and the convexity of the squared speed (Lemma 4.25) we have that

$$S_0(\bar{\mu}) > S_t(\mu_t^{1,2}) + \frac{1}{2} \int_0^t |\dot{\mu}^{1,2}|_r^2 dr + \frac{1}{2} \int_0^t |\nabla_r S_r|^2(\mu_r^{1,2}) dr - \int_0^t \partial_r S_r(\mu_r^{1,2}) dr,$$

whenever these curves are different. But since (202) is applicable to $\mu_t^{1,2}$, this contradicts (171). □

4.5 Dynamic Gradient Flows in Hilbert Spaces

Let $(H, \langle \cdot, \cdot \rangle_t)_{t\in[0,T]}$ be a one-parameter family of separable Hilbert spaces. We assume that (161) holds for the distances $||x-y||_t := \sqrt{\langle x-y, x-y\rangle_t}$. Let $E\colon [0,T]\times H \to \mathbb{R}\cup\{+\infty\}$ be a functional such that $x \mapsto E_t(x)$ is convex and lower semicontinuous. Again we require that the domain $Dom(E_t) = \{x\colon E_t(x) < \infty\}$ is time independent. The subdifferential $D_t^- E_t(x)$ of E_t at some $x \in Dom(E)$ is the set of all $v \in H$ such that

$$E_t(y) - E_t(x) \geq \langle v, y-x\rangle_t \qquad \forall y \in H.$$

It follows from the definition of the subdifferential that $D_t^- E_t$ is *monotone*, i.e. for every $v \in D_t^- E_t(x)$, $w \in D_t^- E_t(y)$ we have

$$\langle v-w, x-y\rangle_t \geq 0. \tag{203}$$

Note that $D_t^- E_t(x)$ is closed and convex. Hence we can set $\nabla_t E_t(x)$ as the element of minimal $||\cdot||_t$-norm in $D_t^- E_t(x)$ as soon as $D_t^- E_t(x) \neq \emptyset$.

Definition 4.34. *We say that $(x_t)_{0\leq t\leq T}$ is a* dynamic gradient flow *for E starting from $x \in H$ if it is locally absolutely continuous and*

$$\partial_t x_t \in -D_t^- E_t(x_t) \text{ for a.e. } t \in (0,T)$$

and $\lim_{t\searrow 0} x_t = x$.

We cannot hope to have a minimal selection result, i.e. $\frac{d^+}{dt} x_t = -\nabla_t E_t(x_t)$. We illustrate this in the following example.

Example 5. *Consider once again the energy functional $E_t(x) = |x-t|$ on $\mathbb{R}$. Then the curve $x_t = t$ defines a gradient flow of E_t since $\partial_t x_t = 1 \in -D^- E_t(x_t)$, but $\partial_t x_t \neq -\nabla E_t(x_t) = 0$.*

In the following we show that the gradient flow in the sense of Definition 4.34 is a dynamic forward EVI$(-L/2,\infty)$ gradient flow introduced in Section 4.3.2. We recall that for $s,t \in [0,T]$, $\gamma \in \mathrm{AC}^2([0,1];H)$ the action of the curve

$$\mathcal{A}_{s,t}(\gamma) := \lim_{h\to 0} \sup \Big\{ \sum_{i=1}^n (a_i - a_{i-1})^{-1} ||\gamma_{a_i} - \gamma_{a_{i-1}}||^2_{s+a(t-s)} \Big\},$$

where the supremum runs over all partitions $0 = a_0 < a_1 < \cdots a_n = 1$ such that $a_i - a_{i-1} \leq h$ for some $h > 0$.

For $x, y \in H$ we define

$$||x-y||^2_{s,t} := \inf A_{s,t}(\gamma),$$

where the infimum runs over all curves $\gamma \in \mathrm{AC}^2([0,1];H)$ such that $\gamma_0 = x$ and $\gamma_1 = y$.

Proposition 4.35. *Let $E\colon [0,T]\times H \to (-\infty,+\infty]$ be a functional such that $x \mapsto E_t(x)$ is convex and lower semicontinuous for each $t \in [0,T]$. Let (x_t) be a gradient flow of E in the sense of Definition 4.34. Then, with L denoting the the logarithmic Lipschitz control* (161) *of the distances, (x_t) is a dynamic forward EVI$(-L/2,\infty)$ gradient flow, i.e. for all $y \in Dom(E)$ and a.e. t*

$$\frac{1}{2}\partial_s^+ ||x_s - y||^2_{s,t}\Big|_{s=t} - \frac{L}{4}||x_t - y||_t^2 \leq E_t(y) - E_t(x_t).$$

Proof. Let $y \in Dom(E)$. Then

$$\begin{aligned}\frac{1}{2}\partial_s^+ ||x_s - y||_{s,t}^2\Big|_{s=t} &= \limsup_{s\to t}\left\{\frac{1}{2(s-t)}(||x_s - y||_{s,t}^2 - ||x_t - y||_t^2)\right\}\\ &\le \limsup_{s\to t}\left\{\frac{1}{2(s-t)}(||x_s - y||_{s,t}^2 - ||x_s - y||_t^2)\right\}\\ &+ \limsup_{s\to t}\left\{\frac{1}{2(s-t)}(||x_s - y||_t^2 - ||x_t - y||_t^2)\right\}\end{aligned}$$

The first limsup can be estimated with the help of Proposition 2.72(iii) by

$$\begin{aligned}&\limsup_{s\to t}\left\{\frac{1}{2(s-t)}(||x_s - y||_{s,t}^2 - ||x_s - y||_t^2)\right\}\\ \le&\limsup_{s\to t}\left\{\frac{1}{2(s-t)}\Big(\frac{e^{L|t-s|}-1}{L|t-s|} - 1\Big)||x_s - y||_s^2\right\}\\ =&\limsup_{s\to t}\left\{\frac{1}{2(s-t)}\Big(\frac{\frac{1}{2}L|t-s|^2 + o(|t-s|^2)}{L|t-s|}\Big)||x_s - y||_s^2\right\}\\ \le&\frac{L}{4}||x_t - y||_t^2,\end{aligned}$$

where the last inequality follows from the continuity of $t \mapsto x_t$ and $t \mapsto ||\cdot||_t$. For the second limsup we apply that (x_t) is supposed to be a gradient flow of E;

$$\begin{aligned}&\limsup_{s\to t}\left\{\frac{1}{2(s-t)}(||x_s - y||_t^2 - ||x_t - y||_t^2)\right\}\\ =&\langle x_t - y, \partial_t x_t\rangle_t \le E_t(y) - E_t(x_t)\end{aligned}$$

for a.e. $t \ge 0$. Combining these two observations we conclude

$$\frac{1}{2}\partial_s^+ ||x_s - y||_{s,t}^2\Big|_{s=t} \le \frac{L}{4}||x_t - y||_t^2 + E_t(y) - E_t(x_t),$$

which proves the claim. □

4.5.1 Existence and Uniqueness

We assume that the following holds for the energy functional, cf. [54].

1. $x \mapsto E_t(x)$ is lower semicontinuous
   and $E_t(x) \ge 0 \quad \forall (t,x) \in [0,T] \times Dom(E)$,
2. $\exists C_1 \, \forall x \in Dom(E) \, \forall s,t \in [0,T]: \quad |E_t(x) - E_s(x)| \le C_1 E_t(x)|t-s|$.

By virtue of the functional's lower semicontinuity we obtain that if $v_n \in D_t^- E_t(x_n)$ and $x_n \to x$, $v_n \rightharpoonup v$, then $v \in D_t^- E_t(x)$.

We write

$$e(x) := \sup_{t\in[0,T]} E_t(x).$$

Note that from the Lipschitz property it follows that there exists a constant $C_2 > 0$ such that for all $x \in Dom(E)$

$$e(x) \le C_2 \inf_{t\in[0,T]} E_t(x). \tag{204}$$

Approximation

We fix a time step $h > 0$ and subdivide the interval $[0, T]$ into

$$\mathcal{P}_h := \{t_0 = 0 < t_1 < \cdots < t_{N-1} < T \leq t_N\}, \qquad t_n = nh, N \in \mathbb{N}.$$

For $0 \leq t \leq T$ we define the piecewise constant interpolants $\overline{h}(t)$ and $\underline{h}(t)$ associated with the partition $\mathcal{P}_h$ in the following way;

$$\overline{h}(0) = 0 = \underline{h}(0), \text{ and for } t \in (t_{n-1}, t_n] \quad \overline{h}(t) = t_n, \quad \underline{h}(t) = t_{n-1}. \tag{205}$$

The definition implies that $\overline{h}(t) \searrow t$ and $\underline{h}(t) \nearrow t$ if $h \searrow 0$.

For a given initial value $\bar{x}$ we recursively define a sequence (x_n^h) of minimizers by

$$x_0^h := \bar{x}, \qquad x_n^h := \arg\min_x \Big\{ E_{t_n}(x) + \frac{1}{2h} ||x - x_{n-1}^h||_{t_n}^2 \Big\}. \tag{206}$$

We can argue as in the proof of Proposition 4.13 and directly obtain for every $\bar{x} \in Dom(E)$ and $h > 0$ a (unique) solution to the minimization problem (206). As in section 4.3.3 we define piecewise constant interpolants by setting

$$\bar{x}_t^h := x_n^h \text{ for } t \in (t_{n-1}, t_n], \qquad \underline{x}_t^h := x_{n-1}^h \text{ for } t \in (t_{n-1}, t_n],$$

and moreover, the piecewise linear interpolant

$$x_t^h = \frac{t - t_{n-1}}{h} x_n^h + \frac{t_n - t}{h} x_{n-1}^h \text{ for } t \in [t_{n-1}, t_n).$$

For $t \in (t_{n-1}, t_n)$ we denote the time derivative of $t \mapsto x_t^h$ by $\dot{x}_t^h$.

Recall that the variational interpolation is a map $t \to \tilde{x}_t^h$ defined by

$$\tilde{x}_t^h = \arg\min_x \Big\{ E_t(x) + \frac{1}{2r} ||x - x_{n-1}^h||_{t_n}^2 \Big\},$$
$$\text{for } t = t_{n-1} + r \in (t_{n-1}, t_n],$$

and $\tilde{x}_0^h = \bar{x}$. Finally we define $t \mapsto \tilde{v}_t^h$ by

$$\tilde{v}_t^h := \frac{\tilde{x}_t^h - x_{n-1}^h}{t - t_{n-1}} \quad \forall t \in (t_{n-1}, t_n].$$

As in Section 4.3.3, in order to extract a converging subsequence, we proof a priori estimates on the discrete solutions. The proof is along the lines of Proposition 6.3 in [53].

Proposition 4.36. *The following inequality holds for the interpolants $\bar{x}^h$, x^h, $\tilde{x}^h$ and $\tilde{v}^h$*

$$\begin{aligned} E_{\overline{h}(t)}(\bar{x}_t^h) + \frac{1}{2} \int_{\overline{h}(s)}^{\overline{h}(t)} ||\dot{x}_r^h||_{\overline{h}(r)}^2 dr + \frac{1}{2} \int_{\overline{h}(s)}^{\overline{h}(t)} ||\tilde{v}_r^h||_{\overline{h}(r)}^2 dr \\ \leq E_{\overline{h}(s)}(\bar{x}_s^h) + C_1 \int_{\overline{h}(s)}^{\overline{h}(t)} e(\tilde{x}_r^h) dr. \end{aligned} \tag{207}$$

In particular there exists a constant M such that for all $h > 0$

$$\sup_{t\in(0,T)} e(\bar{x}_t^h) \leq M, \qquad \sum_{n=0}^{N} \frac{1}{h}||x_n^h - x_{n-1}^h||_{t_n}^2 \leq M, \tag{208}$$

$$\int_0^T ||\dot{x}_r^h||_{\bar{h}(r)}^2 dr \leq M, \qquad \int_0^T ||\tilde{v}_r^h||_{\bar{h}(r)}^2 dr \leq M. \tag{209}$$

Moreover

$$||\tilde{x}_t^h - \underline{x}_t^h||^2 \in O(h), \quad ||x_t^h - \bar{x}_t^h||^2 \in O(h), \quad ||\bar{x}_t^h - \underline{x}_t^h||^2 \in O(h). \tag{210}$$

Proof. Consider the map

$$r \mapsto J_{s,r}(y) := \min_x \Big\{ E_{s+r}(x) + \frac{1}{2r}||x-y||_{s+h}^2 \Big\}$$

for a given $s \in [0,T]$, $y \in D$, $0 < r < T-s$. We claim that this map is differentiable almost everywhere in $(0, T-s)$ and for every $r_0 \in (0, T-s)$ for the minimizer $(0, r_0] \ni r \mapsto x_r$ it holds

$$\begin{aligned} \frac{1}{2r_0}||x_{r_0} - y||_{s+h}^2 + \int_0^{r_0} \frac{1}{2r^2}||x_r - y||_{s+h}^2 dr + E_{s+r_0}(x_{r_0}) \\ \leq E_s(y) + C_1 \int_0^{r_0} e(x_r) dr. \end{aligned} \tag{211}$$

Indeed, arguing similar as in (187) we obtain that for $r_1 < r_2 \in (0, T-s)$

$$\begin{aligned} &J_{s,r_2}(y) - J_{s,r_1}(y) - (E_{s+r_2}(x_{r_1}) - E_{s+r_1}(x_{r_1})) \\ &\leq -\frac{1}{2r_1r_2}(r_2 - r_1)||x_{r_1} - y||_{s+h}^2 \leq 0, \end{aligned} \tag{212}$$

hence the map $r \mapsto J_{s,r}(y)$ is the sum of a locally Lipschitz and of a nonincreasing function

$$J_{s,r_2}(y) \leq J_{s,r_1}(y) + (r_2 - r_1)C_1 e(x_{r_1}),$$

and differentiable almost everywhere. So let $r \in (0, T-s)$ be a differentiable point of $r \mapsto J_{s,r}(y)$. Then with (212) we get

$$\begin{aligned} &\frac{d}{dr}J_{s,r}(y) + \frac{1}{2r^2}||x_r - y||_{s+h}^2 \\ =& \lim_{h\to 0} \Big(\frac{J_{s,r+h}(y) - J_{s,r}(y)}{h} + \frac{1}{2(r+h)r}||x_r - y||_{s+h}^2 \Big) \\ \leq& \liminf_{h\to 0} \frac{E_{s+r+h}(x_r) - E_{s+r}(x_r)}{h} \leq C_1 e(x_r), \end{aligned}$$

and integrating from 0 to r_0 gives us (211).

Applying (211) with $t \in (t_{n-1}, t_n]$, $y = x_{n-1}^h$, $s = t_{n-1}$ and $r_0 = t - s$ we obtain for $\tilde{x}_t^h$

$$\frac{1}{2(t-t_{n-1})}||\tilde{x}_t^h - x_{n-1}^h||_{t_n}^2 + \int_{t_{n-1}}^t \frac{1}{2(r-t_{n-1})^2}||\tilde{x}_r^h - x_{n-1}^h||_{t_n}^2 dr + E_t(\tilde{x}_t^h) \tag{213}$$

$$\leq E_{t_{n-1}}(x_{n-1}^h) + C_1 \int_{t_{n-1}}^t e(\tilde{x}_r^h) dr. \tag{214}$$

Inserting $t = t_n$ we get for the interpolants x_t^h, $\tilde{v}_t^h$

$$\begin{aligned}&\frac{1}{2}\int_{t_{n-1}}^{t_n} ||\dot{x}_r^h||_{t_n}^2 dr + \int_{t_{n-1}}^{t_n} \frac{1}{2}||\tilde{v}_r^h||_{t_n}^2 dr + E_{t_n}(\tilde{x}_{t_n}^h) \\ &\leq E_{t_{n-1}}(x_{n-1}^h) + C_1 \int_{t_{n-1}}^{t_n} e(\tilde{x}_r^h) dr.\end{aligned} \tag{215}$$

Summing over the partition we end up with (207).

Note that the minimality and (204) imply that for $r \in (t_{n-1}, t_n]$, $t = t_{n-1} + r$

$$e(\bar{x}_{t_{n-1}}^h) \geq E_t(\bar{x}_{t_{n-1}}^h) \geq \frac{1}{2h}||\tilde{x}_r^h - \bar{x}_{t_{n-1}}^h||_{t_n}^2 + E_t(\tilde{x}_r^h) \geq E_t(\tilde{x}_r^h) \geq \frac{1}{C_2} e(\tilde{x}_r^h),$$

and hence with (215) we can estimate

$$E_{t_n}(\bar{x}_{t_n}^h) \leq E_{t_{n-1}}(x_{n-1}^h) + C_1 C_2 \int_{t_{n-1}}^{t_n} e(\bar{x}_{t_{n-1}}^h) dr.$$

Summing over the partitions and applying (204) once more we obtain for some constant $C > 0$

$$e(\bar{x}_{t_n}^h) \leq C(E_0(x_0^h) + \int_0^{t_n} e(\underline{x}_r^h) dr).$$

We obtain the first inequality in (208) by applying a discrete Gronwall argument (see e.g. [54, Lemma 4.5]). It directly follows that the right-hand side of (207) is bounded and (209) holds. The second inequality in (208) is a direct consequence of the first estimate in (209).

In order to show the first statement in (210) recall that (213) together with (208) implies (with some different constant M)

$$||\tilde{x}_t^h - \underline{x}_t^h||^2 \leq 2hM.$$

The other two assertions in (210) follow from (209) via Hölder's inequality

$$||x_t^h - x_s^h|| \leq \int_s^t ||\dot{x}_r^h|| dr \leq \sqrt{M(t-s)} \qquad \forall 0 < s < t < T.$$

□

The following result provides the compactness of the approximate solutions.

Proposition 4.37. *For every sequence of time-steps $(h_j)_{j \in \mathbb{N}}$ such that $h_j \to 0$ as $j \to \infty$ there exists a subsequence h_j (not relabeled) and an absolutely continuous curve $(x_t) \in \mathrm{AC}^2([0,T]; H)$ and such that*

$$x_t^{h_j} \to x_t \text{ in } \mathcal{C}^0([0,T]; H),$$

and

$$\dot{x}_t^{h_j} \rightharpoonup \dot{x}_t \text{ in } L^2([0,T]; H).$$

Moreover for each $t \in [0,T]$, $\bar{x}_t^{h_j}, \tilde{x}_t^{h_j} \to x_t$ in H.

Proof. Let $0 < g \le h$ be two stepsizes and $\{t_n^g\}_{n=0}^{N_g}$, $\{t_n^h\}_{n=0}^{N_h}$ the corresponding partitions of the interval $[0,1]$. Let $\{x_n^h\}_{n=0}^{N_h}$ and $\{x_n^g\}_{n=0}^{N_g}$ be the solution to the minimizing problem (206) with respect to the stepsizes h and g respectively with initial condition x_0^h and x_0^g. The Euler-Lagrange equation of x_n^h is

$$\frac{x_n^h - x_{n-1}^h}{h} \in -D_{t_n^h}^- E_{t_n^h}(x_n^h),$$

i.e.

$$h^{-1}\langle x_n^h - x_{n-1}^h, x_n^h - y\rangle_{t_n^h} + E_{t_n^h}(x_n^h) - E_{t_n^h}(y) \le 0 \quad \forall y \in H.$$

Inserting the definition of the piecewise linear interpolation x_t^h at $t \in (t_{n-1}^h, t_n^h)$

$$\begin{aligned} &\langle \dot{x}_t^h, x_t^h - y\rangle_{t_n^h} + E_{t_n^h}(x_t^h) - E_{t_n^h}(y) \\ &\le (t - t_n^h)\Big(||\dot{x}_t^h||_{t_n^h}^2 + \frac{1}{h}(E_{t_n^h}(x_n^h) - E_{t_n^h}(x_{n-1}^h))\Big) \quad \forall y \in H, \end{aligned} \tag{216}$$

where we applied the convexity of $x \mapsto E_t(x)$. The same argumentation for x_t^g at $t \in (t_{m-1}^g, t_m^g)$ yields

$$\begin{aligned} &\langle \dot{x}_t^g, x_t^g - y\rangle_{t_m^g} + E_{t_m^g}(x_t^g) - E_{t_n^g}(y) \\ &\le (t - t_m^g)\Big(||\dot{x}_t^g||_{t_m^g}^2 + \frac{1}{g}(E_{t_m^g}(x_m^g) - E_{t_m^g}(x_{m-1}^g))\Big) \quad \forall y \in H. \end{aligned} \tag{217}$$

For $t \in (t_{n-1}^h, t_n^h) \cap (t_{m-1}^g, t_m^g)$ we get by putting $y = x_t^g$ into (216) and $y = x_t^h$ into (217) and adding them

$$\begin{aligned} &\langle \dot{x}_t^h, x_t^h - x_t^g\rangle_{t_n^h} + \langle \dot{x}_t^g, x_t^g - x_t^h\rangle_{t_m^g} \\ &+ E_{t_n^h}(x_t^h) - E_{t_n^g}(x_t^h) + E_{t_m^g}(x_t^g) - E_{t_n^h}(x_t^g) \\ &\le (t - t_n^h)\Big(||\dot{x}_t^h||_{t_n^h}^2 + \frac{E_{t_n^h}(x_n^h) - E_{t_n^h}(x_{n-1}^h)}{h}\Big) \\ &+ (t - t_m^g)\Big(||\dot{x}_t^g||_{t_m^g}^2 + \frac{E_{t_m^g}(x_m^g) - E_{t_m^g}(x_{m-1}^g)}{g}\Big). \end{aligned} \tag{218}$$

The Lipschitz property (161) of the metric together with the polarization identity gives

$$\begin{aligned} &\langle \dot{x}_t^h, x_t^h - x_t^g\rangle_{t_n^h} + \langle \dot{x}_t^g, x_t^g - x_t^h\rangle_{t_m^g} \\ &\ge \langle \dot{x}_t^h - \dot{x}_t^g, x_t^h - x_t^g\rangle_{t_m^g} - L|t_n^h - t_m^g|\Big(\langle \dot{x}_t^h, x_t^h - x_t^g\rangle_{t_m^g} + \frac{1}{2}||\dot{x}_t^h - (x_t^h - x_t^g)||_{t_m^g}^2\Big), \end{aligned}$$

while the Lipschitz property of the energy yields

$$E_{t_n^h}(x_t^h) - E_{t_n^g}(x_t^h) + E_{t_m^g}(x_t^g) - E_{t_n^h}(x_t^g) \ge -C_1|t_n^h - t_m^g|\Big(E_{t_n^h}(x_t^h) + E_{t_n^h}(x_t^g)\Big).$$

Inserting these two inequalities into (218) we find

$$
\begin{aligned}
&\frac{d}{dt}||x_t^h - x_t^g||_{t_m^g}^2 = \langle \dot{x}_t^h - \dot{x}_t^g, x_t^h - x_t^g \rangle_{t_m^g} \\
&\leq (t - t_n^h)\Big(||\dot{x}_t^h||_{t_n^h}^2 + \frac{E_{t_n^h}(x_n^h) - E_{t_n^h}(x_{n-1}^h)}{h} \Big) \\
&+ (t - t_m^g)\Big(||\dot{x}_t^g||_{t_m^g}^2 + \frac{E_{t_m^g}(x_m^g) - E_{t_m^g}(x_{m-1}^g)}{g} \Big) \\
&+ L|t_n^h - t_m^g|\Big(\langle \dot{x}_t^h, x_t^h - x_t^g \rangle_{t_m^g} + \frac{1}{2}||\dot{x}_t^h - (x_t^h - x_t^g)||_{t_m^g}^2 \Big) \\
&+ C_1|t_n^h - t_m^g|\Big(E_{t_n^h}(x_t^h) + E_{t_n^h}(x_t^g) \Big).
\end{aligned}
\tag{219}
$$

Integrating (219) on the interval $(t_{n-1}^h \vee t_{m-1}^g, t)$ we can estimate

$$
\begin{aligned}
&||x_t^h - x_t^g||_{t_m^g}^2 - ||x_{t_{n-1}^h \vee t_{m-1}^g}^h - x_{t_{n-1}^h \vee t_{m-1}^g}^g||_{t_m^g}^2 \\
&\leq h \int_{t_{n-1}^h \vee t_{m-1}^g}^t \Big(- ||\dot{x}_r^h||_{t_n^h}^2 + \frac{E_{t_n^h}(x_{n-1}^h) - E_{t_n^h}(x_n^h)}{h} \Big)\, dr \\
&+ g \int_{t_{n-1}^h \vee t_{m-1}^g}^t \Big(- ||\dot{x}_r^g||_{t_m^g}^2 + \frac{E_{t_m^g}(x_{m-1}^g) - E_{t_m^g}(x_m^g)}{g} \Big)\, dr \\
&+ L(h \wedge g) \int_{t_{n-1}^h \vee t_{m-1}^g}^t \Big(||\dot{x}_r^h||_{t_m^g}^2\, dr + ||x_r^h - x_r^g||_{t_m^g}^2 \Big)\, dr \\
&+ C_1(h \wedge g) \int_{t_{n-1}^h \vee t_{m-1}^g}^t \Big(E_{t_n^h}(x_r^h) + E_{t_n^h}(x_r^g) \Big)\, dr.
\end{aligned}
\tag{220}
$$

Summing over the partition $\{t_j^{h,g}\}_{j=0}^{N_h+N_g} = \{t_n^h\}_{n=0}^{N_h} \cup \{t_m^g\}_{m=0}^{N_g}$ and exploiting the Lipschitz property of $t \mapsto ||\cdot||_t$

$$
\begin{aligned}
||x_t^h - x_t^g||_{t_m^g}^2 \leq & ||x_0^h - x_0^g||_0^2 + Lg \sum_{j=1}^{n+m} ||x_{t_{j-1}^{h,g}}^h - x_{t_{j-1}^{h,g}}^g||_{g(t_{j-1}^{h,g})}^2 \\
&+ \sum_{j=1}^n \int_{t_{j-1}^h}^{t_j^h} (-h||\dot{x}_r^h||_{t_j^h}^2 + E_{t_j^h}(x_{j-1}^h) - E_{t_j^h}(x_j^h))\, dr \\
&+ \sum_{j=1}^m \int_{t_{j-1}^g}^{t_j^g} (-g||\dot{x}_r^g||_{t_j^g}^2 + E_{t_j^g}(x_{j-1}^g) - E_{t_j^g}(x_j^g))\, dr \\
&+ L(h \wedge g) \int_0^t \Big(||\dot{x}_r^h||_{g(r)}^2 dr + ||x_r^h - x_r^g||_{g(r)}^2 \Big)\, dr \\
&+ C_1(h \wedge g) \int_0^t (E_{h(r)}(x_r^h) + E_{h(r)}(x_r^g))\, dr.
\end{aligned}
$$

Applying once more the Lipschitz property of $t \mapsto E_t(x)$, we can further estimate

$$
\begin{aligned}
&||x_t^h - x_t^g||^2_{t_m^g} \leq ||x_0^h - x_0^g||_0^2 + Lg \sum_{j=1}^{n+m} ||x^h_{t^{h,g}_{j-1}} - x^g_{t^{h,g}_{j-1}}||^2_{g(t^{h,g}_{j-1})} \\
&+ h(E_0(x_0^h) - E_{t_n^h}(x_n^h)) + g(E_0(x_0^g) - E_{t_m^g}(x_m^g)) \\
&+ L(h \wedge g) \int_0^t ||\dot{x}_r^h||^2_{h(r)} dr + C(h \wedge g) \int_0^t ||x_r^h - x_r^g||^2_{g(r)} dr \\
&+ C_1(h \wedge g) \int_0^t (E_{h(r)}(x_r^h) + E_{h(r)}(x_r^g)) dr \\
&+ C_1 \sum_j \int_{t_{j-1}^h}^{t_j^h} (t_j^h - t_{j-1}^h) E_{t_j^h}(x_{j-1}^h) dr + C_1 \sum_k \int_{t_{k-1}^g}^{t_k^g} (t_k^g - t_{k-1}^g) E_{t_k^g}(x_{k-1}^g) dr.
\end{aligned}
\tag{221}
$$

From the positivity of E and from (209) as well as (208) we can deduce (with varying constants)

$$
\begin{aligned}
||x_t^h - x_t^g||^2_{t_m^g} \leq & ||x_0^h - x_0^g||_0^2 + Lg \sum_{j=1}^{n+m} ||x^h_{t^{h,g}_{j-1}} - x^g_{t^{h,g}_{j-1}}||^2_{g(t^{h,g}_{j-1})} \\
&+ hE_0(x_0^h) + gE_0(x_0^g) \\
&+ C(h \wedge g) + C(h \wedge g) \int_0^t ||x_r^h - x_r^g||^2_{g(r)} dr \\
&+ C(h \wedge g) + Ch + Cg \\
\leq & ||x_0^h - x_0^g||_0^2 + Lg \sum_{j=1}^{n+m} ||x^h_{t^{h,g}_{j-1}} - x^g_{t^{h,g}_{j-1}}||^2_{g(t^{h,g}_{j-1})} \\
&+ C(h+g) + C(h \wedge g).
\end{aligned}
\tag{222}
$$

The last inequality follows from

$$
\sup_r ||x_r^h||_{t^*} = \sup_n ||x_n^h||_{t^*} \leq \sup_n (\sqrt{2CnhM} + ||x_0^h||_{t^*}) \leq \sqrt{2CTM} + ||x_0^h||_{t^*},
$$

where we used the definition of $(x_r^h)_{r \in [0,T]}$ in the second equality, triangle inequality and Cauchy-Schwartz inequality in the second, $C = C(L)$ is the constant arising from the log-Lipschitz control (161) of the metric.

For h, g sufficiently small there exists a κ satisfying $1 - L(h \wedge g) \geq \frac{1}{\kappa} > 0$. Applying the discrete Gronwall lemma [54, Lemma 4.5] we finally obtain

$$
||x_t^h - x_t^g||^2_{t_m^g} \leq \kappa C(h+g) e^{\kappa(n+m)(h \wedge g)} \leq \kappa C(h+g) e^{2T\kappa}. \tag{223}
$$

This shows that if h_j is a vanishing sequence of stepsizes, $\{x^{h_j}\}_j \subset \mathcal{C}^0([0,T];H)$ is a Cauchy sequence. Since $\mathcal{C}^0([0,T];H)$ is a Banach space there exists a continuous curve $(x_t)_{t \in [0,T]}$ and a subsequence (not relabeled) such that $x_t^{h_j} \to x_t$ in $\mathcal{C}^0([0,T];H)$ as $j \to \infty$. From (210) it follows immediately that also $\tilde{x}_t^{h_j}, \bar{x}_t^{h_j}$ converge to x_t.

Since $\int_0^T ||\dot{x}_r^{h_j}||^2 dr \leq M$ we can extract a further subsequence (not relabeled) with

$$
\dot{x}^{h_j} \rightharpoonup u \text{ in } L^2([0,T];H)
$$

where u is some function in $L^2([0,T];H)$. As a consequence we obtain that the limit function $(x_t) \in \mathrm{AC}^2([0,T];H)$ since for all $0 < s < t < T$

$$||x_t - x_s||_{t^*} = \lim_{j\to\infty} ||x_t^{h_j} - x_s^{h_j}||_{t^*} = \lim_{j\to\infty} || \int_s^t \dot{x}_r^{h_j} dr||_{t^*} \le \int_s^t ||u_r||_{t^*} dr, \quad (224)$$

where t^* is an arbitrarily fixed timepoint in $[0,T]$. We still have to show that $u_r = \dot{x}_r$ almost everywhere. This follows again straightforward from the weak convergence of $\dot{x}^{h_j}$. Let $y \in H$, then

$$\langle x_t - x_s, y\rangle_{t^*} = \lim \langle x_t^{h_j} - x_s^{h_j}, y\rangle_{t^*} = \lim \langle \int_s^t \dot{x}_r^{h_j} dr, y\rangle_{t^*} = \langle \int_s^t u_r dr, y\rangle_{t^*}.$$

Since $y \in H$ is arbitrary we obtain

$$x_t - x_s = \int_s^t u_r dr,$$

and hence $\lim_{s\to t} \frac{x_t - x_s}{t-s} = u_t$ at every Lebesgue point of u.

□

Theorem 4.38. *Let E be as in the beginning of this section. Then for every $x \in Dom(E)$ there exists a unique map $t \mapsto x_t$ from $[0,T]$ to X with $\lim_{t\searrow 0} x_t = x$ such that*

$$\partial_t x_t \in -D_t^- E_t(x_t) \text{ for a.e. } t \in (0,T).$$

Proof. Recall that the minimizers of (206) with $x_0^h := x$ satisfy the Euler-Lagrange equation, that is in terms of the subdifferential of E, the piecewise linear interpolant x_t^h and the piecewise constant interpolant $\bar{x}_t^h$

$$\langle \dot{x}_t^h, \bar{x}_t^h - y\rangle_{t_n^h} + E_{t_n^h}(\bar{x}_t^h) - E_{t_n^h}(y) \le 0 \quad \forall y \in H, \text{ for every } t \in (t_{n-1}^h, t_n^h).$$

The log Lipschitz property together with the polarization identity gives then for all $y \in H$ and almost every $t \in [0,T]$

$$\langle \dot{x}_t^h, \bar{x}_t^h - y\rangle_t + E_t(\bar{x}_t^h) - E_t(y) \le Lh(||\dot{x}_t^h||_t^2 + ||\bar{x}_t^h - y||_t^2) + C_1 h(e(\bar{x}_t^h) + e(y)).$$

Integrating this inequality over the interval (s,t) for some $0 < s < t < T$ we deduce

$$\begin{aligned} &\int_s^t \langle \dot{x}_r^h, \bar{x}_r^h - y\rangle_r \, dr + \int_s^t E_r(\bar{x}_r^h) - E_r(y)\, dr \\ &\le Lh \int_s^t ||\dot{x}_r^h||^2 + ||\bar{x}_r^h - y||_r^2 \, dr + C_1 h \int_s^t (e(\bar{x}_r^h) + e(y)) dr. \end{aligned} \quad (225)$$

Applying Proposition 4.37 we get existence of a subsequence and a curve $(x_t) \in \mathrm{AC}^2([0,T];H)$ such that $\bar{x}_t^h \to x_t$ in $\mathcal{C}^0([0,T];H)$ and $\dot{x}_t^h \to \dot{x}_t$ weakly in $L^2([0,T];H)$. Hence we get for all $y \in Dom(E)$

$$
\begin{aligned}
&\int_s^t \langle \dot{x}_r, x_r - y\rangle_r dr + \int_s^t E_r(x_r) - E_r(y) dr \\
&\le \liminf_{h\to 0} \int_s^t \langle \dot{x}_r^h, \bar{x}_r^h - y\rangle_r dr + \liminf_{h\to 0} \int_s^t E_r(\bar{x}_r^h) - E_r(y) dr \\
&\le \liminf_{h\to 0} \Big\{ \int_s^t \langle \dot{x}_r^h, \bar{x}_r^h - y\rangle_r dr + \int_s^t E_r(\bar{x}_r^h) - E_r(y) dr \Big\} \\
&\le \liminf_{h\to 0} \Big\{ Lh \int_s^t ||\dot{x}_r^h||^2 + ||\bar{x}_r^h - y||_r^2 \, dr + C_1 h \int_s^t (e(\bar{x}_r^h) + e(y)) dr \Big\} \\
&\le 0,
\end{aligned}
$$

where we applied Fatou's Lemma and the lower semicontinuity of $x \mapsto E_t(x)$ in the first inequality, estimate (225) in the third inequality and the non-negativity of $E_t(y)$, (208) and (209) in the last. Dividing by $t - s$ and letting $s \to t$ we infer from the Lebesgue differentiation theorem that

$$
\langle \dot{x}_t, x_t - y\rangle_t + E_t(x_t) - E_t(y) \le 0
$$

for almost every $t \in (0, T)$ and $y \in X$.

Since $\bar{x}_t^h$ converges to x_t for every t we clearly have that $\lim_{t \searrow 0} x_t = x$.

Suppose there exists two absolutely continuous curves (x_t), $(\tilde{x}_t)_{t\in[0,T]}$ such that for every $y \in X$ and almost every $t \in [0, T]$

$$
\begin{aligned}
&\langle \dot{x}_t, x_t - y\rangle_t + E_t(x_t) - E_t(y) \le 0, \\
&\langle \dot{\tilde{x}}_t, \tilde{x}_t - y\rangle_t + E_t(\tilde{x}_t) - E_t(y) \le 0
\end{aligned}
$$

with $\lim_{t\searrow 0} x_t = \lim_{t\searrow 0} \tilde{x}_t = 0$. Inserting $\tilde{x}_t$ for y into the first inequality and x_t for y into the second we obtain by adding and using (203)

$$
\partial_s \frac{1}{2} ||x_s - \tilde{x}_s||_t^2 \Big|_{s=t} = \langle \dot{x}_t - \dot{\tilde{x}}_t, x_t - \tilde{x}_t\rangle_t \le 0.
$$

From the log-Lipschitz continuity of the metric we deduce

$$
\partial_s \frac{1}{2} ||x_s - \tilde{x}_s||_s^2 \Big|_{s=t} \le L ||x_t - \tilde{x}_t||_t^2.
$$

Applying Gronwall's inequality we conclude $||x_t - \tilde{x}_t||_t^2 \le e^{2Lt} ||x_0 - \tilde{x}_0||_0^2 = 0$ for almost every $t \in [0, T]$ and hence for every $t \in [0, T]$ by continuity. This proves uniqueness. □

4.6 The Heat Equation on Time-dependent Metric Measure Spaces

Let $(X, d_t, m_t)_{t\in[0,T]}$ be a family of Polish metric measure space. We always assume that (161) holds and that there exists a reference measure $m \in \mathcal{P}(X)$ such that $m_t = e^{-f_t} m$ with Borel functions f_t satisfying $|f_t(x)| \le C$ and

$$
|f_t(x) - f_s(x)| \le L^* |t - s|, \quad |f_t(x) - f_t(y)| \le C d_t(x, y). \tag{226}
$$

Let us denote *Cheeger's energy* by $\mathrm{Ch}_t \colon L^2(X, m_t) \to [0, \infty]$

$$\mathrm{Ch}_t(u) = \frac{1}{2} \inf \left\{ \liminf_{n\to\infty} \int_X (\mathrm{lip}_t u_n)^2 dm_t : u_n \in \mathrm{Lip}(X), \int |u_n - u|^2 dm_t \to 0 \right\},$$

where $\mathrm{lip}_t u$ denotes the local slope defined by

$$\mathrm{lip}_t u(x) := \limsup_{y\to x} \frac{|u(x) - u(y)|}{d_t(x, y)}.$$

By making use of the *minimal relaxed gradient* $|\nabla_t u|_*$ ([5, Definition 4.2]), this functional admits the integral representation

$$\mathrm{Ch}_t(u) = \frac{1}{2} \int_X |\nabla_t u|_*^2 dm_t,$$

set equal to $+\infty$ if u has no relaxed slope. This defines a convex and lower semicontinuous functional in $L^2(X, m_t)$ [5, Theorem 4.5].

Lemma 4.39. *Fix $t \in [0, T]$ and let $u \in Dom(\mathrm{Ch}_t)$. Then presuming* (161)

$$|\nabla_t u|_* \leq e^{L|t-s|} |\nabla_s u|_* \quad m\text{-a.e. in } X, \quad \forall s \in [0, T].$$

Proof. Since $u \in Dom(\mathrm{Ch}_t)$ we know $u \in Dom(\mathrm{Ch}_s)$ as well and there exist bounded Borel Lipschitz functions $u_n \in L^2(X, m_s)$ such that

$$u_n \to u, \quad \mathrm{lip}_s u_n \to |\nabla_s u|_* \text{ strongly in } L^2(X, m_s),$$

see e.g. [5, Lemma 4.3 (c)]. This implies that $e^{L|t-s|}|\nabla_s u|_*$ is a relaxed d_t-gradient since

$$u_n \to u, \quad e^{L|t-s|}\mathrm{lip}_s u_n \to e^{L|t-s|}|\nabla_s u|_* \text{ strongly in } L^2(X, m_t)$$

and

$$|\nabla_t u_n|_* \leq e^{L|t-s|}\mathrm{lip}_s u_n,$$

c.f. [5, Lemma 4.3. (a)]. Thus Lemma 4.4 in [5] yields the assertion. □

The domain of Cheeger's energy endowed with the norm

$$\sqrt{||f||^2_{L^2(X,m_t)} + \mathrm{Ch}_t(f)}$$

is a Banach space, cf. [20, Theorem 2.7]. In the following we additionally impose that for each t the space (X, d_t, m_t) is *infinitesimally Hilbertian*, i.e. Cheeger's energy Ch_t defines a quadratic form. In particular the domain is a separable Hilbert space and Lipschitz functions are dense, see [6]. In this case we will denote by $\mathcal{E}_t$ the associated *Dirichlet form*, which is the unique bilinear symmetric form satisfying

$$\mathcal{E}_t(u, u) = 2\mathrm{Ch}_t(u) \quad \forall u \in Dom(\mathrm{Ch}_t) \cap L^2(X, m_t).$$

Moreover $\mathcal{E}_t$ is strongly local [6, Proposition 4.14], i.e.

$$u, v \in Dom(\mathcal{E}), v \text{ constant on } \{u \neq 0\} \Rightarrow \mathcal{E}(u, v) = 0,$$

and admits the integral representation

$$\mathcal{E}_t(u,v) = \int \nabla_t u \cdot \nabla_t v \, dm_t \quad u, v \in Dom(\mathrm{Ch}_t) \cap L^2(X, m_t),$$

where

$$\nabla_t u \cdot \nabla_t v := \lim_{\varepsilon \searrow 0} \frac{|\nabla_t (u + \varepsilon v)|_*^2 - |\nabla_t u|_*^2}{2\varepsilon}$$

and the limit is understood in $L^1(X, m_t)$, see [6, Proposition 4.14].

We define the Laplace operator Δ_t as the generator of $\mathcal{E}_t$, i.e. as the unique non-positive self adjoint operator on $L^2(X, m_t)$ with domain $Dom(\Delta_t) \subset Dom(\mathrm{Ch}_t)$ and

$$-\int_X \Delta_t uv \, dm_t = \mathcal{E}_t(u,v) \quad \forall u \in Dom(\Delta_t), v \in Dom(\mathrm{Ch}_t) \cap L^2(X, m_t).$$

Due to our assumptions the sets $L^2(X, m_t)$ and $Dom(\mathrm{Ch}_t)$ do not depend on t. We set $\mathcal{F} = Dom(\mathcal{E})$ and $\mathcal{H} = L^2(X, m_t)$ and define for $0 \leq s < \tau \leq T$ the Hilbert space

$$\mathcal{F}_{(s,\tau)} = L^2((s,\tau) \to \mathcal{F}) \cap H^1((s,\tau) \to \mathcal{F}^*),$$

equipped with the norm $(\int_s^\tau ||u_t||_{\mathcal{F}}^2 + ||\partial_t u_t||_{\mathcal{F}^*}^2 \, dt)^{1/2}$. According to Lemma 10.3 in [52] we have $\mathcal{F}_{(s,\tau)} \subset \mathcal{C}([s,\tau] \to \mathcal{H})$.

Definition 4.40. *A function u is called* solution to the heat equation

$$\partial_t u = \Delta_t u \text{ on } (s,\tau) \times X$$

if $u \in \mathcal{F}_{(s,\tau)}$ and if for all $v \in \mathcal{F}_{(s,\tau)}$

$$-\int_s^\tau \mathcal{E}_t(u_t, v_t) dt = \int_s^\tau \langle \partial_t u_t, w_t e^{-f_t} \rangle_{\mathcal{F}^*, \mathcal{F}} dt, \tag{227}$$

where $\langle \cdot, \cdot \rangle_{\mathcal{F}^*, \mathcal{F}}$ denotes the dual pairing.

A function v is called solution to the adjoint heat equation

$$-\Delta_s v + \partial_s f \cdot v = \partial_s v \qquad \text{on } (\sigma, t) \times X$$

if $v \in \mathcal{F}_{(\sigma,t)}$ and if for all $w \in \mathcal{F}_{(\sigma,t)}$

$$\int_\sigma^t \mathcal{E}_s(v_s, w_s) ds + \int_\sigma^t \int_X v_s \cdot w_s \cdot \partial_s f_s \, dm_s \, ds = \int_\sigma^t \langle \partial_s v_s, w_s e^{-f_s} \rangle_{\mathcal{F}, \mathcal{F}^*} \, ds.$$

We assume that each static space (X, d_t, m_t) satisfies the Riemannian curvature-dimension condition $\mathrm{RCD}(K, N')$, i.e. (X, d_t, m_t) satisfies the curvature-dimension condition $\mathrm{CD}(K, N')$ in the sense of [5, 57, 43] for some finite numbers K, N' and it is infinitesimally Hilbertian.

By virtue of Theorem 2.29 and Theorem 2.31 we have under combination of the previous assumptions on (X, d_t, m_t) existence and uniqueness to solutions of the heat and the adjoint heat equation with initial condition $u_s = h \in \mathcal{H}$ and terminal condition $v_t = h \in \mathcal{H}$ respectively. We denote these solutions by

$$u_t(x) = P_{t,s} h(x), \qquad v_s(x) = P_{t,s}^* h(x).$$

Both solutions, called heat flow and adjoint heat flow respectively, admit the following kernel representations

$$P_{t,s}h(x) = \int p_{t,s}(x,y)h(y)dm_s(y), \tag{228}$$

$$P^*_{t,s}h(y) = \int p_{t,s}(x,y)h(x)dm_t(x), \tag{229}$$

where $\int p_{t,s}(x,y)\,dm_s(y) = 1$, and $p_{t,r}(x,z) = \int p_{t,s}(x,y)p_{s,r}(y,z)\,dm_s(y)$.

4.6.1 Identification of the Forward Adjoint Heat Flow with the Dynamic EDI-Gradient Flow for the Entropy

We consider the adjoint heat flow $(\rho_t)_{0\le t\le T}$ parametrized forwards in time, i.e. solving

$$\partial_t\rho_t = \Delta_t\rho_t + \rho_t\partial_t f_t \qquad \text{on } (0,T)\times X$$

with nonnegative initial data $\rho_0 = h$. We identify (ρ_t) with the dynamic EDI-gradient flow (μ_t) of S via $\mu_t = \rho_t m_t$. In order to show this we adapt the strategy in [5]. We prove that $\mu_t = \rho_t m_t$ is a dynamic EDI-gradient flow of S. From the uniqueness it follows that both flows coincide.

Lemma 4.41. *Let $h \in \mathcal{H}$ and (ρ_t) be the solution to the forward adjoint heat flow on $(0,T)\times X$ with $\rho_0 = h$.*

1. *The flow (ρ_t) is mass preserving, i.e.*

$$\int \rho_t\,dm_t = \int h\,dm_0 \quad \forall 0\le t\le T. \tag{230}$$

2. *If $e\colon \mathbb{R}\to[0,\infty]$ is a convex lower semicontinuous function and e' is locally Lipschitz in $\mathbb{R}$, it holds for $\mathcal{L}^1$-a.e. $t\in(0,T)$*

$$\frac{d}{dt}\int e(\rho_t)\,dm_t = -\int e''(\rho_t)|\nabla_t\rho_t|^2_*\,dm_t + \int \partial_t f_t(\rho_t e'(\rho_t) - e(\rho_t))\,dm_t. \tag{231}$$

Proof. Since the measure is finite, $1\in\mathcal{H}$, and hence

$$\begin{aligned}\partial_t\int\rho_t\,dm_t &= \langle\partial_t\rho_t, e^{-f_t}\rangle_{\mathcal{F}^*,\mathcal{F}}\\ &=\mathcal{E}_t(\rho_t,1) + \int\rho_t(\partial_t\rho_t)\,dm_t - \int\rho_t(\partial_t\rho_t)\,dm_t = 0,\end{aligned}$$

which shows the first assertion.

In order to prove (231) we assume by a standard approximation that e' is bounded and globally Lipschitz, cf. [5, Theorem 4.16]. Since e is convex and

$\rho \in \mathcal{F}_{0,T}$ we have for $t_0 < t_1$

$$
\begin{aligned}
&\int e(\rho_{t_1})\, dm_{t_1} - \int e(\rho_{t_0})\, dm_{t_0} \\
\geq &\int e'(\rho_{t_0})(\rho_{t_1} - \rho_{t_0})\, dm_{t_1} + \int e(\rho_{t_0})\, d(m_{t_1} - m_{t_0}) \\
= &\int_{t_0}^{t_1} \langle \partial_t \rho_t, e'(\rho_{t_0}) e^{-f_{t_1}} \rangle_{\mathcal{F}^*, \mathcal{F}}\, dt - \int_{t_0}^{t_1} \int e(\rho_{t_0}) \partial_t f_t\, dm_t\, dt \\
\geq &\int_{t_0}^{t_1} \Big(-\frac{1}{2} ||\partial_t \rho_t||^2_{\mathcal{F}^*} - \frac{1}{2} ||e'(\rho_{t_0}) e^{-f_{t_1}}||^2_{\mathcal{F}} - \int_{t_0}^{t_1} \int e(\rho_{t_0}) \partial_t f_t\, dm_t \Big)\, dt,
\end{aligned}
$$

which is integrable. Changing the roles of t_0 and t_1 shows that $s \mapsto \int e(\rho_t)\, dm_t$ is absolutely continuous. Then, since $\rho \in \mathcal{F}_{(0,T)}$, we deduce from the mean value theorem for a.e. t

$$
\begin{aligned}
&\lim_{h \to 0} \frac{1}{h} \Big(\int e(\rho_{t+h})\, dm_{t+h} - \int e(\rho_t)\, dm_t \Big) \\
= &\lim_{h \to 0} \frac{1}{h} \int (e(\rho_{t+h}) - e(\rho_t)) e^{-f_{t+h}}\, dm + \lim_{h \to 0} \frac{1}{h} \int e(\rho_t)(e^{-f_{t+h}} - e^{-f_t})\, dm \\
= &\lim_{h \to 0} \int e'(\rho_t) \frac{\rho_{t+h} - \rho_t}{h}\, dm_t - \int e(\rho_t) \partial_t f_t\, dm_t \\
= &\langle \partial_t \rho_t, e'(\rho_t) e^{-f_t} \rangle_{\mathcal{F}^*, \mathcal{F}} - \int e(\rho_t) \partial_t f_t\, dm_t,
\end{aligned}
$$

cf. [41, Corollary 5.5], [8, Lemma 12.3]. Since ρ is a solution to the forward adjoint heat equation we have

$$
\begin{aligned}
\langle \partial_s \rho_s, e'(\rho_s) e^{-f_s} \rangle_{\mathcal{F}^*, \mathcal{F}} = &- \mathcal{E}_s(\rho_s, e'(\rho_s)) + \int \rho_s e'(\rho_s) \partial_s f_s\, dm_s \\
= &- \int e''(\rho_s) |\nabla_s \rho_s|_*^2 dm_s + \int \rho_s e'(\rho_s) \partial_s f_s\, dm_s,
\end{aligned}
$$

which proves (231). $\square$

Proposition 4.42. *Let $(\rho_t)_{0 \leq t \leq T}$ be the solution of the forward adjoint heat equation with nonnegative initial datum $h \in \mathcal{H}$. Then it holds*

$$
\begin{aligned}
\int_0^t \int_{\{\rho_r > 0\}} \frac{|\nabla_r \rho_r|_*^2}{\rho_r}\, dm_r\, dr \leq &\int h \log h\, dm_0 + \int h\, dm_0 \\
&- m_t(X) + \int_0^t \int (\partial_r f_r) \rho_r\, dm_r\, dr,
\end{aligned}
\tag{232}
$$

and the map $t \mapsto \int \rho_t \log \rho_t dm_t$ is locally absolutely continuous in $[0, T]$ and

$$
\frac{d}{dt} \int \rho_t \log \rho_t\, dm_t = - \int_{\{\rho_t > 0\}} \frac{|\nabla_t \rho_t|_*^2}{\rho_t}\, dm_t + \int (\partial_t f_t) \rho_t\, dm_t \tag{233}
$$

for a.e. $t \in [0, T]$.

Proof. Since all solutions admit a kernel representation, we have $\rho_t \geq 0$ for all $t \in (0,T)$. Applying formula (231) to $\rho_t + \varepsilon$ we get

$$\frac{d}{dt}\int (\rho_t+\varepsilon)\log(\rho_t+\varepsilon)\,dm_t = -\int \frac{|\nabla_t \rho_t|_*^2}{\rho_t+\varepsilon}\,dm_t + \int \partial_t f_t(\rho_t+\varepsilon)\,dm_t.$$

Integrating from 0 to t and letting ε go to 0, we obtain by applying dominated and monotone convergence

$$\begin{aligned}&\int \rho_t \log \rho_t\,dm_t - \int \rho_0 \log \rho_0\,dm_0\\ &= \int_0^t - \int_{\{\rho_r>0\}} \frac{|\nabla_r \rho_r|_*^2}{\rho_r}\,dm_r + \int (\partial_r f_r)\rho_r\,dm_r\,dr.\end{aligned} \tag{234}$$

Using $\rho\log\rho \geq \rho - 1$ and the conservation of total mass (230) leads to

$$\begin{aligned}\int_0^t \int_{\{\rho_r>0\}} \frac{|\nabla_r \rho_r|_*^2}{\rho_r}\,dm_r\,dr \leq &\int h\log h\,dm_0 + \int h\,dm_0\\ &-m_t(X) + \int_0^t \int (\partial_r f_r)\rho_r\,dm_r\,dr,\end{aligned}$$

which proves (232). As a consequence from (232) and (234) we get the local absolute continuity of $s \mapsto \int \rho_s \log \rho_s dm_s$ and (233). □

The following two lemmas are crucial to conclude that the forward adjoint heat flow defines the EDE-gradient flow for the relative entropy. The first lemma gives an estimate of the squared slope of the entropy in terms of the Fisher information, which is an estimate in the static setting, while the second lemma represents a dynamic version of Kuwada's Lemma, see e.g. [28, Proposition 3.7]. The proof of Proposition 4.44 relies on the dual formula of the dynamic distance $W_{s,t}$ (recall Definition 4.9) in terms of subsolutions to a modified Hamilton-Jacobi equation, cf. Section 2.9.1.

Proposition 4.43. *For $\mu = \rho m_t \in Dom(S)$*

$$|\nabla_t S_t|^2(\mu) \leq \int_{\{\rho>0\}} \frac{|\nabla_t \rho|_*^2}{\rho}\,dm_t.$$

Proof. Since each static space (X, d_t, m_t) satisfies $CD(K,\infty)$, Theorem 9.3 in [5] yields the assertion. □

Proposition 4.44. *Let $(\rho_t)_{0\leq t\leq T}$ be the solution to the forward adjoint heat equation with nonnegative initial datum $h \in \mathcal{H}$ such that $\int h\,dm_0 = 1$. Then the curve $t \mapsto \mu_t := \rho_t m_t$ is locally absolutely continuous and satisfies*

$$|\dot\mu_t|_t^2 \leq \int_{\{\rho_t>0\}} \frac{|\nabla_t \rho_t|_*^2}{\rho_t}\,dm_t \qquad \textit{for a.e. } t \in [0,T].$$

Proof. From (230) we know that $\int \rho_t\,dm_t = 1$ for every $0 \leq t \leq T$. Hence each $\mu_t = \rho_t m_t$ is a probability measure.

Let $s < t$ and set $\delta := t - s$. Then, with $\vartheta(a) = s + a$, we define HLS_ϑ as in Section 2.9.1

$$\mathrm{HLS}_\vartheta := \left\{ \varphi \in \mathrm{Lip}_b([0,\delta] \times X) \,\middle|\, \partial_a \varphi_a \leq -\frac{1}{2} |\nabla_{\vartheta(a)}(\varphi_a)|_*^2 \right.$$
$$\left. L^1 \times m \text{ a.e. in } (0,\delta) \times X \right\},$$

and

$$\tilde{W}_\vartheta^2(\mu_s, \mu_t) := 2 \sup_\varphi \left\{ \int \varphi_\delta d\mu_t - \int \varphi_0 d\mu_s \right\},$$

where the supremum runs over all maps $\varphi(a,x) = \varphi_a(x) \in \mathrm{HLS}_\vartheta$. Then we have by Lemma 2.59

$$W_s^2(\mu_s, \mu_t) \leq e^{2L\delta} \delta \tilde{W}_\vartheta^2(\mu_s, \mu_t).$$

By applying [4, Lemma 4.3.4] to the function $(a,b) \mapsto \int \rho_a \varphi_b dm_a$, where $\varphi \in \mathrm{HLS}_\vartheta$, we obtain

$$\begin{aligned}
&\int \varphi_\delta \, d\mu_t - \int \varphi_0 \, d\mu_s = \int_0^\delta \partial_a \int \varphi_a \, d\mu_{s+a} \, da \\
&\leq \int_0^\delta \int -\frac{1}{2} |\nabla_{s+a}(\varphi_a)|_*^2 \, d\mu_{s+a} - \mathcal{E}_{s+a}(\rho_{s+a}, \varphi_a) \, da \\
&\leq \int_0^\delta \int -\frac{1}{2} |\nabla_{s+a}(\varphi_a)|_*^2 \, d\mu_{s+a} \\
&+ \int \frac{1}{2} |\nabla_{s+a}(\varphi_a)|_*^2 \, d\mu_{s+a} + \frac{1}{2} \int_{\{\rho_{s+a} > 0\}} \frac{|\nabla_{s+a}(\rho_{s+a})|_*^2}{\rho_{s+a}} \, dm_{s+a} \, da \\
&= \int_0^\delta \frac{1}{2} \int_{\{\rho_{s+a} > 0\}} \frac{|\nabla_{s+a}(\rho_{s+a})|_*^2}{\rho_{s+a}} \, dm_{s+a} \, da.
\end{aligned}$$

Taking the supremum over all φ

$$W_s^2(\mu_s, \mu_t) \leq e^{2L\delta} \delta \int_0^\delta \int_{\{\rho_{s+a} > 0\}} \frac{|\nabla_{s+a}(\rho_{s+a})|_*^2}{\rho_{s+a}} \, dm_{s+a} \, da.$$

Dividing by δ^2 and letting $\delta \to 0$ we conclude

$$|\dot{\mu}_s|_s^2 \leq \int_{\{\rho_s > 0\}} \frac{|\nabla_s \rho_s|_*^2}{\rho_s} \, dm_s.$$

□

Now we are ready to prove our main result of this section.

Theorem 4.45. *Let $(X, d_t, m_t)_{t \in [0,T]}$ be a family of Polish metric measure spaces with geodesic distances (d_t) satisfying* (161) *such that $m_t = e^{-f_t} m$, where $m \in \mathcal{P}(X)$ and (f_t) are Borel functions satisfying $|f_t(x)| \leq C$ and* (226). *Assume that each static space satisfies $RCD(K, N')$ for finite numbers $K, N' \in \mathbb{R}$. Let $h \in \mathcal{H}$ nonnegative with $\bar{\mu} = h m_0 \in Dom(S)$.*

1. *Let (ρ_t) solve the forward adjoint heat equation starting from h, then $\mu_t = \rho_t m_t$ is the dynamic EDE-gradient flow for the relative entropy S starting in $\bar{\mu}$.*

2. *Conversely, let (μ_t) be the dynamic EDE-gradient flow for S, then $\mu_t = \rho_t m_t$ and (ρ_t) is the solution to the forward adjoint heat equation.*

Proof. Proposition 4.42 applied the forward flow (ρ_t) yields

$$\frac{d}{dt}\int \rho_t \log \rho_t \, dm_t = -\int_{\{\rho_t>0\}} \frac{|\nabla_t \rho_t|_*^2}{\rho_t}\, dm_t + \int (\partial_t f_t)\rho_t \, dm_t.$$

Integrating from 0 to t and using Proposition 4.44 and Proposition 4.43 we obtain

$$S_t(\mu_t) + \frac{1}{2}\int_0^t |\dot{\mu}_r|_r^2 dr + \frac{1}{2}\int_0^t |\nabla_r S_r|^2(\mu_r) dr \leq S_0(\bar{\mu}) + \int_0^t (\partial_r S_r)(\mu_r) dr.$$

Moreover, by virtue of Proposition 2.16, (μ_t) is contained in the sublevel set of the entropy and hence, similarly as in the proof of Theorem 4.33, we get for all t

$$S_t(\mu_t) - S_0(\bar{\mu}) \geq \int_0^t (\partial_r S_r)(\mu_r) dr - \int_0^t |\dot{\mu}|_r |\nabla_r S_r|(\mu_r) dr.$$

Thus we have

$$S_t(\mu_t) + \frac{1}{2}\int_0^t |\dot{\mu}_r|_r^2 dr + \frac{1}{2}\int_0^t |\nabla_r S_r|^2(\mu_r) dr = S_0(\bar{\mu}) + \int_0^t (\partial_r S_r)(\mu_r) dr.$$

To show the converse implication, let $\tilde{\rho}_t$ be the solution to the adjoint heat equation parametrized forwards in time. From the previous argumentation we know that $\tilde{\mu}_t = \tilde{\rho}_t m_t$ is a dynamic EDE-gradient flow of the entropy. From Theorem 4.33 there is at most one gradient flow starting from $\bar{\mu}$, hence $\tilde{\mu}_t = \mu_t$ for every $t \in [0, T]$. ☐

Remark 4.46. *Let us recall the complete picture of forward and backward equation described in Section 4.6. The heat equation (forward in time) induces the adjoint heat equation (backward in time) and vice versa. Then $\mu_s := \rho_s m_s$, where ρ_s denotes the adjoint heat flow (backward in time) is an* upward dynamic EDI gradient flow *in the sense that*

$$S_s(\mu_s) + \frac{1}{2}\int_s^T |\dot{\mu}_r|_r^2 \, dr + \frac{1}{2}\int_s^T |\nabla_r S_r(\mu_r)|^2 \, dr = S_T(\mu_T) + \int_s^T (\partial_r S_r)(\mu_r)\, dr.$$

Equivalently, and this is what we showed, if $\mu_t = \rho_t m_t$, where ρ_t solves the adjoint heat equation forward in time, then μ_t solves

$$S_t(\mu_t) + \frac{1}{2}\int_0^t |\dot{\mu}_r|_r^2 \, dr + \frac{1}{2}\int_0^t |\nabla_r S_r(\mu_r)|^2 \, dr = S_0(\mu_0) + \int_0^t (\partial_r S_r)(\mu_r)\, dr.$$

But then the heat equation is a backward equation.

4.6.2 Identification of the Heat Flow with the Dynamic Gradient Flow for Cheeger's Energy

In the following let $(X, d_t, m_t)_{t\in[0,T]}$ be a family of Polish metric measure spaces. We suppose that (d_t) satisfies (161) and $m_t = e^{-f_t}m$, where m is a σ-finite Borel measure on X and (f_t) are Borel functions satisfying

$$|f_t(x) - f_s(y)| \leq L^*|t-s|. \tag{235}$$

We denote the space of square integrable functions by

$$L^2(X) = \{u \in L^2(X, m_t)| \text{ for some (and hence any) } t \in [0,T]\}$$

and consider $\mathrm{Ch}\colon [0,T] \times L^2(X) \to [0,\infty]$ defined by

$$(t,u) \mapsto \mathrm{Ch}_t(u) = \frac{1}{2}\int_X |\nabla_t u|_*^2 dm_t,$$

where $|\nabla_t u|_*$ denotes the minimal relaxed gradient of u. With $\langle u, v\rangle_t = \int u \cdot v\, dm_t$, $L^2(X, m_t) = (L^2(X), \langle\cdot,\cdot\rangle_t)$ is a separable Hilbert space and since the assumptions on the energy functional from Section 4.5 are satisfied by Ch we directly obtain existence of gradient flows in the sense of Definition 4.34.

Theorem 4.47. *Let $\bar u \in Dom(\mathrm{Ch})$. Then there exists a unique dynamic gradient flow for* Ch *starting in $\bar u$, i.e. an absolutely continuous curve (u_t) solving*

$$\partial_t u_t \in -D_t^- \mathrm{Ch}_t(u_t) \quad \textit{for a.e. } t \in (0,T)$$

and $\lim_{t\to 0} u_t = \bar u$.

Proof. Obviously $\mathrm{Ch}_t \geq 0$ for every $t \in [0,T]$. Moreover $u \mapsto \mathrm{Ch}_t(u)$ is convex and lower semicontinuous by Theorem 4.5 in [5]. From Lemma 4.39 and (235) we obtain

$$\begin{aligned}
&|\mathrm{Ch}_t(u) - \mathrm{Ch}_s(u)| \leq |\int |\nabla_t u|_*^2 - |\nabla_s u|_*^2\, dm_t| + |\int |\nabla_s u|_*^2\, d(m_t - m_s)| \\
&\leq 2L|t-s|\int |\nabla_s u|_*^2\, dm_t + L^* e^{L^*|t-s|}|t-s|\int |\nabla_s u|_*^2\, dm_s \\
&\leq 2L|t-s|e^{C|t-s|}\int |\nabla_s u|_*^2\, dm_s + L^* e^{L^*|t-s|}|t-s|\int |\nabla_s u|_*^2\, dm_s \\
&\leq (2L + L^*)e^{L^*|t-s|}|t-s|\mathrm{Ch}_s(u).
\end{aligned}$$

We get the result as a consequence of Theorem 4.38. □

In the case when the underlying space satisfies $\mathrm{RCD}(K, N')$ we may identify the gradient flow for Cheeger's energy with the heat flow $\partial_t u_t = \Delta_t u_t$.

Theorem 4.48. *Let $(X, d_t, m_t)_{t\in[0,T]}$ be a family of Polish metric measure spaces with geodesic distances (d_t) satisfying* (161) *such that $m_t = e^{-f_t}m$, where $m \in \mathcal{P}(X)$ and (f_t) are Borel functions satisfying $|f_t(x)| \leq C$ and* (226)*. Assume that each static space satisfies $RCD(K, N')$ for finite numbers $K, N' \in \mathbb{R}$. Let $(\tilde u_t)$ be the solution to the heat equation $\partial_t \tilde u_t = \Delta_t \tilde u_t$ on $(0,T) \times X$ starting in some $\bar u \in Dom(\mathrm{Ch})$. Then $(\tilde u_t)$ satisfies*

$$\partial_t \tilde u_t \in -D_t^- \mathrm{Ch}_t(\tilde u_t) \quad \textit{for a.e. } t \in (0,T),$$

and can be constructed as the limit of a minimizing movement scheme. Conversely, let (u_t) *be the dynamic gradient flow of Cheeger's energy* Ch *starting in* $\bar{u} \in \mathit{Dom}(\mathrm{Ch})$*. Then* (u_t) *solves the heat equation*

$$\partial_t u_t = \Delta_t u_t \quad on\ (0,T) \times X.$$

In particular $u_t = \tilde{u}_t$ *in* $L^2(X)$ *for every* $t \geq 0$*.*

Proof. Both flows satisfy the dynamic $\mathrm{EVI}(-L/2, \infty)$ gradient flow inequality almost everywhere by virtue of Proposition 4.35 and Theorem 2.24. Hence from the contraction estimate (174)

$$||u_t - \tilde{u}_t||_t^2 \leq e^{7L(t-s)} ||u_s - \tilde{u}_s||_s^2 \quad \text{for a.e. } t \geq s,$$

we obtain

$$||u_t - \tilde{u}_t||_t^2 \leq \lim_{s \to 0} e^{7L(t-s)} ||u_s - \tilde{u}_s||_s = 0 \quad \text{for a.e. } t,$$

and hence by continuity $||u_t - \tilde{u}_t||_t = 0$ for every t. □

References

[1] Luigi Ambrosio, Matthias Erbar, and Giuseppe Savaré. Optimal transport, Cheeger energies and contractivity of dynamic transport distances in extended spaces. *Nonlinear Analysis: Theory, Methods & Applications*, 137:77–134, 2016.

[2] Luigi Ambrosio and Nicola Gigli. A user's guide to optimal transport, 2011.

[3] Luigi Ambrosio, Nicola Gigli, Andrea Mondino, and Tapio Rajala. Riemannian Ricci curvature lower bounds in metric measure spaces with σ-finite measure. *Trans. Amer. Math. Soc.*, 367(7):4661–4701, 2015.

[4] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the Space of Probabiliy Measures*. Springer Science & Business Media, 2008.

[5] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below. *Invent. Math.*, 195(2):289–391, 2013.

[6] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. Metric measure spaces with Riemannian Ricci curvature bounded from below. *Duke Math. J.*, 163(7):1405–1490, 2014.

[7] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. Bakry-Émery curvature-dimension condition and Riemannian Ricci curvature bounds. *Ann. Probab.*, 43(1):339–404, 2015.

[8] Luigi Ambrosio, Andrea Mondino, and Giuseppe Savaré. Nonlinear diffusion equations and curvature conditions in metric spaces. *arXiv:1509.07273*, 2015.

[9] Luigi Ambrosio, Giuseppe Savaré, and Lorenzo Zambotti. Existence and stability for Fokker-Planck equations with log-concave reference measure. *Probab. Theory Related Fields*, 145(3-4):517–564, 2009.

[10] Marc Arnaudon, Kolehe Abdoulaye Coulibaly, and Anton Thalmaier. Brownian motion with respect to a metric depending on time; definition, existence and applications to Ricci flow. *Comptes Rendus Mathematique*, 346(13):773–778, 2008.

[11] Kathrin Bacher and Karl-Theodor Sturm. Ricci bounds for Euclidean and spherical cones. In *Singular phenomena and scaling in mathematical models*, pages 3–23. Springer, Cham, 2014.

[12] 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*, volume 1123 of *Lecture Notes in Math.*, pages 145–174. Springer, Berlin, 1985.

[13] D. Bakry and Michel Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

[14] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.

[15] Heinz Bauer. *Probability theory and elements of measure theory*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London-New York, 1981. Second edition of the translation by R. B. Burckel from the third German edition, Probability and Mathematical Statistics.

[16] Heinz Bauer. *Wahrscheinlichkeitstheorie*. de Gruyter Lehrbuch. [de Gruyter Textbook]. Walter de Gruyter & Co., Berlin, fifth edition, 2002.

[17] Vladimir I. Bogachev. *Measure Theory*, volume 1. Springer-Verlag, 2007.

[18] Francois Bolley, Ivan Gentil, Arnaud Guillin, and Kazumasa Kuwada. Equivalence between dimensional contractions in Wasserstein distance and the curvature-dimension condition. *arXiv:1510.07793*, 2015.

[19] Haïm Brezis. *Operateurs Maximaux Monotones et semi-groupes de contractions dans les espaces de Hilbert*, volume 5. Elsevier, 1973.

[20] Jeff Cheeger. Differentiability of Lipschitz functions on metric measure spaces. *Geom. Funct. Anal.*, 9(3):428–517, 1999.

[21] Zhen-Qing Chen and Masatoshi Fukushima. *Symmetric Markov processes, time change, and boundary theory*, volume 35 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2012.

[22] Dario Cordero-Erausquin, Robert J. McCann, and Michael Schmuckenschläger. A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. *Invent. Math.*, 146(2):219–257, 2001.

[23] Matthias Erbar. The heat equation on manifolds as a gradient flow in the Wasserstein space. *Annales de l'I. H. P. Probabilités et Statistiques*, 46(1):1–23, 2010.

[24] Matthias Erbar, Kazumasa Kuwada, and Karl-Theodor Sturm. On the equivalence of the entropic curvature-dimension condition and Bochner's inequality on metric measure spaces. *Invent. Math.*, 201:993–1071, 2015.

[25] Lucas Ferreira and Julio Valencia-Guevara. Gradient flows of time-dependent functionals in metric spaces and applications for PDE. *arXiv:1509.0416161v1*, 2015.

[26] Nicola Gigli. On the heat flow on metric measure spaces: existence, uniqueness and stability. *Calc. Var.*, 39(1-2):101–120, 2010.

[27] Nicola Gigli. Nonsmooth differential geometry-An approach tailored for spaces with Ricci curvature bounded from below. *arXiv:1407.0809*, 2014.

[28] Nicola Gigli, Kazumasa Kuwada, and Shin-ichi Ohta. Heat flow on Alexandrov spaces. *Communications on Pure and Applied Mathematics*, 66(3):307–331, 2013.

[29] Robert Haslhofer and Aaron Naber. Weak solutions for the Ricci flow I. *arXiv:1504.00911*, 2015.

[30] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29:1–17, 1998.

[31] Leonid V. Kantorovich. On mass transportation. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 312(Teor. Predst. Din. Sist. Komb. i Algoritm. Metody. 11):11–14, 2004.

[32] Bruce Kleiner and John Lott. Singular Ricci flows I. *arXiv:1408.2271*, 2014.

[33] Martin Knott and Cyril S. Smith. On the optimal mapping of distributions. *J. Optim. Theory Appl.*, 43(1):39–49, 1984.

[34] Eva Kopfer. Gradient flow for the Boltzmann entropy and Cheeger's energy on time-dependent metric measure spaces. *arXiv:1611.09522*, 2017.

[35] Eva Kopfer and Karl-Theodor Sturm. Heat flows on Time-dependent Metric Measure Spaces and Super-Ricci Flows. *arXiv:1611.02570*, 2017.

[36] Kazumasa Kuwada. Duality on gradient estimates and Wasserstein controls. *J. Funct. Anal.*, 258(11):3758–3774, 2010.

[37] Kazumasa Kuwada. Space-time Wasserstein controls and Bakry–Ledoux type gradient estimates. *Calculus of Variations and Partial Differential Equations*, 54(1):127–161, 2015.

[38] Kazumasa Kuwada and Robert Philipowski. Coupling of Brownian motions and Perelman's L-functional. *Journal of Functional Analysis*, 260(9):2742–2766, 2011.

[39] Kazumasa Kuwada and Robert Philipowski. Non-explosion of diffusion processes on manifolds with time-dependent metric. *Mathematische Zeitschrift*, 268(3-4):979–991, 2011.

[40] Sajjad Lakzian and Michael Munn. Super Ricci flow for disjoint unions. *arXiv preprint arXiv:1211.2792*, 2012.

[41] Janna Lierl and Laurent Saloff-Coste. Parabolic Harnack inequality for time-dependent non-symmetric Dirichlet forms. *arXiv:1205.6493*, 2012.

[42] Jacques Louis Lions and Enrico Magenes. *Non-homogeneous boundary value problems and applications*, volume 1. Springer Science & Business Media, 2012.

[43] John Lott and Cédric Villani. Ricci curvature for metric-measure spaces via optimal transport. *Annals of Mathematics*, 169(2):903–991, 2009.

[44] Zhi Ming Ma and Michael Röckner. *Introduction to the theory of (nonsymmetric) Dirichlet forms*. Springer-Verlag, Berlin, 1992.

[45] Robert J. McCann. *A convexity theory for interacting gases and equilibrium crystals*. ProQuest LLC, Ann Arbor, MI, 1994. Thesis (Ph.D.)–Princeton University.

[46] Robert J. McCann and Peter M. Topping. Ricci flow, entropy and optimal transportation. *American Journal of Mathematics*, 132(3):711–730, 2010.

[47] Gaspard Monge. *Mémoire sur la théorie des déblais et des remblais*. De l'Imprimerie Royale, 1781.

[48] Ricardo H. Nochetto, Giuseppe Savaré, and Claudio Verdi. A posteriori error estimates for variable time-step discretizations of nonlinear evolution equations. *Communications on Pure and Applied Mathematics*, 53(5):525–589, 2000.

[49] Shin-Ichi Ohta and Karl-Theodor Sturm. Heat flow on Finsler manifolds. *Comm. Pure Appl. Math.*, 62(10):1386–1433, 2009.

[50] Felix Otto and Cédric Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2):361–400, 2000.

[51] Svetlozar T. Rachev and Ludger Rüschendorf. *Mass transportation problems. Vol. I*. Probability and its Applications (New York). Springer-Verlag, New York, 1998.

[52] Michael Renardy and Robert C. Rogers. *An introduction to partial differential equations*, volume 13. Springer Science & Business Media, 2006.

[53] Riccarda Rossi, Alexander Mielke, and Giuseppe Savaré. A metric approach to a class of doubly nonlinear evolution equations and applications. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, 7:97–169, 2008.

[54] Riccarda Rossi and Giuseppe Savaré. Gradient flows of non convex functionals in Hilbert spaces. *ESAIM Control Optim. Calc. Var.*, 12:564–614, 2006.

[55] Giuseppe Savaré. Self-improvement of the Bakry-Émery condition and Wasserstein contraction of the heat flow in RCD(K,∞) metric measure spaces. *Discr. Cont. Dyn. Syst. A*, 34(4):1641–1661, 2014.

[56] Karl-Theodor Sturm. Analysis on local Dirichlet spaces. II. Upper Gaussian estimates for the fundamental solutions of parabolic equations. *Osaka J. Math.*, 32(2):275–312, 1995.

[57] Karl-Theodor Sturm. On the geometry of metric measure spaces. I and II. *Acta Math.*, 169(1):65–131, 2006.

[58] Karl-Theodor Sturm. Metric measure spaces with variable Ricci bounds and couplings of Brownian motions. In *Festschrift Masatoshi Fukushima*, volume 17 of *Interdiscip. Math. Sci.*, pages 553–575. World Sci. Publ., Hackensack, NJ, 2015.

[59] Karl-Theodor Sturm. Super Ricci flows for metric measure spaces. I. *arXiv:1603.02193*, 2016.

[60] Peter Topping. $\mathcal{L}$-optimal transportation for Ricci flow. *J. Reine Angew. Math.*, 636:93–122, 2009.

[61] Peter M. Topping. *Lectures on the Ricci flow*, volume 325. Cambridge University Press, 2006.

[62] Cédric Villani. *Optimal transport, old and new.* Springer-Verlag, Berlin, Heidelberg, 2009.

[63] Max-K von Renesse and Karl-Theodor Sturm. Transport inequalities, gradient estimates, entropy and Ricci curvature. *Communications on pure and applied mathematics*, 58(7):923–940, 2005.

[64] Feng-Yu Wang. Equivalent semigroup properties for the curvature-dimension condition. *Bull. Sci. Math.*, 135(6-7):803–815, 2011.