Heat Flow on Time-Dependent Metric Measure Spaces and Super-Ricci Flows

EVA KOPFER
Universität Bonn

KARL-THEODOR STURM
Universität Bonn

Abstract

We study the heat equation on time-dependent metric measure spaces (as well as the dual and the adjoint heat equation) and prove existence, uniqueness, and regularity. Of particular interest are properties that characterize the underlying space as a super-Ricci flow as previously introduced by the second author [51]. Our main result yields the equivalence of

- dynamic convexity of the Boltzmann entropy on the (time-dependent) $L^2$-Wasserstein space,
- monotonicity of $L^2$-Kantorovich-Wasserstein distances under the dual heat flow acting on probability measures (backward in time),
- gradient estimates for the heat flow acting on functions (forward in time), and
- a Bochner inequality involving the time-derivative of the metric.

Moreover, we characterize the heat flow on functions as the unique forward EVI flow for the (time-dependent) energy in $L^2$-Hilbert space and the dual heat flow on probability measures as the unique backward EVI flow for the (time-dependent) Boltzmann entropy in $L^2$-Wasserstein space. © 2018 Wiley Periodicals, Inc.

Contents

1. Introduction and Statement of Main Results 2501
2. The Heat Equation for Time-Dependent Dirichlet Forms 2522
3. Heat Flow and Optimal Transport on Time-Dependent Metric Measure Spaces 2538
4. Towards Transport Estimates 2549
5. From Transport Estimates to Gradient Estimates and the Bochner Inequality 2559
6. From Gradient Estimates to Dynamic EVI 2576
Appendix: EVI and Dynamic Convexity 2597
Bibliography 2605

Communications on Pure and Applied Mathematics, Vol. LXXI, 2500–2608 (2018)
© 2018 Wiley Periodicals, Inc.
1 Introduction and Statement of Main Results

1.1 Introduction

The present paper has two main objectives:

(i) to define and study the heat flow on time-dependent metric measure spaces
and

(ii) to characterize super-Ricci flows of metric measure spaces by properties
of optimal transports and heat flows.

The former is regarded as the “parabolic” analogue to the analysis of heat flow,
optimal transport, and functional inequalities on “static” metric measure spaces.
The latter should be considered as a first contribution to a theory of Ricci flows
of metric measure spaces. Our approach will combine and extend two previous—
hitherto unrelated—lines of developments: the analysis on (static) metric measure
spaces and the analysis on (smooth) time-dependent Riemannian manifolds.

Heat Flow on (Static) Metric Measure Spaces

The heat equation is one of the most fundamental and well-studied PDEs on Rie-
mannian manifolds. It is intimately linked to other important objects like Dirich-
let energy, Boltzmann entropy, optimal transport, and Brownian motion. On one
hand, it is a very robust object and admits an integral representation in terms
of the heat kernel. Without any extra assumptions, its existence and basic properties
are always guaranteed. On the other hand, its more subtle properties reveal deep
information on the underlying space, like curvature, genus, index, etc.

In the last few decades, the heat flow was also successfully studied on more
general spaces, in particular, on metric measure spaces (mm-spaces for short) [14,
21, 47, 49]. The foundational work of Ambrosio, Gigli, and Savaré [4–6] clarified
the picture, allowing the unification of various of the previous approaches and made
clear that for each metric measure space \( (X, d, m) \)

\[
\int \exp(-C d^2(x, z)) d m(x) < \infty
\]

(for some \( C, z \)) there exists a unique solution to the heat equation, most conve-
niently defined as gradient flow in \( L^2(X, m) \) for the Dirichlet energy (“Cheeger
energy”) \( \mathcal{E}(u) = \int_X |\nabla u|^2 \, d m \).

Synthetic Lower Ricci Bounds

The heat flow on Riemannian manifolds—and more generally on metric mea-
ure spaces—turned out to be a powerful tool for characterizing (synthetic) lower
bounds on the Ricci curvature. Such curvature bounds are indeed necessary and
sufficient for various important properties of the heat flow \( t \mapsto P_t u \). Moreover,
they imply that \( t \mapsto (P_t u)m \) is the gradient flow for the Boltzmann entropy
\( S(u m) = \int u \log u \, d m \) in the space \( \mathcal{P}(X) \) of probability measures equipped with
the \( L^2 \)-Kantorovich-Wasserstein distance \( W \).
For instance, nonnegative Ricci curvature is equivalent to:

▷ the gradient estimate $|\nabla P_t u|^2 \leq P_t |\nabla u|^2$,
▷ the existence of coupled pairs of Brownian motions with $d(X_t, Y_t) \leq d(X_0, Y_0)$,
▷ the transport estimate $W((P_t u)m, (P_t v)m) \leq W(um, vm)$, and
▷ convexity of the Boltzmann entropy $S$ on the geodesic space $(P(X), W)$.

Indeed, in the Lott-Sturm-Villani approach to synthetic lower Ricci bounds [38, 50], the latter property was used to define nonnegative Ricci curvature for metric measure spaces. Furthermore, the previous properties—gradient estimate, coupling property of Brownian motions, and transport estimate—illustrate the effect of nonnegative Ricci curvature in a very graphical way, well suited for applications and modeling, and also make perfect sense in discrete settings; cf. Ollivier [41], Tan-nnenbaum et al. [19], and Sandhu, Georgiou, and Tannenbaum [46].

Heat Flow on Time-Dependent Metric Measure Spaces

New phenomena emerge and novel challenges arise for the heat flow if the underlying geometric objects (Riemannian manifolds, metric measure spaces) will vary in time, e.g., if they will change their “shape” or “material properties.” This might result from exterior forces or from an interior dynamic, like mean curvature flow or Ricci flow. To model such time-dependent geometric objects, one typically considers families $(M, g_t), t \in I$, consisting of a manifold $M$ and a one-parameter family of metric tensors $g_t, t \in I \subset \mathbb{R}$. We will consider more generally time-dependent metric measure spaces $(X, d_t, m_t), t \in I$, consisting of a Polish space $X$ equipped with one-parameter families of metrics (= distance functions) $d_t$ and measures $m_t, t \in I$. The main questions to be addressed are:

(a) In which generality does existence and uniqueness hold for solutions to the heat equation on time-dependent metric measure spaces?

(b) Is the heat flow the gradient flow for the energy? Does it coincide with the gradient flow for the entropy? More generally, is there a meaningful concept of gradient flows for time-dependent functionals on time-dependent geodesic spaces?

(c) What is the time-dependent counterpart to nonnegative Ricci curvature or, more generally, to the CD(0, $\infty$)-condition? More precisely, which kind of curvature bound is necessary and/or sufficient for (the time-dependent counterpart to) the gradient estimate? Which for the corresponding transport estimate? Is there a synthetic version of such a curvature bound?

In contrast to the static case, until now nothing seemed to be known for the heat flow on general time-dependent metric measure spaces.

For time-dependent Riemannian manifolds $(M, g_t), t \in I$ (with smoothly varying, nondegenerate $g_t$) question (a) allows for an easy, affirmative answer. Surprisingly enough, Brownian motion was constructed only recently [9, 16]. Question (b) had been unsolved so far. McCann-Topping 2010 [39], Arnaudon-Coulibaly-Thal-
maier [8], and Haslhofer-Naber [24] proved that the first three questions in (c) have one common answer:

$$\text{Ric}_{g_t} + \frac{1}{2} \partial_t g_t \geq 0.$$  

Finally, in [51], the second author presented a synthetic definition for the latter, formulated as “dynamic convexity” of the Boltzmann entropy $S_t$ in the Wasserstein space $(P(X), W_t)$.

The current paper, regarded as an accompanying paper to [51], will provide complete answers to the previous questions in the setting of time-dependent metric measure spaces. We will prove existence, uniqueness, and regularity results for the heat equation and its dual. The former will be identified as the forward gradient flow for the Dirichlet energy $E_t$ in $L^2(X, m_t)$, and the latter as the backward gradient flow for the Boltzmann entropy $S_t$ in $(P(X), W_t)$. A general discussion on gradient flows for time-dependent functionals on time-dependent geodesic spaces will be included. Our main result provides a comprehensive characterization of super-Ricci flows $(X, d_t, m_t)$, $t \in I$, by the equivalence of dynamic convexity of the Boltzmann entropy, monotonicity of transport estimates under the dual heat flow, monotonicity of gradient estimates under the primal heat flow, and the time-dependent Bochner inequality.

In the static case, synthetic lower Ricci bounds will play its role to the full only in combination with an upper bound on the dimension that led to the formulation of the so-called curvature-dimension condition $\text{CD}(K, N)$. The time-dependent counterpart to the $\text{CD}(K, N)$-condition will be the so-called $\text{supcr}(K, N)$-Ricci flows. Taking into account the role of the parameter $N \in \mathbb{R}_+$ requires quite some effort. However, we expect this to be worthwhile for future applications. The case $K \neq 0$, however, can be reduced to the case $K = 0$ by means of a simple scaling of space and time; see Theorem 1.11. To simplify the presentation, throughout this paper we will thus restrict ourselves to the curvature bound $K = 0$.

### Ricci Flows, Super-Ricci Flows, and Super-$N$-Ricci Flows

Given a manifold $M$ and a smooth 1-parameter family $(g_t)$, $t \in I$, of Riemannian tensors on $M$, we say that the “time-dependent Riemannian manifold” $(M, g_t)$, $t \in I$, evolves as a Ricci flow if $\text{Ric}_{g_t} = -\frac{1}{2} \partial_t g_t$ for all $t \in I$. It is called a super-Ricci flow if instead only $\text{Ric}_{g_t} \geq -\frac{1}{2} \partial_t g_t$ holds true on $M \times I$ (regarded as inequalities between quadratic forms on the tangent bundle of $(M, g_t^x)$ for each $(x, t) \in M \times I$). In other words, super-Ricci flows are supersolutions to the Ricci flow equation and Ricci flows are minimal super-Ricci flows.

Thanks to the groundbreaking work of Hamilton [22][23] and Perelman [42][44] (see also [13][26][40]), Ricci flow has attracted a lot of attention and has proved itself as a powerful tool and inspiring source for many new developments. Currently, one of the major challenges is to extend the theory of Ricci flows and the scope of its applications beyond the setting of smooth Riemannian manifolds. In particular, one
Aims to define and analyze (Ricci) flows through singularities and to study evolutions of spaces with changing dimension and/or topological type. Kleiner-Lott [27] and Haslhofer-Naber [24] presented notions of singular and weak solutions for Ricci flows. In [24], Ricci flows of “regular” (i.e., smooth with uniform bounds on curvature and derivatives of it) time-depending Riemannian manifolds \( (M, g_t), t \in I \), of arbitrary dimension are characterized by means of functional inequalities on the path space (spectral gap or logarithmic Sobolev inequalities for the Ornstein-Uhlenbeck operator). In [27], Ricci flow of “singular” 3-dimensional Riemannian manifolds \( (M, g_t), t \in I \) (regarded as 4-dimensional Ricci flow space-times) is defined and analyzed in detail, allowing also for Ricci flows through singularities.

Compared to Ricci flows, super-Ricci flows allow for much larger classes of examples. This is an advantage if one is interested in analysis (e.g., functional inequalities, heat kernel estimates, etc.) on huge classes of singular spaces or if one tries to extend tools and insights from the study of classical Ricci flows to more general time evolutions of geometric objects. It is a disadvantage if one aims for uniqueness results or for properties close to those of Ricci flows.

The defining property of super-Ricci flows for mm-spaces \((X, d_t, m_t)_t\) contains no constraint on the evolution of the measures \(m_t\) but only a lower bound on the evolution of the distances \(d_t\). Moreover, super-Ricci flows can increase the dimension in order to match the constraint imposed by the lower bound on the Ricci curvature.

These distracting effects can be ruled out by considering the more restrictive class of “super-N-Ricci flows.” A time-dependent weighted \(n\)-dimensional Riemannian manifold \((M, g_t, e^{-f_t} \text{d} \text{vol}_{g_t})_t\), for instance, is a super-\(n\)-Ricci flow if and only if \(g_t\) satisfies (1.1) and if \(f_t\) is constant for each \(t\); see [51, theorem 2.9].

In [51], the second author of this paper presented a synthetic definition for super-\(N\)-Ricci flows in the general setting of time-dependent metric measure spaces. A work in progress deals with synthetic upper Ricci bounds [52], which—in combination with the former—will then also allow for characterizations of Ricci flows of mm-spaces. For most of our results, we request a controlled \(t\)-dependence for \(d_t\) and \(m_t\). Of course, this is a severe limitation and rules out various challenging applications. Even more, one might wish to replace \(X\) by varying \(X_t\), e.g., allowing for changing topological type. However, in contrast to the static case, so far there are no existence and uniqueness results for the heat flow on time-dependent mm-spaces that hold in full generality. The current paper will lay the foundations for further work devoted to enlarge the scope and to include singularities and degenerations.

### 1.2 Some Examples

Let us give some motivating examples of super-Ricci flows as defined in [51, def. 2.4]. We also discuss whether they are super-\(N\)-Ricci flows or Ricci flows.

**Example 1.1** (Vertebral column). Consider a surface of revolution with piecewise constant negative curvature \(\text{Ric} = -Kg\) for some \(K > 0\) depicted in Figure 1.1.
Under the evolution of a Ricci flow, the curvature of the surface where $\text{Ric} = -Kg$ will increase, while the curvature of the “rims” ($\text{Ric} = +\infty$) 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—or it may start to smoothen them at any later time or with smaller speed.

**Example 1.2 (Wandering Gaussian).** Let $X = \mathbb{R}^n$, $d_t(x, y) = \|x - y\|$, and $m_t = e^{-V_t} \text{Leb}^n$ with

$$V_t(x) = (x, \alpha_t)^2 + (x, \beta_t) + \gamma_t$$

where $\alpha, \beta : I \to \mathbb{R}^n$ and $\gamma : I \to \mathbb{R}$ are arbitrary functions. Then $(X, d_t, m_t)$, $t \in I$, is a super-Ricci flow. For each $N \in [n, \infty)$ it will be a super-$N$-Ricci flow if and only if $\alpha \equiv \beta \equiv 0$.

**Example 1.3 (Exploding point).** Let $(M, g_0)$ be a compact, $n$-dimensional Riemannian manifold of constant Ricci curvature $-Kg_0 < 0$ (e.g., a compact quotient of a hyperbolic space) and let

$$g_t = \begin{cases} (1 + 2Kt)g_0, & t > t_*, \\ 0, & t \leq t_* \end{cases}$$

for $t_* = -\frac{1}{2K}$. Let $(X, d_t, m_t), t \in \mathbb{R}$, be the induced time-dependent mm-space with normalized volume $m_t$ where $(X, d_t)$ for $t \leq t_*$ will be identified with a 1-point space (and $m_t$ with the Dirac mass in this point); see also Figure 1.2. Then this is a super-Ricci flow—provided we slightly enlarge the scope of [51] to also admit degenerate distances $d_t$ (or varying spaces $X_t$). It will not be a super-$N$-Ricci flow for $N < n$.

More generally, consider $(\tilde{M}, \tilde{g}_t) = (M' \times M, g' \otimes g_t)$ with $(M', g')$ being a compact $n'$-dimensional Ricci-flat Riemannian manifold. Then the induced time-dependent mm-space is a super-Ricci flow but no super-$N$-Ricci flow for $N < n' + n$. For any $N \in [n', n' + n)$, up to isometry the only super-$N$-Ricci flow that
Figure 1.2. Point exploding to a hyperbolic quotient

coincides with the given mm-space for $t \leq t_*$ is the static mm-space induced by $(M', g')$.

Example 1.4 (Singular suspension). 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 $S := M \times \{0\}$ and $N := M \times \{\pi\}$ to a point, the south and the north pole, respectively. The resulting space is called spherical suspension and is denoted by $\Sigma(M)$. We endow $\Sigma(M)$ with the measure $d\hat{m}(x, s) := dm(x) \otimes (\sin^4 s \, ds)$ and the 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 $m$ and $d$ are the volume and metric of $M$ and $(x, s), (x', s') \in M \times [0, \pi]$. Since $M$ is an RCD$^*$(3, 4) space, the cone of it is an RCD$^*$ (4, 5) space [25].

The punctured cone $\Sigma_0 := \Sigma(M) \setminus \{S, N\}$ is an incomplete 5-dimensional Riemannian manifold. Let $g_0$ denote the metric tensor of $\Sigma_0$. The curvature of the punctured cone can be calculated explicitly and is given by $\text{Ric}(g_0) = 4g_0$. Then $g(t) := (1 - 8t)g_0$ defines a solution to the Ricci flow $\text{Ric}(g_t) = -\frac{1}{2} \frac{\partial}{\partial t} g_t$ with $g(0) = g_0$, which collapses to a point at time $T = \frac{1}{4}$.

We claim that the associated metric measure space $(\Sigma(M), d_{\Sigma(M)}(t), \hat{m}_t)$, $t \in I$, for $I = (0, T)$ is a super-Ricci flow. Fix $t \in I$ and let $\mu_0, \mu_1 \in \text{Dom}(S_t)$ on $\Sigma(M)$ be given. Let $(\mu_a)_{a \in [0, 1]}$ be a $W_t$-geodesic connecting $\mu_0, \mu_1$. Then, $\mu_a = (e_a)_* \nu$, where $\nu$ 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 $e_a : \Gamma(\Sigma(M)) \to \Sigma(M)$ denotes the
evaluation map. According to [10, theorem 3.3], every optimal path measure \( \nu \) 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 [51, def. 2.4], 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 \( \Sigma_0 \) is bounded neither 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

\[
\begin{align*}
\hat{u}_i &= \frac{1}{\sin r}(u_i, 0, 0), & \hat{v}_i &= \frac{1}{\sin r}(0, v_i, 0), & \hat{w} &= (0, 0, 1),
\end{align*}
\]

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

\[
\begin{align*}
\text{Sec}_{(x,y,r)}(\hat{u}_1, \hat{u}_2) &= \frac{3 - \cos^2 r}{\sin^2 r}, & \text{Sec}_{(x,y,r)}(\hat{u}_1, \hat{v}_1) &= \frac{\cos^2 r}{\sin^2 r}, \\
\text{Sec}_{(x,y,r)}(\hat{u}_1, \hat{v}_1) &= -\frac{\cos^2 r}{\sin^2 r}, & \text{Sec}_{(x,y,r)}(\hat{u}_1, \hat{w}) &= 1,
\end{align*}
\]

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 \( \text{Ric}_{(x,y,r)}(\xi, \xi) = 4 \), but for \( r \to 0 \) and \( r \to \pi \), \( \text{Sec}_{(x,y,r)}(\hat{u}_1, \hat{u}_2) \to +\infty \) and \( \text{Sec}_{(x,y,r)}(\hat{u}_1, \hat{v}_1) \to -\infty \).

We also point out ongoing work [18] indicating that \((\Sigma(M), d_{\Sigma(M)}(t), \hat{m}_t), t \in I\), will not be a Ricci flow in the sense of [52].

1.3 Main Results

Setting

Throughout this introductory section, we fix a time-dependent metric measure space \((X, d_t, m_t), t \in I\), where \( I = (0, T) \) and \( X \) is a compact space equipped with 1-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

\[
(1.2) \quad \left| \log \frac{d_t(x, y)}{d_s(x, y)} \right| \leq L \cdot |t - s|.
\]

(“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 [2][17]. (In various respects, the latter is not really a restriction; see Remark [1.13])
Thus for each $t$ under consideration, there is a well-defined Laplacian $\Delta_t$ 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

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

is defined 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 $\text{lip}_t u$. Here $\text{lip}_t u(x) = \limsup_{y \to x} \frac{|u(x) - u(y)|}{\delta_t(x, y)}$ and $\text{Lip}(X, d_t)$ denotes the class of Lipschitz functions $v : X \to \mathbb{R}$ (which is independent of $t$).

**Heat Equation**

Our first important result concerns existence and uniqueness for solutions to the heat equation—as well as for the adjoint heat equation—on the time-dependent metric measure space $(X, d_t, m_t)_{t \in I}$. Moreover, it yields regularity of solutions and representation as integrals w.r.t. a heat kernel. See Theorems 3.3 and 3.5 for the precise formulations in a slightly more general context.

**Theorem 1.5.** 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

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

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

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

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

Many properties that are self-evident for the heat semigroup on static mm-spaces (e.g., operator and semigroup commute or the semigroup maps $L^2$ into the domain of the operator) no longer hold true for the heat propagator on time-dependent mm-spaces—or require detailed, sophisticated proofs. Let us emphasize here that in general $\text{Dom}(\Delta_t)$ will depend on $t$. 
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 1.6.**

(i) The heat flow is uniquely characterized as the dynamic forward EVI($-L/2$, $\infty$)-flow for $\frac{1}{2}$ $\times$ the Dirichlet energy 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 $\tau \leq T$ and all $w \in \text{Dom}(\mathcal{E})$,

$$-\frac{1}{2} \partial_t^+ \|u_s - w\|_{s,t}^2 \bigg|_{s=t} + \frac{L}{4} \cdot \|u_s - w\|_{s,t}^2 \geq \frac{1}{2} \mathcal{E}_t(u_t) - \frac{1}{2} \mathcal{E}_t(w).$$

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

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

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

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

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

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

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

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

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

**Characterization of Super-Ricci flows**

In [51], the second author 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

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

and $S_t(\mu) = \infty$ if $\mu \ll m_t$. 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 $d_t, t \in I$. This property was proven to be stable under an appropriate space-time version of measured Gromov-Hausdorff convergence and suitably bounded families of super-Ricci flows were shown to be compact—a
far-reaching analogue to the stability and compactness results in the Lott-Sturm-Villani theory of metric measure spaces with synthetic lower Ricci bounds. Furthermore, in the case of time-dependent Riemannian manifolds this novel, synthetic definition of super-Ricci flows was proven to be equivalent to the classical one: \( \text{Ric}_g + \frac{1}{2} \partial_t g_t \geq 0 \).

The main goal of the current paper is to characterize super-Ricci flows in terms of the heat flow (acting on functions, forward in time) and of the dual heat flow (acting on probability measures, backward 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, and Savaré [6].

**Theorem 1.7.** The following assertions are equivalent:

1. For a.e. \( t \in (0, T) \) and every \( W_t \)-geodesic \( (\mu^a)_{a \in [0,1]} \in P \) with \( \mu^0, \mu^1 \in \text{Dom}(S) \),
   \[
   \partial_a^+ S_t(\mu^a) \big|_{a=1} - \partial_a^- S_t(\mu^a) \big|_{a=0} \geq -\frac{1}{2} \partial_t W_t^2(\mu^0, \mu^1)
   \]
   (“dynamic convexity”).
2. For all \( 0 \leq s < t \leq T \) and \( \mu, \nu \in P \)
   \[
   W_s(\hat{P}_{t,s} \mu, \hat{P}_{t,s} \nu) \leq W_t(\mu, \nu)
   \]
   (“transport estimate”).
3. For all \( u \in \text{Dom}(E) \) and all \( 0 < s < t < T \)
   \[
   |\nabla_t (P_{t,s} u)|^2 \leq P_{t,s}(|\nabla_s u|^2)
   \]
   (“gradient estimate”).
4. For all \( 0 < s < t < T \) and for all \( u_s, g_t \in \mathcal{F} \) with \( g_t \geq 0, g_t \in L^\infty, \ u_s \in \text{Lip}(X) \), and for a.e. \( r \in (s,t) \)
   \[
   \Gamma_2,r(u_r)(g_r) \geq \frac{1}{2} \int \hat{\Gamma}_r(u_r)(g_r) \, dm_r
   \]
   (“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 following regularity assumption is satisfied:
   \[
   u_r \in \text{Lip}(X) \text{ for all } r \in (s, t) \text{ with } \sup_{r,x} \text{lip}_r u_r(x) < \infty.
   \]

Here and in what follows,

\[
\Gamma_2,r(u_r)(g_r) := \int \left[ \frac{1}{2} \Gamma_r(u_r) \Delta_r g_r + (\Delta_r u_r)^2 g_r + \Gamma_r(u_r, g_r) \Delta_r u_r \right] dm_r
\]

denotes the distribution-valued \( \Gamma_2 \)-operator (at time \( r \)) applied to \( u_r \) and tested against \( g_r \) and \( \hat{\Gamma}_r(u_r) := \lim_{\delta \to 0} \frac{1}{\delta} (\Gamma_{r+\delta}(u_r) - \Gamma_r(u_r)) \).
denotes any subsequential weak limit of $\frac{1}{2\delta}(\Gamma_{t+\delta} - \Gamma_{t-\delta}) (\mu_t)$ in $L^2((s, t) \times X)$.

**EVI Characterization of the Dual Heat Flow**

Recall that we started with the heat equation (acting on functions, forward in time) as a forward gradient flow for the time-dependent Dirichlet energy. By duality, we defined the dual heat flow (acting on probability measures, backward in time). This turns out to be the backward gradient flow for the Boltzmann entropy—in a very precise, strong sense—and it is the only one with this property.

**Theorem 1.8.** Each of the assertions of the previous theorem implies that the dual heat flow $t \mapsto \mu_t = \tilde{P}_{t, s} \mu$ is the unique dynamical (backward) EVI$^-$-gradient flow for the Boltzmann entropy $S$ in the following sense: For every $\mu \in \text{Dom}(S)$ and every $\tau < T$ the absolutely continuous curve $t \mapsto \mu_t$ satisfies

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

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

**Characterization of Super-$N$-Ricci Flows**

For static metric measure spaces, it turns 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 led to the so-called curvature-dimension condition $\text{CD}(K, N)$. In a similar spirit, in [51] the notion of super-Ricci flows for time-dependent metric measure spaces was tightened up towards super-$N$-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 the second author [17].

**Theorem 1.9.** For each $N \in (0, \infty)$ the following are equivalent:

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

$$\partial^+_a S_t(\mu^a)|_{a = 1} - \partial^-_a S_t(\mu^a)|_{a = 0^+} \geq$$

$$-\frac{1}{2} \frac{d}{dt} W^2_t(\mu^0, \mu^1) + \frac{1}{N} |S_t(\mu^0) - S_t(\mu^1)|^2.$$

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

$$W^2_s(\tilde{P}_{t, s} \mu, \tilde{P}_{t, s} \nu) \leq W^2_t(\mu, \nu) - \frac{2}{N} \int_s^t \left[ S_r(\tilde{P}_{t, r} \mu) - S_r(\tilde{P}_{t, r} \nu) \right]^2 dr.$$
(IIIₜ) For all \( u \in \text{Dom}(\mathcal{E}) \) and all \( 0 < s < t < T \)

\[
\|\nabla_t (P_{t,s}u)\|^2 \leq P_{t,s}(|\nabla_s (u)|^2) - \frac{2}{N} \int_s^t (P_{t,r} \Delta_r P_{r,s} u)^2 \, dr.
\]

(IVₜ) For all \( 0 < s < t < T \) and for all \( u, g_t \in \mathcal{F} \) with \( g_t \geq 0, g_t \in L^\infty, u_s \in \text{Lip}(X) \), the regularity assumption (1.7) is satisfied and for a.e. \( r \in (s,t) \)

\[
\Gamma_{2,r}(u_r)(g_r) \geq \frac{1}{2} \int \Gamma_r (u_r) g_r \, dm_r + \frac{1}{N} \left( \int \Delta_r u_r g_r \, dm_r \right)^2
\]

\begin{align*}
\text{("dynamic Bochner inequality" or “dynamic Bakry-Emery condition")} \\
\text{where } u_r = P_{r,s} u_s \text{ and } g_r = P^*_{r,t} g_t.
\end{align*}

Remark 1.10.

(a) In (Iₜ), the requested property for a.e. \( t \) will imply that it holds true for all \( t \in (0, T) \).

(b) The transport estimate (IIₜ) implies the stronger property

\[
W^2_2(\tilde{P}_{t,s} \mu, \tilde{P}_{t,s} \nu) \leq W^2_2(\mu, \nu) - \frac{2}{N} \int_0^1 \int \left( \delta_a S_r(\rho^a) \right)^2 \, da \, dr
\]

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

(c) Under slightly more restrictive assumptions on \((X, d_1, m_1)\)—namely, \( C \)-dependence of \( t \mapsto \log d_t \) instead of Lipschitz continuity—in a subsequent work of the first author [29] a refined version of the dynamic Bochner inequality (IVₜ) will be deduced with estimate (1.11) for every \( r \) and all \( u_r, g_r \) in respective domains, without requiring that they are solutions to heat and adjoint heat equations, respectively.

(d) Note that the regularity assumption (1.7) 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 (1.2) and the RCD\((-K, \infty)\)-condition for the static mm-spaces \((X, d_1, m_1)\).

\[\text{Super-}(K, N)\text{-Ricci Flows}\]

A more general version of the previous theorem will deal with the equivalences to \textit{dynamic} \((K, N)\)-convexity of the Boltzmann entropy as introduced in [51]. To simplify the presentation, however, we will restrict ourselves here to the case \( K = 0 \). Indeed, we would not expect new challenges or novel insights from the more general case \((K, N)\) since this can be easily transformed into the case \((0, N)\) by means of a simple rescaling of time and space.

\[\text{Theorem 1.11. Assume the time-dependent mm-space } (X, d_t, m_t), \ t \in I, \text{ is a super-}\( (K, N)\)-Ricci flow in the sense that for a.e. } t \in I \text{ and every } W_t\text{-geodesic}\]
\( (\mu^a), a \in [0, 1], \) in \( \mathcal{P} \) with \( \mu^0, \mu^1 \in \text{Dom}(S) \)

\[
\frac{1}{2} \partial^+_a W^2_{\tau(t)}(\mu^0, \mu^1) \geq \frac{1}{2} \partial^- W^2_{\tau(t)}(\mu^0, \mu^1) + \frac{1}{N} |S_{\tau(t)}(\mu^0) - S_{\tau(t)}(\mu^1)|^2 + K W^2_{\tau(t)}(\mu^0, \mu^1).
\]

Then for each \( C \in \mathbb{R} \) the time-dependent mm-space \((X, \tilde{d}_t, \tilde{m}_t), t \in \tilde{T},\) is a super-\(N\)-Ricci flow if we let

\[
\tilde{d}_t = e^{-K \tau(t)} d_{\tau(t)}, \quad \tilde{m}_t = m_{\tau(t)}, \quad \tau(t) = \frac{-1}{2K} \log(C - 2Kt),
\]

and \( \tilde{T} = \{ \tau(t) : t \in I, 2Kt < C \} \).

**Proof.** Put \( \tilde{d} = e^{-K \tau(t)} d_{\tau(t)} \). Then every \( \tilde{W}_\tau \)-geodesic will be a \( W_{\tau(t)} \)-geodesic. Therefore, neither the term \( \frac{1}{N} |S_{\tau(t)}(\mu^0) - S_{\tau(t)}(\mu^1)|^2 \) nor the term

\[
\frac{1}{2} \partial^+_a S_{\tau(t)}(\mu^a)|_{a=1} - \partial^- S_{\tau(t)}(\mu^a)|_{a=0+}
\]

in (1.12) will be changed by the transformation \( d \mapsto \tilde{d} \). Moreover,

\[
\frac{1}{2} \partial^- \tilde{W}^2_{\tau(t)}(\mu^0, \mu^1) = e^{-2K \tau(t)} \left[ -K \partial_t \tau(t) \cdot W_{\tau(t)} + (\partial^- W_{\tau(t)})(\tau(t) -) \cdot \partial_t \tau(t) \right] \cdot W_{\tau(t)}
\]

\[
= e^{-2K \tau(t)} \cdot \partial_t \tau(t) \cdot \left[ -K \cdot W^2 + \frac{1}{2} \partial^- W^2 \right](\tau(t) -)
\]

\[
= \left[ -K \cdot W^2 + \frac{1}{2} \partial^- W^2 \right](\tau(t) -).
\]

Thus (1.12) implies

\[
\frac{1}{2} \partial^- \tilde{W}^2_{\tau(t)}(\mu^0, \mu^1) = \left[ -K \cdot W^2 + \frac{1}{2} \partial^- W^2 \right](\tau(t) -)
\]

\[
\geq -\partial^+_a S_{\tau(t)}(\mu^a)|_{a=1} + \partial^- S_{\tau(t)}(\mu^a)|_{a=0+} + \frac{1}{N} |S_{\tau(t)}(\mu^0) - S_{\tau(t)}(\mu^1)|^2
\]

which proves the dynamic \(N\)-convexity of \( \tilde{S} \) and thus the super-\(N\)-Ricci flow property of \((X, \tilde{d}_t, \tilde{m}_t), t \in \tilde{T},\).

**Corollary 1.12.** For each \( N \in (0, \infty) \) and \( K \in \mathbb{R} \) the following are equivalent:
(I<sub>K,N</sub>) For a.e. \( t \in (0,T) \) and every \( W_\tau \)-geodesic \((\mu^a)_{a \in [0,1]}\) in \( \mathcal{P} \) with \( \mu^0, \mu^1 \in \text{Dom}(S) \)

\[
\dot{\gamma}_a^+ S_t(\mu^a)|_{a=1-} - \dot{\gamma}_a^- S_t(\mu^a)|_{a=0+} \\
\geq -\frac{1}{2} \dot{\gamma}_t W_\tau^2(\mu^0, \mu^1) + K \cdot W_\tau^2(\mu^0, \mu^1) \\
+ \frac{1}{N} |S_t(\mu^0) - S_t(\mu^1)|^2.
\]

(1.13)

(II<sub>K,N</sub>) For all \( 0 \leq s < t \leq T \) and \( \mu, v \in \mathcal{P} \)

\[
e^{-2Ks} W_s^2(\tilde{P}_{t,s} \mu, \tilde{P}_{t,s} v) \leq e^{-2Kt} W_t^2(\mu, v) - \frac{2}{N} \int_s^t e^{-2Kr} [S_r(\tilde{P}_{t,r} \mu) - S_r(\tilde{P}_{t,r} v)]^2 \, dr.
\]

(1.14)

(III<sub>K,N</sub>) For all \( u \in \text{Dom}(\mathcal{E}) \) and all \( 0 < s < t < T \)

\[
e^{2Kt} |\nabla_t (P_{t,s} u)|^2 \leq e^{2Ks} P_{t,s} (|\nabla_s (u)|^2) - \frac{2}{N} \int_s^t e^{2Kr} (P_{t,r} \Delta_r P_{r,s} u)^2 \, dr.
\]

(1.15)

(IV<sub>K,N</sub>) For all \( 0 < s < t < T \) and for all \( u_s, g_t \in \mathcal{F} \) with \( g_t \geq 0, g_t \in L^\infty, u_s \in \text{Lip}(X) \), the regularity assumption \([17]\) is satisfied and for a.e. \( r \in (s,t) \)

\[
\Gamma_{2,r}(u_r)(g_r) \geq \frac{1}{2} \int \Gamma_r(u_r) g_r \, dm_r + K \int \Gamma_r(u_r) g_r \, dm_r \\
+ \frac{1}{N} \left( \int \Delta_r u_r g_r \, dm_r \right)^2,
\]

(1.16)

where \( u_r = P_{r,s} u_s \) and \( g_r = P_{t,r}^* g_t \).

PROOF. As in the proof of the previous theorem, consider the time-dependent mm-space \((X, d_t, \tilde{m}_t), t \in \tilde{T}\), with \( d_t = e^{-K\tau(t)} d_{\tau(t)}, \tilde{m}_t = m_{\tau(t)} \), and \( \tilde{T} = \{\tau(t) : t \in I, 2Kt < C\} \) where \( \tau(t) = \frac{1}{2K} \log(C - 2Kt) \). Then

\[
\tilde{W}_2^2 = e^{-2K\tau} W_2^2, \quad \tilde{\Gamma}_t = e^{2K\tau} \Gamma_t, \quad \tilde{\Delta}_t = e^{2K\tau} \Delta_t, \\
\Gamma_{2,t} = e^{2K\tau} \Gamma_{2,\tau}, \quad \tau_t = e^{2K\tau}.
\]

Moreover, \( \tilde{P}_{t,s} = P_{\tau(t),\tau(s)} \). Thus for \((X, d_t, \tilde{m}_t), t \in \tilde{T}\), each of the statements (I<sub>N</sub>)–(IV<sub>N</sub>) is obviously equivalent to the corresponding statement (I<sub>K,N</sub>)–(IV<sub>K,N</sub>) for \((X, d_t, m_t), t \in I\). For instance, the equivalence “(II<sub>N</sub>) for \((X, d_t, \tilde{m}_t) \Leftrightarrow (II<sub>K,N</sub>) for \((X, d_t, m_t)\)” follows from the fact that

\[
e^{-2K\tau} W_2^2 - e^{-2K\sigma} W_2^2 = \tilde{W}_2^2 - \tilde{W}_s^2
\]
for $\tau = \tau(t)$ and $\sigma = \tau(s)$ and
\[
\frac{2}{N} \int_s^\tau \left[ \tilde{S}_r(\tilde{P}_{t,r} \mu) - S_r(\tilde{P}_{t,r} \nu) \right]^2 \, dr = \frac{2}{N} \int_0^\tau e^{-2Kr} \left[ S_r(\tilde{P}_{t,r} \mu) - S_r(\tilde{P}_{t,r} \nu) \right]^2 \, dr.
\]

\section*{Discussion of Standing Assumptions}

Let us briefly comment on the assumptions that we imposed throughout this introduction and for major parts of this paper.

Let us start with the discussion on the a priori assumption that each of the static spaces satisfies a Riemannian curvature-dimension condition.

\begin{remark}
Let a time-dependent mm-space $(X, d_t, m_t)$, $t \in I$, be given that satisfies all the assumptions mentioned in the beginning of this section but with no Riemannian curvature-dimension condition 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 $(I_N)$ of the Main Theorem \ref{thm:main} implies that for a.e. $t \in I$ the static space $(X, d_t, m_t)$ satisfies an $\text{RCD}^* (-L, N)$ condition.
\end{remark}

\begin{proof}
$(I_N)$ together with the log-Lipschitz bound (1.2) implies that along all $W_t$-geodesics
\[
\partial_a^+ S_t(\mu^a)|_{a=1} - \partial_a^- S_t(\mu^a)|_{a=0+} \geq - L \cdot W_t^2(\mu^0, \mu^1) + \frac{1}{N} \left| S_t(\mu^0) - S_t(\mu^1) \right|^2.
\]
This yields the $\text{RCD}^* (-L, N)$-condition in combination with the absolute continuity of $a \mapsto S_t(\mu^a)$; cf. \cite{51}.
\end{proof}

Next, we will discuss the assumption (1.2) concerning log-Lipschitz continuity of $t \mapsto d_t$.

\begin{remark}
Let $(M, g_t)$, be a time-dependent Riemannian manifold and let $(X, d_t, m_t)$ be the induced time-dependent mm-space.

(i) Then for any $L_1, L_2 \in [-\infty, \infty]$
\[
L_1 \leq \frac{1}{t-s} \log \frac{d_t}{d_s} \leq L_2 \iff L_1 g_t \leq \frac{1}{2} \partial_t g_t \leq L_2 g_t.
\]
Moreover, if $(M, g_t)$ evolves as a Ricci flow, then the previous assertions are equivalent to
\begin{equation}
-L_2 g_t \leq \text{Ric}_g t \leq -L_1 g_t.
\end{equation}
\end{remark}
If \((M, g_t)\) is a super-Ricci flow, then instead we merely have the implications
\[
\frac{1}{t-s} \log \frac{d_t}{d_s} \leq L_2 \quad \implies \quad -L_2 g_t \leq \text{Ric}_g,
\]
and
\[
L_1 \leq \frac{1}{t-s} \log \frac{d_t}{d_s} \quad \iff \quad \text{Ric}_g \leq -L_1 g_t.
\]
The proof is obvious. Similar assertions hold for the log-Lipschitz continuity of \(t \mapsto m_t\).

(ii) For Ricci flows of Riemannian manifolds, we can write \(m_t = e^{-(f_t - f_s)} m_s\) for all \(s < t\) with \(f_t - f_s = \int_s^t \text{scal}_g \, dr\). Thus
\[
L_1 \leq \frac{1}{t-s} \log \frac{d m_t}{d m_s} \leq L_2 \quad \iff \quad -L_2 \leq \text{scal}_g \leq -L_1.
\]
Super-Ricci flows allow for arbitrary time dependence of the exponential weight functions \(f_t\). Their regularity in time does not impose any a priori restriction on the metric tensors of the underlying space.

(iii) Condition (1.17) with finite \(L_1, L_2\) rules out Ricci flows running through singularities. In particular, it will not allow collapsing or changing topological type.

Related Works

Our main results, Theorem 1.7 and Theorem 1.9, combine and extend two previous—hitherto unrelated—lines of developments:

- results in the setting of smooth families of time-dependent Riemannian manifolds that characterize solutions to \(\text{Ric} + \frac{1}{2} \partial_t g_t \geq 0\) on \(I \times M\) (super-Ricci flows), e.g., by means of the monotonicity property (II) in terms of the \(L^2\)-Wasserstein metric for the dual heat flow, initiated by work by McCann and Topping [39]; for subsequent work in this direction that also includes equivalences with gradient estimates (III) and coupling properties of backward Brownian motions; see, e.g., Topping [53], Philipowski-Kuwada [32,33], Arnaudon-Coulibaly-Thalmaier [9], Lakzian-Munn [34], Li-Li [35].
- results for (static) metric measure spaces by Ambrosio-Gigli-Savare [6] as well as by Erbar-Kuwada-Sturm [17].

Indeed, Theorem 1.7 and Theorem 1.9 extend the main results from [6] and from [17] (cf. also [7]) to the time-dependent setting. Partly, our proofs also provide new and simpler arguments in the static setting, for instance, for the implication (III\(_N\)) \(\Rightarrow\) (II\(_N\)). Even though we benefited very much from the powerful, detailed calculus on mm-spaces developed in [4–6] and pushed forward in [1,2,7,20], in many cases we had to develop entirely new strategies and to derive numerous auxiliary estimates and regularity assertions. For the proof of implication (II\(_N\)) \(\Rightarrow\) (III\(_N\)),...
we followed the argumentation of [12] and carried over their arguments from the static to the dynamic setting.

The analysis of the heat flow on time-dependent spaces (either Dirichlet spaces or metric measure spaces) seems to be completely new.

Even in the smooth case, the characterization (I) of super-Ricci flows in terms of the so-called dynamic convexity (as introduced in the accompanying paper [51] by the second author) was not known before.

Work in Progress

The current paper, together with the previous paper by the second author [51], will lay the foundations for a broad systematic study of (super-)Ricci flows in the context of mm-spaces with various subsequent publications in preparation that, among others, will address the following challenges:

- time-discrete gradient flow scheme à la Jordan-Kinderlehrer-Otto for the heat equation and its dual as gradient flows of energy and entropy, respectively, [28];
- improved dynamic Bochner inequality; $L^p$-gradient and $L^q$-transport estimates; construction and optimal coupling of Brownian motions on time-dependent mm-spaces [29];
- geometric functional inequalities on time-dependent mm-spaces—in particular, local Poincaré, logarithmic Sobolev, and dimension-free Harnack inequalities—and characterization of super-Ricci flows in terms of them [30];
- synthetic approaches to upper Ricci bounds [52] and rigidity results for Ricci flat metric cones [18].

Preliminary Remarks

We use $\partial_t$ as a short-hand notation for $\frac{d}{dt}$. Moreover, we let

$$\partial_t^+ u(t) = \limsup_{s \to t} \frac{1}{t-s} (u(t) - u(s)) \quad \text{and} \quad \partial_t^- u(t) = \liminf_{s \to t} \frac{1}{t-s} (u(t) - u(s)).$$

In what follows, $r,s,t$ always denote time parameters whereas $a,b$ denote curve parameters.

1.4 Sketch of the Argumentation for the Main Result

Structure of Proof of Theorem[1.9] In Section[4] we present the implications $(\text{I}_N) \implies (\text{II}_N)$ and $(\text{III}_N) \implies (\text{II}_N)$ as well as the converse of the latter in the case $N = \infty$. Section[5] is devoted to the proof of the equivalence $(\text{III}_N) \iff (\text{IV}_N)$ as well as to the proof of the implication $(\text{II}_N) \implies (\text{IV}_N)$.

In Section[6] 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 the Appendix (p. 2597), which is devoted to the study of dynamical EVI flows in a general framework. Here in particular, it will also be shown that $(\text{III}_N)$ and $\text{EVI}^{-} \implies (\text{I}_N)$. \hfill $\square$

Let us now briefly sketch the arguments for each of the implications.

$$(\text{I}_N) \implies (\text{II}_N)$$

Given two solutions to the dual heat flow $(\mu_t, v_t)$, for fixed $t$ we connect the measures $\mu_t = u m_t$ and $v_t = v m_t$ by a $W_t$-geodesic $(\eta^a)$, $a \in [0, 1]$, and choose a pair of functions $\phi, \psi$ in duality w.r.t. $\frac{1}{2} W_t^2$ and optimal for the pair $\mu_t, v_t$ (Kantorovich potentials); see Figure 1.3. (Note that in the smooth Riemannian setting the $W_t$-geodesic and the Kantorovich potentials are linked through the relation $\eta^a \equiv \exp(-a \nabla\phi) \ast \mu_t = (\exp(- (1 - a) \nabla\psi)) \ast v_t$.)

In the general setting, we deduce with $u = \frac{d\mu_t}{dm_t}, v = \frac{dv_t}{dm_t}$

- $\frac{1}{2} \partial_r^- W_t^2 (\mu_t, v_t)|_{r=t^+} \geq \mathcal{E}_t (\phi, u) + \mathcal{E}_t (\psi, v)$ from Kantorovich duality,
- $\mathcal{E}_t (\phi, u) + \mathcal{E}_t (\psi, v) \geq -\partial_a S_t (\eta^a)|_{a=0^+} + \partial_a S_t (\eta^a)|_{a=1^-}$ from semic-\nconvexity of $S_t$, and
- $\frac{1}{2} \partial_r^- W_t^2 (\mu_t, v_t)|_{r=t^-} \geq -\partial_a S_t (\eta^1^-) + \partial_a S_t (\eta^{0^+}) + \frac{1}{N} [S_t (\mu_t) - S_t (v_t)]^2$ from the defining property of a super-$N$-Ricci flows.

Adding up these estimates yields

$$\frac{1}{2} \partial_r^- W_t^2 (\mu_t, v_t)|_{r=t^+} + \frac{1}{2} \partial_r^- W_r^2 (\mu_t, v_t)|_{r=t^-} \geq \frac{1}{N} [S_t (\mu_t) - S_t (v_t)]^2.$$ 

We can now replace the left-hand side by $\frac{1}{2} \partial_r^- W_t^2 (\mu_t, v_t)$ by employing a careful time shift argument, which then proves the claim.

$$(\text{II}_N) \implies (\text{IV}_N)$$

Given a Lipschitz function $u$ and a probability density $g$ (w.r.t. $m_\tau$), let $g_r = P_{\tau, r}^* g$, $u_r = P_{r, \sigma} u$, and $h_r := \int g_r \Gamma_r (u_r) dm_r$ for $0 < \sigma < r < \tau < T$. 

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure13.png}
\caption{Figure 1.3.}
\end{figure}
By duality we already know that the transport estimate (IIₙ) implies the infinite-dimensional gradient estimate (III), which helps us to deduce that

$$h₂ - hₚ \geq \int_σ^\tau \left[ -2\Gamma_{2,r}(u_r)(g_r) + \int \tilde{\Gamma}_r(u_r) g_r m_r \right] dr.$$

To improve this inequality, we follow the approach initiated by [12] and consider the perturbation of $g_τ$ given by

$$g_τ^{α,a} := g_τ(1 - α[Δ_τ u_σ + Γ_τ(log g_τ, u_σ)])$$

for small $α > 0$. It can be interpreted as the Taylor expansion of the $W_τ$-geodesic starting in $g_τ$ with initial velocity $u_σ$. The transport estimate (IIₙ) applied to the probability measures $g_τ m_τ$ and $g_τ^{α,a} m_τ$ gives us for all $α > 0$

$$W_α^2(\tilde{P}_τ, g_τ m_τ, \tilde{P}_τ, g_τ^{α,a} m_τ)$$

$$- W_α^2(g_τ m_τ, g_τ^{α,a} m_τ)$$

$$\leq -\frac{2}{N} \int_σ^\tau [S_τ(\tilde{P}_τ, g_τ m_τ) - S_τ(\tilde{P}_τ, g_τ^{α,a} m_τ)]^2 dr.$$

In the limit $α \to 0$ we eventually end up with

$$h₂ - hₚ \leq -\frac{2}{N} \int_σ^\tau \left( \int Δ_r u_r g_r dm_r \right)^2 dr.$$

Together with the previous lower estimate for $h₂ - hₚ$, this proves the claim.

(IVₙ) $\iff$ (IIIₙ)

This is—modulo regularity issues—a simple, well-known (cf. [51, theorem 5.5]) differentiation-integration argument for the function

$$r \mapsto \int P_{t,r}^* g \cdot Γ_r(P_{r,s} u) dm_r.$$

(IIIₙ) $\implies$ (IIₙ)

Given any regular curve $(μ²_a)_{a ∈ [0,1]}$ and $τ ∈ I$ we will study the evolution of this curve under the dual heat flow. More precisely, we analyze the growth of the action

$$A_τ(μ_τ) := \int_0^1 |\dot{μ}²_t|_t da = \int_0^1 \int_X |∇_τ Φ_τ|^2 dμ²_t da$$

of the curve $(μ²_a), a ∈ [0,1]$, for $t < τ$ where $μ²_t = \tilde{P}_{τ,t} μ²_t = u²_t m_t$ and $(Φ²_t)_{a ∈ [0,1]}$ denotes the velocity potentials in the static space $(X, d_t, m_t)$. For $s < t$ we approximate the action $A_s(μ_τ²)$ by

$$\sum_{i=k}^1 \frac{1}{a_i - a_i-1} W_²{s}(μ²_{a_i-1}, μ²_{a_i}).$$
the latter in terms of $W_s$-Kantorovich potentials, and finally by means of the interpolating Hopf-Lax semigroup. By applying the Bakry-Ledoux gradient estimate, $(\text{III}_N)$ then allows us to estimate

$$
2\varepsilon + \frac{1}{t-s} [A_t(\mu_t') - A_s(\mu_s')] \geq \frac{2}{N + \varepsilon} \left| \int_0^1 \int_X \nabla_x \Phi_t^a \cdot \nabla_x \log u_t^a \, d\mu_t^a \, da \right|^2 = \frac{2}{N + \varepsilon} \left| S_t(\mu_t^1) - S_t(\mu_t^0) \right|^2
$$

for each $\varepsilon > 0$ provided that $s$ is sufficiently close to $t$. Passing to the limit $s \uparrow t$ and integrating the result from $s$ to $\tau$ yields

$$A_s(\mu_s') \leq A_\tau(\mu_\tau') - \frac{2}{N} \int_s^\tau \left[ S_t(\mu_t^0) - S_t(\mu_t^1) \right]^2 \, dt.$$

This indeed proves the claim since

$$W_\tau^2(\mu_0^0, \mu_1^1) = \inf \{ A_t(\mu_t') : (\mu_t^a)_{a \in [0,1]} \text{ regular curve connecting } \mu_0^0, \mu_1^1 \}$$

for any $\mu_0^0, \mu_1^1$, and $\tau$ whereas $W_\tau^2(\mu_s^0, \mu_s^1) \leq A_s(\mu_s')$ for all $s < \tau$.

$(\text{III}_N) \implies (\text{I}_N)$

To deduce the dynamic convexity of the Boltzmann entropy $S_t$, let a $W_t$-geodesic $(\mu_t^a), a \in [0,1]$, be given and consider its evolution $\mu_s^a := \hat{P}_{t,s} \mu_t^a, s < t$, under the dual heat flow. Then on one hand

$$W_t^2(\mu_t^0, \mu_t^1) = \frac{1}{a} W_t^2(\mu_t^0, \mu_t^a) + \frac{1}{1 - 2a} W_t^2(\mu_t^a, \mu_t^{1-a}) + \frac{1}{a} W_t^2(\mu_t^{1-a}, \mu_t^1)$$

(1.18)

for all $a \in (0, \frac{1}{2})$, whereas on the other

$$W_s^2(\mu_t^0, \mu_t^1) \leq \frac{1}{a} W_s^2(\mu_t^0, \mu_s^a) + \frac{1}{1 - 2a} W_s^2(\mu_s^a, \mu_s^{1-a}) + \frac{1}{a} W_s^2(\mu_s^{1-a}, \mu_t^1).$$

(1.19)

We already know that the gradient estimate $(\text{III}_N)$ implies the transport estimate $(\text{II}_N)$, and the latter yields

$$\liminf_{s \nearrow t} \frac{1}{t-s} \frac{1}{1 - 2a} \left[ W_t^2(\mu_t^a, \mu_t^{1-a}) - W_s^2(\mu_s^a, \mu_s^{1-a}) \right] \geq \frac{2}{N} \frac{1}{1 - 2a} \left[ S_t(\mu_t^a) - S_t(\mu_t^{1-a}) \right]^2.$$
The EVI property to be discussed below will allow us to estimate
\[
\liminf_{s \to t} \frac{1}{t-s} \left[ W_t^2(\mu^0_t, \mu^a_t) - W_s^2(\mu^0_s, \mu^a_s) \right] \geq \frac{2}{a} \left[ S_t(\mu^a_t) - S_t(\mu^0_t) \right] - L a W_t^2(\mu^0_t, \mu^1_t),
\]
as well as
\[
\liminf_{s \to t} \frac{1}{t-s} \left[ W_t^2(\mu^{1-a}_t, \mu^1_t) - W_s^2(\mu^{1-a}_s, \mu^1_s) \right] \geq \frac{2}{a} \left[ S_t(\mu^{1-a}_t) - S_t(\mu^1_t) \right] - L a W_t^2(\mu^0_t, \mu^1_t).
\]
Using (1.18) together with (1.19) and adding up the last three inequalities we obtain after letting \(a \searrow 0\) (see Figure 1.4):
\[
\liminf_{s \to t} \frac{1}{t-s} \left[ W_t^2(\mu^0_t, \mu^1_t) - W_s^2(\mu^0_s, \mu^1_s) \right] \geq \frac{2}{N} \left[ S_t(\mu^0_t) - S_t(\mu^1_t) \right]^2 + 2 \partial_a^- S_t(\mu^a_t) \big|_{a=0^+} - 2 \partial_a^+ S_t(\mu^a_t) \big|_{a=1^-}.
\]

In order to prove the EVI property, we follow the approach in [6,17] and extend their arguments to the time-dependent setting. We show that the gradient estimate implies that the dual heat flow is a dynamic EVI-\(-\)gradient flow. For this we introduce in Section 6.1 a dual formulation \(\bar{W}_{s,t}\) of our time-dependent distance \(W_{s,t}\).

For each fixed \(s < t\) we take a regular curve \((\rho^a), a \in [0, 1]\), approximating the \(W_t\)-geodesic joining \(\sigma\) and \(\mu_t := \bar{P}_{t,s} \mu\) where \(\mu, \sigma \in \mathcal{P}(X)\) are fixed. We then apply the dual heat flow \(\rho_{a,t} := \bar{P}_{t,s+a(t-s)} \rho^a\) to the regular curve (cf. Figure 1.5) and eventually show using (III) that
\[
\frac{1}{2} \bar{W}_{s,t}^2(\rho_{1,\theta}, \rho_{0,\theta}) - (t-s)(S_t(\rho_{1,\theta}) - S_s(\rho_{0,\theta})) \leq \int_0^1 \left[ \frac{1}{2} |\dot{\theta}|^2_t + (t-s)^2 \int \dot{\theta}_a d\rho_{a,\theta} \right] da.
\]
\[
\frac{1}{2} \widetilde{W}_{s,t}^2(\mu_s, \sigma) - (t-s)(S_t(\sigma) - S_s(\mu_s)) \leq \frac{1}{2} W_t^2(\mu_t, \sigma) - (t-s)^2 \int_0^1 \int f_{\partial(a)} d\rho_{a,\theta} \, da.
\]

In contrast to the static case we obtain the additional error term
\[
(t-s)^2 \int_0^1 \int f_{\partial(a)} d\rho_{a,\theta} \, da,
\]
which, however, vanishes after dividing by \(t-s\) and letting \(s \nearrow t\). Thus
\[
S_t(\mu_t) - S_t(\sigma) \leq \liminf_{s \nearrow t} \frac{1}{2(t-s)} \left( W_t^2(\mu_t, \sigma) - \widetilde{W}_{s,t}^2(\mu_s, \sigma) \right) = \frac{1}{2} \partial_s^- W_{s,t}^2(\mu_s, \sigma) \bigg|_{s=t-}.
\]

Note that the log-Lipschitz continuity of the distance allows us to estimate the last term from above by
\[
\frac{1}{2} \partial_s^- W_{s,t}^2(\mu_s, \sigma) \bigg|_{s=t-} + \frac{L}{2} W_t^2(\mu_t, \sigma).
\]

2 The Heat Equation for Time-Dependent Dirichlet Forms

2.1 The Heat Equation

Let us choose here a setting that is slightly more general than for the rest of the paper. We assume that we are given a Polish space \(X\) and a \(\sigma\)-finite reference measure \(m_\circ\) on it that is assumed to have full topological support. Moreover, we assume that we are given a strongly local Dirichlet form \(\mathcal{E}_\circ\) with domain \(\mathcal{D}(\mathcal{E}_\circ)\) on \(\mathcal{H} = L^2(X, m_\circ)\) and with square field operator \(\Gamma_\circ\) such that \(\mathcal{E}_\circ(u, v) = \int_X \Gamma_\circ(u,v) \, dm_\circ\) for all functions \(u, v \in \mathcal{F}\). 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 \(\int uv \, dm_\circ\) and \(\mathcal{E}_\circ(u, v) + \int uv \, dm_\circ,\) 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.

Recall that a Dirichlet form $\mathcal{E}_\circ$ on $L^2(X, m_\circ)$ is a densely defined, nonnegative symmetric form on $L^2(X, m_\circ)$ that is closed (which is equivalent to saying that the quadratic form is lower-semicontinuous on $L^2(X, m_\circ)$) and that satisfies the Markov property

$$\mathcal{E}_\circ(\xi \circ u) \leq \mathcal{E}_\circ(u) \quad \text{for all } \xi : \mathbb{R} \to \mathbb{R} \text{ 1-Lipschitz such that } \xi(0) = 0.$$ Here and in what follows, the same symbol will be used for a bilinear form and the quadratic form associated with it, i.e., $\mathcal{E}_\circ(u) = \mathcal{E}_\circ(u, u)$. The Dirichlet form $\mathcal{E}_\circ$ is called strongly local if $\mathcal{E}_\circ(u, v) = 0$ whenever $(u + c)v = 0 m_\circ$-a.e. for some $c \in \mathbb{R}$. We refer to [15] for a comprehensive study of Dirichlet forms and to [11] for the important role of the square field operator.

Let $I \subset \mathbb{R}$ be a bounded open interval, say $I = (0, T)$ for simplicity. In order to deal with time-dependent evolutions, following [48] we consider for $0 < s < T$ the Hilbert spaces

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

equipped with the respective norms $(\int_s^\tau \|u_t\|^2_{\mathcal{F}} + \|\partial_t u_t\|^2_{\mathcal{F}, \tau} \, dt)^{1/2}$. According to [45], lemma 10.3, the embeddings $\mathcal{F}(s, \tau) \subset C([s, \tau] \to \mathcal{H})$ hold true, which guarantee that values at $t = s$ and $t = \tau$ are well-defined.

Moreover, assume that we are given a 1-parameter family $(m_t)_{t \in (0, T)}$ of measures on $X$ such that $m_t = e^{-\int_t^\tau f_t \, dt} m_\circ$ for some bounded measurable function $f$ on $I \times X$ with $f_t \in \mathcal{F}$ and $\exists C$ s.t. $\forall t, x$

$$\Gamma_\circ(f_t)(x) \leq C.
$$

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

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

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

\(\forall k \in \mathbb{N}, \forall v, u_1, \ldots, u_k \in \mathcal{F} \cap L^\infty(X, m_\circ), \forall \Psi \in C^1(\mathbb{R}^k)\) with $\Psi(0) = 0$ [11],

- and all of them being uniformly comparable (“uniformly elliptic”) w.r.t. the reference form $\Gamma_\circ$ on $\mathcal{F}$, i.e., $\exists C$ s.t. $\forall t \in (0, T), \forall u \in \mathcal{F}, \forall x \in X$

$$\frac{1}{C} \Gamma_\circ(u)(x) \leq \Gamma_t(u)(x) \leq C \Gamma_\circ(u)(x). \tag{2.2}$$

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 $\text{Dom}(\mathcal{E}_t) = \mathcal{F}$ and a self-adjoint,
nonpositive operator $A_t$ on $L^2(X, m_t)$ with domain $\text{Dom}(A_t) \subset \mathcal{F}$ uniquely determined by the relations
\[
\int_X \Gamma_t(u, v) d m_t = \mathcal{E}_t(u, v) = -\int_X A_t u v d m_t
\]
for $u, v \in \mathcal{F}$. Recall that $u \in \text{Dom}(A_t)$ if and only if $u \in \mathcal{F}$ and $\exists C'$ such that $\mathcal{E}_t(u, v) \leq C' \cdot \|v\|_{L^2(m_t)}$ for all $v \in \mathcal{F}$.

**Definition 2.1.** A function $u$ is called the solution to the heat equation
\[
A_t u = \partial_t u \quad \text{on } (s, \tau) \times X
\]
if $u \in \mathcal{F}_{(s, \tau)}$ and if for all $w \in \mathcal{F}_{(s, \tau)}$
\[
(2.3) \quad -\int_s^\tau \mathcal{E}_t(u_t, w_t) dt = \int_s^\tau \langle \partial_t u_t, w_t e^{-f_t} \rangle_{\mathcal{F}^*, \mathcal{F}} dt
\]
where $\langle \cdot, \cdot \rangle_{\mathcal{F}^*, \mathcal{F}} = \langle \cdot, \cdot \rangle$ denotes the dual pairing. Note that thanks to (2.1), $w \in L^2((s, \tau) \to \mathcal{F})$ if and only if $w e^{-f} \in L^2((s, \tau) \to \mathcal{F})$.

Since $u_t \in \text{Dom}(A_t)$ (and thus $\partial_t u_t \in L^2$) for almost every $t$ by virtue of Theorem 2.12, we may equivalently rewrite the right-hand side of the above equation as
\[
\int_s^\tau \langle \partial_t u_t, w_t e^{-f_t} \rangle_{\mathcal{F}^*, \mathcal{F}} dt = \int_s^\tau \int_X \partial_t u_t \cdot (w_t e^{-f_t}) d m_\sigma dt
\]
\[
= \int_s^\tau \int_X \partial_t u_t \cdot w_t d m_t dt
\]
which allows for a more intuitive, alternative formulation of (2.3) as follows:
\[
-\int_s^\tau \mathcal{E}_t(u_t, w_t) dt = \int_s^\tau \int_X \partial_t u_t \cdot w_t d m_t dt.
\]

**Theorem 2.2.** For all $0 \leq s < \tau \leq T$ 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 $\lim_{t \to s^-} u_t = h$).

**Proof.** For each $t$ the bilinear form $\mathcal{E}_t^\circ$ on $\mathcal{F}$ is defined by
\[
\mathcal{E}_t^\circ(u, v) = -\int_X A_t u v d m_\sigma = \int_X \Gamma_t(u, v e^{f_t}) e^{-f_t} d m_\sigma
\]
\[
= \int_X [\Gamma_t(u, v) + v \Gamma_t(u, f_t)] d m_\sigma
\]
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 $w \in \mathcal{F}_{(s, \tau)}$
\[
-\int_s^\tau \mathcal{E}_t^\circ(u_t, w_t) dt = \int_s^\tau \int_X \partial_t u_t \cdot w_t d m_\sigma dt.
\]
(Indeed, we simply have to replace the test function $w_t$ by $w_t e^{f_t}$.)
Our assumptions on $\Gamma_t$ and $f_t$ guarantee that $\mathcal{E}_t^\circ$ for each $t$ is a closed coercive form with domain $\mathcal{F} = \text{Dom}(\mathcal{E}_\circ)$ on $\mathcal{H} = L^2(X, m_\circ)$, uniformly comparable to $\mathcal{E}_\circ$. For each $t$, the operator $A_t$ is a bounded linear operator from $\mathcal{F}$ to $\mathcal{F}^\ast$. Indeed,

$$\|A_t\|_{\mathcal{F} \to \mathcal{F}^\ast} = \sup_{u,v \in \mathcal{F}} \frac{|\mathcal{E}_t^\circ(u, v)|}{\|u\|_{\mathcal{F}}^{1/2} \cdot \|v\|_{\mathcal{F}}^{1/2}} \leq \sup_{u,v \in \mathcal{F}} \frac{1}{\|u\|_{\mathcal{F}}^{1/2} \cdot \|v\|_{\mathcal{F}}^{1/2}} \int_X |\Gamma_t(u, v)| \, dm_\circ$$

$$+ \sup_{u,v \in \mathcal{F}} \frac{1}{\|u\|_{\mathcal{F}}^{1/2} \cdot \|v\|_{\mathcal{F}}^{1/2}} \int_X |v \Gamma_t(u, f_t)| \, dm_\circ$$

$$\leq C \left( 1 + \|\Gamma(f_t)\|_{\mathcal{F}}^{1/2} \right)$$

if $C$ is chosen such that $|\Gamma_t(u, v)| \leq C \cdot \Gamma_\circ(u)^{1/2} \cdot \Gamma_\circ(v)^{1/2}$ 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 [37, chap. III, theorem 4.1, and remark 4.3]; see also [43, theorem 10.3]. (Note, however, that the latter assumes a continuity of $t \mapsto A_t$ in the operator norm, which is not really necessary.)

Remark 2.3. We denote this solution by $u_t(x) = P_{t,s} h(x)$. Then $(P_{t,s})_{0 \leq s \leq t < T}$ is a family of bounded linear operators on $\mathcal{H}$ that has the propagator property

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

for all $r \leq s \leq t$. For fixed $s$ and $h$ the function $t \mapsto P_{t,s} h$ is continuous in $\mathcal{H}$ (due to the embedding $\mathcal{F}(s,T) \subset C([s, T] \to \mathcal{H})$). 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 $h \in \mathcal{H}$

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

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 $m_\circ$ nor with respect to $m_t$ nor with respect to $m_s$.

2.2 The Adjoint Heat Equation

Definition 2.4. Given $0 \leq \sigma < t \leq T$, a function $v$ is called a solution to the adjoint heat equation

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

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 2.5. Assume (2.1) and

\begin{equation}
|f_t(x) - f_s(x)| \leq L|t - s|.
\end{equation}

(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

\[ P_{t,r}^* = P_{s,r}^* \circ P_{t,s}^* \quad (\forall r \leq s \leq t). \]

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

\[ \int P_{t,s} h \cdot g \, dm_t = \int h \cdot P_{t,s}^* g \, dm_s \quad (\forall g, h \in \mathcal{H}). \]

PROOF. For (i) and (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 $-\partial_s v = A_s v - (\partial_s f_s)v$.

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

\[ \int u_t v_t \, dm_t - \int u_s v_s \, dm_s = \int_t^s \partial_r u_r \, v_r \, dm_r \, dr + \int_t^s u_r \, \partial_r v_r \, dm_r \, dr \\
- \int_s^t u_r \, v_r \, \partial_r f_r \, dm_r \, dr \\
= \int_t^s \mathcal{E}_r(u_r, v_r) \, dr - \int_s^t \mathcal{E}_r(u_r, v_r) \, dr = 0. \]

Note, however, that—even under the assumption $m_\sigma(X) < \infty$—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 2.6. 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: $\text{Dom}(A_t) = \text{Dom}(A_t^*)$.

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

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

Boundedness of $\partial_t f$ implies that this is equivalent to $v \in \text{Dom}(A_t)$. \qed
In contrast to the form domains, the operator domains $\text{Dom}(A_t)$ in general will depend on $t$.

**Example 2.7.** Consider $\mathcal{H} = L^2(\mathbb{R}, dx)$ with $m_t(dx) = dx$ and

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

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

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

Thus $\text{Dom}(A_s) \neq \text{Dom}(A_t)$ for all $s \neq t$.

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

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 2.8.**

(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 $m_\sigma(X) < \infty$.

**Proof.** In both cases, straightforward calculations yield

$$e^{Lt} \partial_t \left[ e^{-Lt} \int u_t \log u_t \, dm_t \right] \leq \int (\log u_t + 1) \partial_t u_t \, dm_t$$

$$= -\int \Gamma_t(\log u_t) \, u_t \, dm_t \leq 0$$

and

$$\partial_s \int v_s \log v_s \, dm_s = \int (\log v_s + 1) \partial_s v_s \, dm_s - \int v_s \log v_s \cdot \partial_s f_s \, dm_s$$

$$= \int \Gamma_s(\log v_s) \, v_s \, dm_s + \int v_s \cdot \partial_s f_s \, dm_s$$

$$\geq -L \int v_s \, dm_s.$$ \qed
2.3 Energy Estimates

Throughout this section, assume (2.1) as well as (2.5) and in addition

\[(2.6) \quad |\Gamma_t(u) - \Gamma_s(u)| \leq 2L \cdot \int_s^t \Gamma_r(u) \, dr\]

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 \(u \in L^2((s, \tau) \to \mathcal{F}) \cap H^1((s, \tau) \to \mathcal{F}^*) \subset C((s, \tau) \to \mathcal{H})\) and

\[(2.7) \quad \int_s^t \mathcal{E}_t(u_t) \, dt \leq \frac{1}{2} \|u_s\|^2_{L^2(m_s)}.

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, nonpositive operator \(\tilde{A}_t\) on \(L^2(X, m_\circ)\) by

\[-\int_X \tilde{A}_t uv \, dm_\circ = \mathcal{E}_t(u, v) := \int_X \Gamma_t(u, v) \, dm_\circ\]

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

\[\tilde{A}_t u = A_t u + \Gamma_t(u, f_t)\]

Indeed,

\[-\int A_t uv \, dm_\circ = \int \Gamma_t(u, ve^{f_t})e^{-f_t} \, dm_\circ\]

\[= -\int \tilde{A}_t uv \, dm_\circ + \int \Gamma_t(u, f_t)v \, dm_\circ.

Next, consider the Hille-Yosida approximation \(\tilde{A}_t^\beta := (I - \delta \tilde{A}_t)^{-1}\tilde{A}_t\) of \(\tilde{A}_t\) on \(L^2(X, m_\circ)\), let \(\mathcal{E}_t^\beta(u, v) := -\int \tilde{A}_t^\beta uv \, dm_\circ\), and recall the well-known fact that \(\mathcal{E}_t^\beta(u, u) \ngeq \mathcal{E}_t(u, u)\) for each \(u \in \mathcal{F}\) as \(\delta \searrow 0\). More generally, we have the following:

**Lemma 2.9.** For all \(\alpha, \beta > 0\) with \(\beta - \alpha \leq \frac{1}{2}\), \(\mathcal{F} \subset \text{Dom}((I - \delta \tilde{A}_t)^{-\alpha} \tilde{A}_t^\beta)\) and for all \(u \in \mathcal{F}\):

\[u \in \text{Dom}(\tilde{A}_t^\beta) \iff \sup_{\delta > 0} \|(I - \delta \tilde{A}_t)^{-\alpha} \tilde{A}_t^\beta u\|_{L^2} < \infty\]

with \(\|(I - \delta \tilde{A}_t)^{-\alpha} \tilde{A}_t^\beta u\|_{L^2} \nless \|\tilde{A}_t^\beta u\|_{L^2}\) for \(\delta \searrow 0\).

**Proof.** For fixed \(t\) we apply the spectral theorem to the nonnegative 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 semibounded \(\Phi : \mathbb{R}_+ \to \mathbb{R}\)

\[\text{Dom}(\Phi(-\tilde{A}_t)) = \left\{ u \in \mathcal{H} : \int_0^\infty |\Phi(\lambda)|^2 \, dE_\lambda(u, u) \right\}\]
and \((\Phi(-\tilde{A}_t)u, v)_H = \int_0^\infty \Phi(\lambda)dE_\lambda(u, v)\). Thus, in particular, 
\[
\mathcal{F} = \left\{ u \in H : \int_0^\infty \lambda dE_\lambda(u, u) \right\}
\] 
and 
\[
\text{Dom}\left( (I - \delta \tilde{A}_t)^{-\alpha} \tilde{A}_t^\beta \right) = \left\{ u \in H : \int_0^\infty \left| \frac{\lambda^\beta}{(1 + \delta \lambda)^\alpha} \right|^2 dE_\lambda(u, u) \right\}.
\]
Moreover, by monotone convergence as \(\delta \searrow 0\), 
\[
\| (I - \delta \tilde{A}_t)^{-\alpha} \tilde{A}_t^\beta u \|^2_{L^2} = \int_0^\infty \left| \frac{\lambda^\beta}{(1 + \delta \lambda)^\alpha} \right|^2 dE_\lambda(u, u) \int_0^\infty \lambda^2 dE_\lambda(u, u) = \| \tilde{A}_t^\beta u \|^2_{L^2}.
\]

**Lemma 2.10.** For all \(\delta > 0\) and all \(u, v \in \mathcal{F}\), the map \(t \mapsto \tilde{E}_t^\delta(u, v)\) is absolutely continuous with 
\[
|\partial_t \tilde{E}_t^\delta(u, v)| \leq \frac{L}{2} \left[ \tilde{E}_t(u, u) + \tilde{E}_t(v, v) \right].
\]

**Proof.** For all \(\delta, u, v\) as above, let \(u_t^\delta = (I - \delta \tilde{A}_t)^{-1}u\) and \(v_t^\delta = (I - \delta \tilde{A}_t)^{-1}v\). Then 
\[
\partial_t \tilde{E}_t^\delta(u, v) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int \left[ (I - \delta \tilde{A}_{t+\epsilon})^{-1} \tilde{A}_{t+\epsilon}u - (I - \delta \tilde{A}_t)^{-1} \tilde{A}_tu \right] \cdot v \, dm_0
\]
\[
= \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int \left[ (I - \delta \tilde{A}_{t+\epsilon})^{-1} (\tilde{A}_{t+\epsilon} - \tilde{A}_t)(1 - \delta \tilde{A}_t)^{-1} u \right] \cdot v \, dm_0
\]
\[
= \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left[ \tilde{E}_t(u_t^\delta, v_t^\delta) - \tilde{E}_{t+\epsilon}(u_t^\delta, v_t^\delta) \right]
\]
\[
\leq \frac{L}{2} \lim_{\epsilon \to 0} \left[ \tilde{E}_t(u_t^\delta, u_t^\delta) + \tilde{E}_{t+\epsilon}(v_t^\delta, v_t^\delta) \right] = \frac{L}{2} \left[ \tilde{E}_t(u, u) + \tilde{E}_t(v, v) \right].
\]
Here we also used the fact that \(\tilde{E}_t(u_t^\delta, u_t^\delta) \geq \tilde{E}_t(u_t, u_t)\) as \(\delta \to 0\). 

**Lemma 2.11.** 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 \(\delta > 0\), 
\[
(2.8) \quad \int_s^\tau \int_X |(I - \delta \tilde{A}_t)^{-1/2} \tilde{A}_tu_t|^2 \, dm_0 \, dt \leq C \cdot [E_s(u_s) + \|u_s\|_{L^2(m_0)}^2].
\]
Thus, in particular, if \(u_s \in \mathcal{F}\) then \(u_t \in \text{Dom}(\tilde{A}_t)\) for a.e. \(t \in (s, \tau)\) and 
\[
(2.9) \quad \int_s^\tau \int_X |\tilde{A}_tu_t|^2 \, dm_0 \, dt \leq C \cdot [E_s(u_s) + \|u_s\|_{L^2(m_0)}^2].
\]
\textbf{THEOREM 2.12.} For all \( 0 < s < \tau < T \) and for all solutions \( u \in \mathcal{F}_{(s,T)} \) to the heat equation:

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

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

\[ u \in L^2((s, \tau) \rightarrow \text{Dom}(A_s)) \cap H^1((s, \tau) \rightarrow \mathcal{H}). \]

More precisely,

\( e^{-3Ls} \mathcal{E}_s(u_t) + 2 \int_s^\tau e^{-3Lt} |A_t u_t|^2 \, dm_t \, dt \leq e^{-3Ls} \cdot \mathcal{E}_s(u_s). \) (2.10)

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

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

Moreover, \( v_s \in \text{Dom}(A_s) \) for a.e. \( s \in (\sigma, t) \).
PROOF.

(i) In the case where \( u_s, \sigma \in \mathcal{F} \), this follows from the previous lemma and the fact that \( \text{Dom}(A_t) = \text{Dom}(A_t) \). In the general case \( u_s, \sigma \in \mathcal{H} \), by the very definition of the heat equation it follows that \( u_s, \sigma \in \mathcal{F} \) for a.e. \( \sigma \in (s, \tau) \). Applying the previous argument now with \( \sigma \) in the place of \( s \) yields that \( u_t \in \text{Dom}(A_t) \) 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 (2.6) states

\[
|\partial_t \Gamma_t(v)| \leq 2L \cdot \Gamma_t(v).
\]

Together with (2.5) this implies \( \partial_s \mathcal{E}_s(u_t) \mid_{s=t} \leq 3L \cdot \mathcal{E}_t(u_t) \). Therefore,

\[
e^{3Lt} \partial_t \left[ e^{-3Lt} \mathcal{E}_t(u_t) \right] \leq \partial_s \mathcal{E}_t(u_s) \mid_{s=t} = -2 \int |A_t u_t|^2 \, dm_t
\]

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 that will allow us to prove that \( v_s \in \text{Dom}(A_s) \) for a.e. \( s \in (\sigma, t) \). Therefore, we may conclude

\[
\partial_s \mathcal{E}_s(v_s) \geq 2 \int |A_s v_s|^2 \, ds - 3L \cdot \mathcal{E}_s(v_s) - 2 \int A_s v_s \cdot v_s \cdot \partial_s f_s \, ds
\]

\[
\geq -3L \cdot \mathcal{E}_s(v_s) - \frac{L}{2} \int v_s^2 \, ds
\]

and thus

\[
\partial_s \left[ \mathcal{E}_s(v_s) + \|v_s\|_{L^2(m_s)}^2 \right] \geq -3L \cdot \mathcal{E}_s(v_s) - \frac{L}{2} \int v_s^2 \, ds
\]

\[
+ 2 \int \left[ \Gamma_s(v_s) + v_s^2 \cdot v_s \cdot \partial_s f_s \right] \, ds - \int v_s^2 \, ds
\]

\[
\geq -3L \cdot \left[ \mathcal{E}_s(v_s) + \|v_s\|_{L^2(m_s)}^2 \right].
\]

\[ \square \]

**Remark 2.13.** For fixed \( s \) and a.e. \( \sigma \geq s \) the operator \( P_{\sigma,s} \) maps \( \mathcal{H} \) into \( \text{Dom}(\mathcal{E}) \), and then for a.e. \( t > \sigma \) the operator \( P_{t,\sigma} \) maps \( \text{Dom}(\mathcal{E}) \) into \( \text{Dom}(A_t) \). Thus by composition, for a.e. \( t > s \) the operator \( P_{t,s} \) maps \( \mathcal{H} \) into \( \text{Dom}(A_t) \).

A simple restatement of the assertions of Proposition 2.14 will yield that for all \( s \leq t \) and all \( h \in \mathcal{H} \)

- \( 0 \leq h \leq 1 \Rightarrow 0 \leq P_{t,s} h \leq 1 \),
- \( P_{t,s} 1 = 1 \) provided \( m_\sigma(X) < \infty \),
- \( (P_{t,s} h)^2 \leq P_{t,s} (h^2) \).

**Proposition 2.14.** The following holds true.

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

\[
u_s \geq 0 \text{ a.e. on } X \quad \implies \quad u_t \geq 0 \text{ a.e. on } X.
\]

More generally, for any \( M \geq 0 \)

\[
u_s \leq M \text{ a.e. on } X \quad \implies \quad u_t \leq M \text{ a.e. on } X.
\]
If $m_\phi(X) < \infty$ 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 $s < t$

$$v_t \geq 0 \text{ a.e. on } X \implies v_s \geq 0 \text{ a.e. on } X.$$ 

More generally, for any $M \geq 0$

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

If $m_\phi(X) < \infty$, 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 $p \in [1, \infty]$

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

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

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

$$u_s^2 \leq g_s \text{ a.e. on } X \implies u_t^2 \leq g_t \text{ a.e. on } X.$$

**Proof.**

(i) Assume that $u$ solves the heat equation. Let $w = (u - M)_+$. Then for each $t$, strong locality of the Dirichlet form $E_t$ implies

$$E_t(u_t, (u_t - M)_+) = E_t((u_t - M)_+, (u_t - M)_+).$$

The chain rule applied to $\Phi(x) = (x)_+$ implies that a.e.0 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. $t$

$$0 \leq E_t((u_t - M)_+, (u_t - M)_+) = E_t(u_t, (u_t - M)_+)$$

$$= -\int \partial_t u_t \cdot (u_t - M)_+ e^{-\phi t} \, dm_\phi$$

$$= -\int \partial_t (u_t - M)_+ (u_t - M)_+ e^{-\phi t} \, dm_\phi$$

$$\leq -\frac{1}{2} e^{Lt} \cdot \partial_t \left[ e^{-Lt} \int_X (u_t - M)_+^2 \, dm_t \right],$$

where we used (2.5) in the last inequality. Thus $u_s \leq M$ will imply $u_t \leq M$ for all $t > s$.

In the case $m_\phi(X) < \infty$, 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. $s$

$$\frac{1}{2} \partial_s \int (v_s - e^{L(t-s)} M)_+^2 \, dm_s$$

$$= \int (v_s - e^{L(t-s)} M)_+ \partial_s (v_s - e^{L(t-s)} M)_+ \, dm_s$$

$$- \frac{1}{2} \int (v_s - e^{L(t-s)} M)_+^2 \partial_s f_s \, dm_s$$

$$= \int (v_s - e^{L(t-s)} M)_+ (\partial_s v_s + L e^{L(t-s)} M)_+ \, dm_s$$

$$- \frac{1}{2} \int (v_s - e^{L(t-s)} M)_+^2 \partial_s f_s \, dm_s$$

$$\geq -\frac{3}{2} L \int (v_s - e^{L(t-s)} M)_+^2 \, dm_s.$$

Applying Gronwall’s inequality yields

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

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

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

maps $L^1(m_s)$ boundedly into $L^1(m_t)$ and $L^\infty(m_s)$ boundedly into $L^\infty(m_t)$. 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. $t$

$$\frac{1}{2} e^{L t} \cdot \partial_t \left[ e^{-L t} \int_X w_t^2 \, dm_t \right] \leq \int \partial_t (u_t^2 - g_t) w_t \, dm_t$$

$$= \int \partial_t u_t (2 u_t w_t) \, dm_t - \int \partial_g g_t w_t \, dm_t =$$
where we applied the strong locality in the last equation. Thus
\[
\int w_t^2 \, dt \leq e^{L(t-s)} \int w_s^2 \, ds
\]
for all \( t > s \). This proves the claim. \( \Box \)

As a direct consequence we obtain the following corollary.

**Corollary 2.15.** For all \( s < t \)

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

(ii) \( \| P_{t,s} \|_{L^2(m_s) \to L^2(m_t)} \leq e^{L(t-s)} \), \( \| P_{t,s} \|_{L^\infty(m_s) \to L^\infty(m_t)} \leq e^{L(t-s)} \),

(iii) \( \| P_{t,s} \|_{L^4(m_s) \to L^4(m_t)} \leq e^{L(t-s)/2} \), \( \| P_{t,s} \|_{L^2(m_s) \to L^2(m_t)} \leq e^{L(t-s)/2} \).

The next result yields that the heat flow is a dynamic EVI\((-L/2, \infty)\)-flow for \( \frac{1}{2} \) times the Dirichlet energy \( \mathcal{E}_t \) on \( L^2(X, m_t) \). For the definition of dynamic EVI flows, we refer to Section 6.4.

**Theorem 2.16.**

(i) The heat flow is a dynamic forward EVI\((-L/2, \infty)\)-flow for \( \frac{1}{2} \times \) the Dirichlet energy on \( L^2(X, m_t) \), see the Appendix. More precisely, for all solutions \( (u_t)_{t \in [s, t]} \) to the heat equation, for all \( t \leq T \), and all \( w \in \text{Dom}(\mathcal{E}) \),

\[
\frac{1}{2} \partial^+ \| u_s - w \|_{L^2}^2 \big|_{s=t} + \frac{L}{4} \cdot \| u_t - w \|_t^2 \geq \frac{1}{2} \mathcal{E}_t(u_t) - \frac{1}{2} \mathcal{E}_t(w)
\]

where \( \| \cdot \|_{s,t} \) is defined according to Definition \([A.1]\) with

\[
d_t(v, w) = \| v - w \|_t = \left( \int |v - w|^2 \, dm_t \right)^{1/2}.
\]

(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 \([2.5]\) implies \( \partial_t \| v \|^2_t \leq L \| v \|^2_t \) as well as (following the argument from Proposition \([A.2]\))

\[
\partial_s \| v \|^2_s \big|_{s=t} \leq \frac{L}{2} \| v \|^2_t.
\]
for all $v$ and $t$. Therefore, we can estimate
\[
\frac{1}{2} \partial^+_s \| u_s - w \|_t^2 \leq \limsup_{s \to t} \frac{1}{2(s-t)} \left( \| u_s - w \|^2_t - \| u_t - w \|^2_t \right) + \limsup_{s \to t} \frac{1}{2(s-t)} \left( \| u_s - w \|^2_t - \| u_s - w \|^2_t \right) \\
\leq \langle u_t - w, \partial_t u_t \rangle_t + \frac{L}{4} \| u_t - w \|^2_t \\
= -\mathcal{E}_t(u, u) + \mathcal{E}_t(w, u) + \frac{L}{4} \| u_t - w \|^2_t \\
\leq -\frac{1}{2} \mathcal{E}_t(u, u) + \frac{1}{2} \mathcal{E}_t(w, w) + \frac{L}{4} \| u_t - w \|^2_t.
\]

(ii) Uniqueness and the growth estimate immediately follow from the EVI property. Indeed, the distance \( \| \cdot \|_t \) and the function \( \mathcal{E} \) on the time-dependent geodesic space \( L^2(X, m_t), t \in I \), satisfy all assumptions mentioned in the Appendix on EVI flows. In particular, the distance is log-Lipschitz: \( \| \partial_t \|_t^2 \leq L \| \|_t^2 \) 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 2.17.** Let \( u, g \in \text{Dom}(\mathcal{E}), 0 < r \leq t < T \). Then
\[
\lim_{s \nearrow t} P^*_{t,s} g = g \quad \text{in} \quad (\text{Dom}(\mathcal{E}), \sqrt{\mathcal{E}}), \quad \lim_{s \nearrow r} P^{*,r} u = u \quad \text{in} \quad (\text{Dom}(\mathcal{E}), \sqrt{\mathcal{E}}).
\]

**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 2.12(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 \( P^*_{t,s} g \to g \) in \( L^2(X) \),
\[
\mathcal{E}_t(g) \geq \limsup_{s \nearrow t} e^{-L(t-s)}(\mathcal{E}_s(P^*_{t,s} g) + \| P^*_{t,s} g \|_{L^2(m_s)}) - \| g \|_{L^2(m_t)}
\]
\[
\geq \limsup_{s \nearrow t} \mathcal{E}_s(P^*_{t,s} g) = \limsup_{s \nearrow t} \mathcal{E}_t(P^*_{t,s} g),
\]
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
\[ \limsup_{s \not= t} \mathcal{E}_t(P_{t,s}^* g - g) = \limsup_{s \not= t} (2\mathcal{E}_t(g) + 2\mathcal{E}_t(P_{t,s}^* g) - \mathcal{E}_t(g + P_{t,s}^* g)) \]
\[ \leq 4\mathcal{E}_t(g) - \liminf_{s \not= t} \mathcal{E}_t(g + P_{t,s}^* g) \leq 4\mathcal{E}_t(g) - \mathcal{E}_t(2g) = 0, \]
where the last inequality is a consequence of the lower semicontinuity of \( \mathcal{E}_t \).

The second assertion follows the same lines by replacing Theorem 2.12(iii) with Theorem 2.12(ii). \[ \Box \]

\section{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

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

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

\begin{lemma}
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,
\begin{equation}
|\mathcal{E}_t(u_t, v_t) - \mathcal{E}_s(u_s, v_s)| \leq C(u_s, v_t) \cdot |t - s|^{1/2}
\end{equation}
for a.e. \( s, t \in (\sigma, \tau) \) with \( s \leq t \) where
\begin{equation}
C(u_s, v_t) = C \cdot \left[ \mathcal{E}_s(u_s) + \mathcal{E}_t(v_t) + \|v_t\|_{L^2(m_t)}^2 \right]
\end{equation}
with \( C := Le^{3(L+1)T} \).
\end{lemma}

In other words, the commutator lemma states
\begin{equation}
\left| \int_X [A_t(P_{t,s} u_s) - P_{t,s}(A_s u_s)] v_t \, dm_t \right| \leq C(u_s, v_t) \cdot |t - s|^{1/2}.
\end{equation}

\begin{proof}
Obviously, the function \( r \mapsto \mathcal{E}_r(u_r, v_r) \) is finite (even locally bounded) and measurable on \( (\sigma, \tau) \). Therefore, by Lebesgue’s density theorem for a.e. \( s, t \in (\sigma, \tau) \)
\[ \mathcal{E}_t(u_t, v_t) = \lim_{\delta \searrow 0} \frac{1}{\delta} \int_{t-\delta}^t \mathcal{E}_r(u_r, v_r) \, dr, \]
\[ \mathcal{E}_s(u_s, v_s) = \lim_{\delta \searrow 0} \frac{1}{\delta} \int_{s}^{s+\delta} \mathcal{E}_r(u_r, v_r) \, dr, \]
and thus
\[ \mathcal{E}_t(u_t, v_t) - \mathcal{E}_s(u_s, v_s) = \lim_{\delta \searrow 0} \frac{1}{\delta} \int_s^{t-\delta} (\mathcal{E}_{r+\delta}(u_{r+\delta}, v_{r+\delta}) - \mathcal{E}_r(u_r, v_r)) \, dr. \]
\end{proof}
To proceed, we decompose the integrand into three terms:

\[
\frac{1}{\delta} [E_{r+\delta}(u_{r+\delta}, v_{r+\delta}) - E_r(u_r, v_r)]
\]

\[
= \frac{1}{\delta} [E_{r+\delta}(u_{r+\delta}, v_{r+\delta}) - E_{r+\delta}(u_r, v_{r+\delta})] + \frac{1}{\delta} [E_{r+\delta}(u_r, v_{r+\delta}) - E_r(u_r, v_r)] + \frac{1}{\delta} [E_r(u_r, v_{r+\delta}) - E_r(u_r, v_r)]
\]

\[=: \alpha_r(\delta) + \beta_r(\delta) + \gamma_r(\delta).
\]

Let us first estimate the second term:

\[
\beta_r(\delta) = \frac{1}{4\delta} \left[ E_{r+\delta}(u_r + v_{r+\delta}) + E_{r+\delta}(u_r - v_{r+\delta}) - E_r(u_r + v_{r+\delta}) - E_r(u_r - v_{r+\delta}) \right]
\]

\[ \leq \frac{3L}{4} e^{3L\delta} [E_r(u_r + v_{r+\delta}) + E_r(u_r - v_{r+\delta})]
\]

\[ \leq \frac{3L}{2} e^{6L\delta} [E_r(u_r) + E_{r+\delta}(v_{r+\delta})]
\]

due to the fact that \( |\partial_r E_r(w)| \leq 3L E_r(w) \) for each \( w \in \mathcal{F} \). According to Theorem 2.12, the final expressions can be estimated (uniformly in \( \delta \)) in terms of \( E_s(u_s) \) and \( E_t(v_t) \) \( + \|v_t\|^2_{L^2(m_t)} \). Thus we finally obtain

\[
\lim_{\delta \to 0} \int_{s}^{t-\delta} \beta_r(\delta) \, dr \leq \frac{3L}{2} \int_{s}^{t} [E_r(u_r) + E_r(v_r)] \, dr
\]

\[ \leq (t - s) \frac{3L}{2} e^{3L(t-s)} [E_s(u_s) + E_t(v_t) + \|v_t\|^2_{L^2(m_t)}].
\]

Now let us consider jointly the first and third terms:

\[
\int_{s}^{t-\delta} [\alpha_r(\delta) + \gamma_r(\delta)] \, dr
\]

\[ = \frac{1}{\delta} \int_{s}^{t-\delta} [E_{r+\delta}((u_{r+\delta} - u_r), v_{r+\delta}) + E_r(u_r, (v_{r+\delta} - v_r))] \, dr
\]

\[ = -\frac{1}{\delta} \int_{s}^{t-\delta} \int_{X} [(u_{r+\delta} - u_r) \cdot A_{r+\delta} v_{r+\delta} \cdot e^{-f_{r+\delta}}
\]

\[ + A_r u_r \cdot (v_{r+\delta} - v_r) \cdot e^{-f_r} \] \, dm_{\partial} \, dr
\]

\[ = -\frac{1}{\delta} \int_{0}^{\delta} \int_{s}^{t-\delta} \int_{X} [A_{r+\epsilon} u_{r+\epsilon} + A_r u_r \cdot (-A_{r+\epsilon} v_{r+\epsilon} + f_{r+\epsilon} v_{r+\epsilon}) \cdot e^{-f_{r+\epsilon}}] \, dm_{\partial} \, dr \, d\epsilon.
\]
Integrability of $|A_r u_r|^2$ w.r.t. $dm_r \, dr$ implies that $\int_{t-\delta}^{t} |A_r u_r|^2 \, dm_r \, dr \to 0$ as $\delta \to 0$ as well as $\int_{s}^{s+\delta} |A_r u_r|^2 \, dm_r \, dr \to 0$. Thus, together with Lipschitz continuity of $t \mapsto f_t$, this implies

$$\frac{1}{\delta} \int_{0}^{\delta} \int_{s}^{t-\delta} \int_{X} \left[ A_r u_r + \epsilon \cdot A_r \cdot v_r + \epsilon \cdot e^{-f_r} \right] \, dm_r \, dr \, d\epsilon \to 0$$

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

$$\lim_{\delta \to 0} \left| \int_{s}^{t-\delta} \left[ \alpha_r(\delta) + \gamma_r(\delta) \right] \, dr \right| \leq -\frac{1}{\delta} \int_{0}^{\delta} \int_{s}^{t-\delta} \int_{X} \left[ A_r u_r \cdot \dot{f}_r + \epsilon \cdot v_r + \epsilon \cdot e^{-f_r} \right] \, dm_r \, dr \, d\epsilon$$

$$\leq L \cdot |t - s|^{1/2} \cdot \left( \int_{s}^{t} |A_r u_r|^2 \, dm_r \, dr \right)^{1/2} \cdot \|v_t\|_{L^2(m_t)}$$

$$\leq L \cdot |t - s|^{1/2} \cdot \left( \frac{1}{2} e^{3L(t-s)} \mathcal{E}_s(u_s) \right)^{1/2} \cdot \|v_t\|_{L^2(m_t)}.$$

To summarize, we have

$$|\mathcal{E}_t(u_t, v_t) - \mathcal{E}_s(u_s, v_s)|$$

$$= \lim_{\delta \to 0} \left| \int_{s}^{t-\delta} \left[ \alpha_r(\delta) + \beta_r(\delta) + \gamma_r(\delta) \right] \, dr \right|$$

$$\leq |t - s| \frac{3L}{2} e^{3L(t-s)} \left[ \mathcal{E}_s(u_s) + \mathcal{E}_t(v_t) + \|v_t\|_{L^2(m_t)}^2 \right]$$

$$+ L \cdot |t - s|^{1/2} \cdot \left( \frac{1}{2} e^{3L(t-s)} \mathcal{E}_s(u_s) \right)^{1/2} \cdot \|v_t\|_{L^2(m_t)}$$

$$\leq C \cdot |t - s|^{1/2} \cdot \left[ \mathcal{E}_s(u_s) + \mathcal{E}_t(v_t) + \|v_t\|_{L^2(m_t)}^2 \right]$$

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

### 3 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 $(X, d_t, m_t)$, $t \in I$. 
3.1 The Setting

Here and for the rest of the paper, 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 $\sigma$-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

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

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, by choosing some reference measure $\nu_\otimes$ the measures can be represented as $m_t \propto e^{-f_t}d\nu_\otimes$ with functions $f_t$ satisfying $|f_t(x)| \leq C$, $|f_t(x) - f_t(y)| \leq C \cdot d_t(x, y)$, and

$$|f_s(x) - f_t(x)| \leq L \cdot |s - t|$$

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 $\text{CD}(K, N')$ in the sense of [4, 38, 50].

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 $X$:

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

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

The log-Lipschitz bound (3.1) implies that for all $s, t \in I$ and all $\mu, \nu \in \mathcal{P}$

$$\log \frac{W_t(\mu, \nu)}{W_s(\mu, \nu)} \leq L \cdot |t - s|$$

see corollary 2.2 in [51]. Note that the latter is equivalent to weak differentiability of $t \mapsto W_t(\mu, \nu)$ and $|\partial_t W_t(\mu, \nu)| \leq L \cdot W_t(\mu, \nu)$ for all $\mu, \nu \in \mathcal{P}$.

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

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

where the supremum is taken among all continuous and bounded functions $\varphi, \psi$. Closely related to this is the $d_t$-Hopf-Lax semigroup defined on bounded Lipschitz
functions \( \varphi \) by
\[
Q^t_a \varphi(x) := \inf_{y \in X} \left\{ \varphi(y) + \frac{1}{2a} \inf_{y \in X} \left( x, y \right)^2 \right\}, \quad a > 0, \ x \in X.
\]
The map \((a, x) \mapsto Q^t_a \varphi(x)\) satisfies the Hamilton-Jacobi equation
\[
\tag{3.4}
\partial_a Q^t_a \varphi(x) = -\frac{1}{2} (\text{lip}_t Q^t_a \varphi)^2(x), \quad \lim_{a \to 0} Q^t_a \varphi(x) = \varphi(x).
\]
In addition, since \((X, d_t)\) is assumed to be geodesic,
\[\text{Lip}(Q^t_a \varphi) \leq 2 \text{Lip}(\varphi), \quad \text{Lip}(Q^t_a f(x)) \leq 2[\text{Lip}(\varphi)]^2,\]
where \(\text{Lip} = \sup_x \text{lip}_t \varphi(x)\). See, for instance, [6, sec. 3] for these facts.

For \(\mu, \nu \in \mathcal{P}(X)\) the Kantorovich duality can be written as
\[
\tag{3.5}
\frac{1}{2} W^2_t(\mu_0, \mu_1) = \sup_{\varphi} \left\{ \int Q^t_a \varphi \ d\mu_1 - \int \varphi \ d\mu_0 \right\}.
\]
We say that a curve \(\mu : J \to \mathcal{P}(X)\) belongs to \(AC^p(J; \mathcal{P}(X))\) if
\[
W_t(\mu^a, \mu^b) \leq \int_a^b g(r) \ dr \quad \forall a < b \in J
\]
for some \(g \in L^p(J)\). We will exclusively treat the case \(p = 2\) and call \(\mu\) a 2-\emph{absolutely continuous curve}. Recall that there exists a minimal function \(g\), called \emph{metric speed} and denoted by \(|\dot{\mu}_t|_t\), such that
\[
|\dot{\mu}_t|_t := \lim_{b \to a} \frac{W_t(\mu^a, \mu^b)}{|b - a|}.
\]
See, for example, [3, theorem 1.1.2]. For continuous curves \(\mu \in C([0, 1], \mathcal{P}(X))\) satisfying \(\mu^a = u^a m\) with \(u^a \leq R\), \(\mu\) belongs to \(AC^2([0, 1], \mathcal{P}(X))\) if and only if for each \(t \in (0, T)\) there exists a velocity potential \((\Phi^a_t)_t\) such that
\[
\int_0^1 \int \Gamma_t(\Phi^a_t) d\mu^a \ da < \infty
\]
and
\[
\tag{3.6}
\int \varphi \ d\mu^a - \int \varphi \ d\mu^0 = \int_{a_0}^{a_1} \int \Gamma_t(\varphi, \Phi^a_t) d\mu^a \ da \quad \forall \varphi \in \text{Dom}(\mathcal{E}).
\]
Moreover, we can express the metric speed in the following way:
\[
\tag{3.7}
|\dot{\mu}_t|_t^2 = \int \Gamma_t(\Phi^a_t) d\mu^a.
\]
See sections 6 and 8 in [7] for a detailed discussion.

Occasionally, we have to measure the distance between points \(x, y \in X\) that belong to different time sheets. In this case, for \(s, t \in I\) and \(\mu, \nu \in \mathcal{P}(X)\), we define
\[ W_{s,t}(\mu, \nu) := \inf_{h \to 0} \lim_{t \to 0} \sup_{\mu_0 = \mu_1 = \nu} \left\{ \sum_{i=1}^{n} (d_i - d_{i-1})^{-1} W_{s+a_{i-1}(t-\tau)}^{2}(\mu^{a_{i-1}}, \mu^{a_i}) \right\}^{1/2} \]

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

\[ W_{s,t}(\mu, \nu) = \inf \left\{ \int_{0}^{1} |\dot{\mu}|_{W_{s+a(t-\tau)}}^{2} \, da \right\}^{1/2} \]

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

In the following we will make frequent use of the concept of regular curves, which has already been successfully used in \([6, 7, 17]\). We use the refined version of \([7]\).

**Definition 3.1.** 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 \( \rho \) is regular (w.r.t. \( m_t \)) if:

1. \( u \in C^1([0, 1], L^1(X)) \cap \text{Lip}([0, 1], \mathcal{F}^*) \),
2. there exists a constant \( R > 0 \) such that \( u^a \leq R \) m.a.e. \( \forall a \in [0, 1] \),
3. there exists a constant \( E > 0 \) such that \( E_{\xi}(\sqrt{u^a}) \leq E \forall a \in [0, 1] \).

**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. \( m_\theta \) where \( \theta \) is a function belonging to \( C^1([0, 1], \mathbb{R}) \). So we will just say regular curve.

We will use the following approximation result, which is a combination of \([7]\) lemma 12.2 and \([17]\) lemma 4.11. For this we define for a fixed time \( t \) the semi-group mollification \( h^t_\varepsilon \) given by

\[ h^t_\varepsilon \psi = \frac{1}{\varepsilon} \int_{0}^{\infty} H^t_\varepsilon \psi \kappa \left( \frac{a}{\varepsilon} \right) da, \]

where \((H^t_\varepsilon)_{a \geq 0}\) denotes the semi-group associated to the Dirichlet form \( E_t \), and \( \kappa \in C^\infty_0((0, \infty)) \) with \( \kappa \geq 0 \) and \( \int_{0}^{\infty} \kappa(a) da = 1 \). Recall that for \( \psi \in L^2(m_t) \cap L^\infty(m_t), h^t_\varepsilon \psi, \Delta_t(h^t_\varepsilon \psi) \in \text{Dom}(\Delta_t) \cap \text{Lip}_B(X) \). Moreover, \( \|h^t_\varepsilon \psi - \psi\| \to 0 \) in \( \text{Dom}(E) \) as \( \varepsilon \to 0 \) for \( \psi \in \text{Dom}(E) \).

**Lemma 3.2.** Let \( X \) be an \( RCD(K, \infty) \) 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 \((\rho^a_n), a \in [0, 1], n \in \mathbb{N} \), such that

\[ W_t(\rho^a_n, \rho^a) \to 0 \quad \forall a \in [0, 1], \]

\[ \limsup_{n \to \infty} \int_{0}^{1} |\rho^a_n|^2 \, da \leq W_t^2(\rho^0_0, \rho^1_0). \]
If we additionally impose that \( \rho^0, \rho^1 \in \text{Dom}(S) \), then
\[
S_t(\rho^a_n) \to S_t(\rho^a) \quad \forall a \in [0, 1],
\]
and
\[
\limsup_{n \to \infty} \sup_{a \in [0, 1]} S_t(\rho^a_n) \leq \sup_{a \in [0, 1]} S_t(\rho^a) = \max_{a \in [0, 1]} S_t(\rho^a).
\]

**Proof.** We follow the argument in [7, lemma 12.2] and approximate \( \rho^0, \rho^1 \) by two sequences of measures \( \{\sigma^0_n\}_n \) with bounded densities. Then as in [6, prop. 4.11] one employs a threefold regularization procedure to the \( W_t \)-geodesic \( (v^a_n)_n \) connecting \( a_n \) and \( a_n \) for some smooth kernel \( \chi \in C_c(\mathbb{R}) \). Finally, we set
\[
\rho^a_n, k = h^t_{1/k} \rho^a_{n, k, 2},
\]
where \( h^t_{1/k} \) is given by (3.9). Then by a standard diagonal argument one obtains a sequence of regular curves in the sense of Definition 3.1 satisfying (3.10) and (3.11).

In order to show (3.12) and (3.13), note that since \( X \) is an RCD\((K, \infty)\) space we have that \( a \mapsto S_t(\rho^a) \) is \( K \)-convex, where \( S_t(\rho^a) \) 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 argument in [17, lemma 4.11] and insert the explicit formulas of the regularization \( (\rho^a_n)_n \) to obtain
\[
\begin{align*}
S_t(\rho^a_n) &\leq S_t(\rho^a_{n, 2}) \\
&\leq \int_{\mathbb{R}} \chi_n(a') S_t(\rho^{a-a'}) da' \\
&\leq S_t(\rho^a) + \int_{\mathbb{R}} \chi_n(a') |S_t(\rho^{a-a'}) - S_t(\rho^a)| da'.
\end{align*}
\]
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 (3.12). One obtains (3.13) from (3.14) by exploiting the uniform continuity of the entropy along geodesics on compact intervals once more. \(\Box\)

In Section 4.2 we will see that there is an easier construction of regular curves based on the “dual heat flow” to be introduced next.

### 3.2 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 [3–6] applies. In particular, for each \( t \) there is a well-defined energy functional
\[
\mathcal{E}_t(u) = \int_X |\nabla_t u|^2 \, dm_t = \liminf_{v \to u \text{ in } L^2(X, m_t)} \int_X (\text{lip}_t v)^2 \, dm_t
\]
for \( u \in L^2(X, m_t) \) where \( \text{lip}_t u(x) \) 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 what follows, we freely switch between these two notations for the same object.

The Laplacian \( \Delta_t \) is defined as the generator of \( \mathcal{E}_t \), i.e., as the unique nonpositive self-adjoint operator on \( L^2(X, m_t) \) with domain \( \mathcal{D}(\Delta_t) \subseteq \mathcal{D}(\mathcal{E}_t) \) and
\[
- \int_X \Delta_t u v \, dm_t = \mathcal{E}_t(u, v) \quad \forall u \in \mathcal{D}(\Delta_t), \ v \in \mathcal{D}(\mathcal{E}_t).
\]
Thanks to the RCD\((K, \infty)\)-condition, for each \( t \) the domain of the Laplacian coincides with the domain of the Hessian \([20]\), i.e., \( \text{Dom}(\Delta_t) = W^{2,2}(X, d_t, m_t) \). Indeed, the self-improved Bochner inequality implies that
\[
\Gamma_{2,t}(u) \geq K |\nabla_t u|^2 + |\nabla_t^2 u|_{HS}^2.
\]
which after integrating w.r.t. \( m_t \), integrating by parts, and applying the Cauchy-Schwarz inequality gives
\[
(3.16) \quad \| \nabla_t^2 u \|^2 \leq (1 + K_-/2) \cdot (\| \Delta_t u \|^2 + \| u \|^2)
\]
with \( K_- := \max\{-K, 0\} \) and \( \| \cdot \|^2 := \| \cdot \|^2_{L^2(m_t)} \).

Note that in general, \( \text{Dom}(\Delta_t) \) may depend on \( t \); see Example \[2.7\].

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 set \( \mathcal{H} = L^2(X, m_0) \) and \( \mathcal{F} = \mathcal{D}(\mathcal{E}_o) \) as well as
\[
\mathcal{F}_{(s, \tau)} = L^2((s, \tau) \to \mathcal{F}) \cap H^1((s, \tau) \to \mathcal{F}^*) \subset C([s, \tau] \to \mathcal{H})
\]
for each \( 0 \leq s < \tau \leq T \). For the definition of the solution to the heat equation and for the existence of the heat propagator we refer to the previous section.

**Theorem 3.3.**

(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.,
\[
(3.17) \quad P_{t,s} h(x) = \int p_{t,s}(x, y) h(y) dm_s(y).
\]
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
\[
\int p_{t,s}(x, y) \, dm_s(y) = 1 \quad \forall s < t, \forall x
\]
and has the propagator property
\[
p_{t,r}(x, z) = \int p_{t,s}(x, y) p_{s,r}(y, z) \, dm_s(y) \quad \forall r < s < t, \forall s, z.
\]

\textbf{Proof.}

(i) It remains to verify the boundedness and regularity assumptions on \( f_t \) and \( \xi_t \) that were made for Theorem 2.2. Choose a reference point \( t_0 \in I \) and put \( \Gamma_\circ = \Gamma_{t_0} \). Then \( \mathcal{E}_\circ(u) = \int \Gamma_{t_0}(u) e^{-f_{t_0}} \, dm_\circ \). The uniform bounds on \( f_t \) and on \( \Gamma_\circ(f_t) \) are stated as assumption (3.2). The log-Lipschitz bound (3.1) on \( d_t \) implies the requested uniform bound on \( \Gamma_t \). The claim thus follows from Theorem 2.2.

(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 \( f_t \), \( \Gamma_t(\cdot) \), and \( \Gamma_t(f_t) \), this allows us to apply results of [36], which provide all the assertions of the theorem. \( \square \)

Remark 3.4. Formula (3.17) allows us to give a pointwise definition for \( P_{t,s} h(x) \) for each \( h \in L^2(X, m_\circ) \) (or, in other words, to select a “nice” version) and, moreover, it allows us 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 \((m_t, m_s, m_\circ)\) and that in general the operator norm in \( L^p \) for \( p \neq \infty \) will not be bounded by 1.

3.3 The Dual Heat Equation

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:

\[
\int u \, d(\hat{P}_{t,s} \mu) = \int (P_{t,s} u) \, d\mu \quad \forall u \in C_b(X), \forall \mu \in \mathcal{P}(X).
\]

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
\[
\partial_t v = \Delta_t v,
\]
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: $\int u \, d(\hat{\Delta}_s \mu) = \int \Delta_s u \, d\mu \, \forall u, \forall \mu$.

If we define Markov kernels $p_{t,s}(x, dy)$ for $s \leq t$ by

$$p_{t,s}(x, dy) = p_{t,s}(x, y) \, dm(x),$$

then

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

and the dual propagator is given by

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

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 3.5.**

(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 terms of the adjoint heat propagator

$$P_{t,s}^* g(y) = \int p_{t,s}(x, y) \, g(x) \, dm(x). \quad (3.20)$$

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 Lipschitz continuity of $t \mapsto f_t$ implies that all the regularity assumptions requested in [36] also hold for the time-dependent operators $\Delta_s - (\partial_s f_s)$ (which then are just the operators $\Delta_s$ perturbed by multiplication operators in terms of bounded functions). Thus all the previous results apply without any changes. \qed

**Corollary 3.6.** For all $g, h \in L^1(X)$

$$\int h \cdot P_{t,s}^* g \, dm = \int P_{t,s} h \cdot g \, dm \quad \text{and}$$

$$\hat{P}_{t,s}(g \cdot m_t) = (P_{t,s}^* g) \cdot m_s. \quad (3.21)$$

**Lemma 3.7.**

(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 \( W_r, r \in I \), see the next section). More precisely, there exists a constant \( C \) such that for all \( s, s' < t \), all \( \tau \), and all \( \mu \),

\[
W_\tau^2(\mu_s, \mu_{s'}) \leq C \cdot |s - s'|. 
\]

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

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

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

(iv) For \( \mu \in \mathcal{P}(X) \) such that \( \mu \in \text{Dom}(S) \), 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) Let \( \mu_s = \hat{P}_{t,s} \mu \) and \( \mu_{s'} = \hat{P}_{t,s'} \mu \) for \( s < s' < t \) be given. Then

\[
W_\tau^2(\mu_s, \mu_{s'}) \leq \iint d_\tau^2(x, y) p_{s',s}(x, y) \, dm_s(y) \, d\mu_{s'}(x).
\]

According to \([36][38]\), the heat kernel admits upper Gaussian estimates of the form

\[
p_{s',s}(x, y) \leq \frac{C}{m_\tau(B_\tau(\sqrt{\sigma}, x))} \cdot \exp\left(-\frac{d_\tau^2(x, y)}{C\sigma}\right)
\]

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

\[
A(R, x) \leq \left(\frac{R}{r}\right)^{N-1} \cdot e^{R\sqrt{|K|}/(N-1)} \cdot A(r, x)
\]

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

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

for \( R \geq \sqrt{\sigma} \). Hence, we finally obtain

\[
W_\tau^2(\mu_s, \mu_{s'}) \leq \iint d_\tau^2(x, y) p_{s',s}(x, y) \, dm_s(y) \, d\mu_{s'}(x) \leq \frac{C}{m_\tau(B_\tau(\sqrt{\sigma}, x))} \cdot \exp\left(-\frac{d_\tau^2(x, y)}{C\sigma}\right)
\]
\[ \leq \int_X \left[ \frac{C}{m_t(B_t(\sqrt{\sigma}, x))} \cdot \int_X d_t^2(x, y) \cdot \exp \left( -\frac{d_t^2(x, y)}{C\sigma} \right) dm_t(y) \right] d\mu_s'(x) \]
\[ \leq C\sigma + C \int_X \int_0^\infty R^2 \cdot \exp \left( -\frac{R^2}{C\sigma} \right) N \frac{R^{N-1}}{\sigma^{N/2}} \cdot e^{R\sqrt{K(N-1)}} dR d\mu_s'(x) \]
\[ \leq C' \cdot \sigma. \]

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

(iv) In a similar calculation as in Proposition 2.8, we find for \( \mu = \nu m_t, \mu_s = \tilde{P}_{t,s}\mu \), since the dual heat flow is mass preserving,
\[ \int_s^t \int \Gamma_t(\log v_r) d\mu_r \, dr = S_t(\mu) - S_s(\mu_s) - \int_s^t \int v_r \partial_t f_r \, dm_r \, dr \]
\[ \leq S_t(\mu) + m_t(X) + L(t - s). \]

Now choose \( \phi \in \text{Dom}(\mathcal{E}) \) with \( \phi, \Gamma(\phi) \in L^\infty(X) \). Then
\[ \left| \int \phi v_t \, dm_t - \int \phi v_s \, dm_s \right| \]
\[ = \left| \int_s^t \mathcal{E}_r(\phi, v_r) \, dr \right| \]
\[ \leq \int_s^t \left( \int \Gamma_r(\phi) v_r \, dm_r \right)^{1/2} \left( \int \Gamma_r(\log v_r) v_r \, dm_r \right)^{1/2} \, dr \]
\[ \leq \int_s^t \left( \int \Gamma_t(\phi) v_r \, dm_r \right)^{1/2} \left( e^{2L(s-t)} \int \Gamma_r(\log v_r) v_r \, dm_r \right)^{1/2} \, dr. \]

Then, theorem 7.3 in \[1\] yields
\[ |\tilde{\mu}_t|^2 \leq e^{2L(s-t)} \int \Gamma_r(\log v_r) v_r \, dm_r \in L^1_{\text{loc}}((0, t)), \]
where the last conclusion is due to our previous calculation.

\[ \square \]

**Lemma 3.8.** Let \( u, g \in \text{Dom}(\mathcal{E}) \) and \( t \in (0, T) \) with \( g \in L^1(X, m_t) \). Then
\[ \lim_{h \searrow 0} \frac{1}{h} \left( \int u g \, dm_t - \int u P_{t, t-h} g \, dm_{t-h} \right) = \int \Gamma_t(u, g) \, dm_t, \]
and for a.e. \( s < t \)
\[ \lim_{h \searrow 0} \frac{1}{h} \left( \int u P_{t,s+h}^* g \, dm_{s+h} - \int u P_{t,s}^* g \, dm_s \right) = \int \Gamma_s(u, P_{t,s}^* g) \, dm_s. \]
PROOF. Without loss of generality, assume that \( g \geq 0 \) and \( \int g \, dm = 1 \). The general case can be obtained by considering the positive and negative parts separately and normalization. We first prove that for \( g \in \text{Dom}(\mathcal{E}) \) and \( u \in \text{Lip}(X) \)

\[
(3.23) \quad \frac{1}{h} \left( \int u g \, dm - \int u P_{t,t-h}^* g \, dm \right) = \int_0^1 \int \Gamma_{t-rh}(u, P_{t,t-rh}^* g) \, dm \, dr.
\]

Note that for \( 0 \leq r_1 \leq r_2 \leq 1 \)

\[
\left| \int u P_{t,t-rh}^* g \, dm - \int u P_{t,t-rh_1}^* g \, dm \right| \leq \text{Lip}(u) W_2(\hat{P}_{t,t-rh}(gm_t), \hat{P}_{t,t-rh_1}(gm_t)),
\]

and hence, as a consequence of Lemma 3.7(ii), the map \( r \mapsto \int u P_{t,t-rh}^* g \, dm \) is absolutely continuous. Thus

\[
\frac{1}{h} \left( \int u g \, dm - \int u P_{t,t-h}^* g \, dm \right) = -\frac{1}{h} \int_0^1 \partial_r \int u P_{t,t-rh}^* g \, dm \, dr
\]

\[
= -\frac{1}{h} \int_0^1 \int u e^{-f_{t-rh}} \partial_r P_{t,t-rh}^* g \, dm \, dr
\]

\[
= -\frac{1}{h} \int_0^1 \int u P_{t,t-rh}^* g \partial_r e^{-f_{t-rh}} \, dm \, dr
\]

\[
= \int_0^1 \mathcal{E}_{t-rh}(P_{t,t-rh}^*, u e^{-f_{t-rh}}) \, dr
\]

\[
+ \int_0^1 P_{t,t-rh} g u e^{-f_{t-rh}} \partial_r f_{t-rh} \, dm \, dr
\]

\[
- \int_0^1 P_{t,t-rh} g u e^{-f_{t-rh}} \partial_r f_{t-rh} \, dm \, dr
\]

\[
= \int_0^1 \mathcal{E}_{t-rh}(P_{t,t-rh}^*, u e^{-f_{t-rh}}) \, dr = \int_0^1 \mathcal{E}_{t-rh}(P_{t,t-rh}^*, u) \, dr,
\]

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 (3.23) for every \( u \in \text{Dom}(\mathcal{E}) \) by approximating with Lipschitz functions \( u_n \), satisfying \( u_n \to u \) strongly in

\[
(\text{Dom}(\mathcal{E}), \sqrt{\| \cdot \|_{L^2(X)}^2 + \mathcal{E}(\cdot)}).
\]
see [5] prop. 4.10. Hence
\[
\lim_{h \to 0} \frac{1}{h} \left( \int u g \, dm_t - \int u P_{t,t-h}^* g \, dm_{t-h} \right) \\
= \lim_{h \to 0} \int_0^1 \Gamma_{t-rh} (u, P_{t,t-rh}^* g) \, dm_{t-rh} \, dr \\
= \int_0^1 \lim_{h \to 0} \Gamma_{t-rh} (u, P_{t,t-rh}^* g) \, dm_{t-rh} \, dr \\
= \int \Gamma_t (u, g) \, dm_t,
\]
where the third equality directly follows from Lemma 2.17 and the second equality follows from dominated convergence.

Similarly, for the second claim we write for \( h < t - s \)
\[
\frac{1}{h} \left( \int u P_{t,s+h}^* g \, dm_{s+h} - \int u P_{t,s}^* g \, dm_s \right) = \frac{1}{h} \int_s^{s+h} \partial_r \int u P_{t,r}^* g \, dm_r \, dr \\
= \frac{1}{h} \int_s^{s+h} \Gamma_r (u, P_{t,r}^* g) \, dm_r \, dr,
\]
which converges for a.e. \( s \) to \( \int \Gamma_s (u, P_{t,s}^* g) \, dm_s \) as \( h \searrow 0 \).

To summarize:
\begin{itemize}
\item 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 \( u_s = h \).
\end{itemize}

In Markov process theory, this is the Kolmogorov backward equation (in reverse time direction).

\begin{itemize}
\item By duality we obtain the dual propagator \( \hat{P}_{t,s} \) acting on probability measures. Given any \( \nu \in (P(X), W_t) \), the probability measures \( (s, y) \mapsto \mu_s = \hat{P}_{t,s} \nu \) solve the dual heat equation \( -\partial_s \mu_s = \hat{A}_s \mu_s \) in \( [0, t) \times X \) with terminal condition \( \mu_t = \nu \).
\end{itemize}

\begin{itemize}
\item 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 the adjoint heat equation.
\end{itemize}

### 4 Towards Transport Estimates

In the following, \( N \) will always denote an extended number in \( (0, \infty) \). The assumptions from Section 3.1 will always be in force (in particular, we assume \( \text{RCD}^*(K, N') \) and the bounds (3.1) and (3.2)). Moreover, \( X \) will be assumed to be bounded (and thus compact).
4.1 From Dynamic Convexity to Transport Estimates

**Definition 4.1.** 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 \(\mu^0, \mu^1 \in \text{Dom}(S)\)

\[
\frac{\partial^+ a}{\partial t} S_t (\mu^a) \bigg|_{a=1} - \frac{\partial^- a}{\partial t} S_t (\mu^a) \bigg|_{a=0} \geq \\
- \frac{1}{2} \frac{\partial^- a}{\partial t} W^2_{t-} (\mu^0, \mu^1) + \frac{1}{N} |S_t (\mu^0) - S_t (\mu^1)|^2.
\]

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

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

**Proposition 4.2.** Given probability measures \(\mu, v \in \mathcal{D} \subset \mathcal{P}\), then the \(W_t\)-geodesic \((\nu^t), a \in [0, 1]\), connecting \(\mu\) and \(v\) has uniformly bounded densities \(\frac{\partial \nu^t}{\partial m_t} \leq C\) and there exist \(W_t\)-Kantorovich potentials \(\phi\) from \(\mu\) to \(v\) and \(\psi\) from \(v\) to \(\mu\) (both conjugate to each other) such that

\[
\frac{\partial a}{\partial t} S_t (\eta^a) \bigg|_{a=0} \geq -\mathcal{E}_t (\phi, u), \quad \frac{\partial a}{\partial t} S_t (\eta^a) \bigg|_{a=1} \leq +\mathcal{E}_t (\psi, v).
\]

**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 [2]. Note that due to our (upper and lower) boundedness assumption on \(u, v\), no extra regularization is required.

**Proposition 4.3.** Given \(\tau \leq T\) and \(\mu, v \in \mathcal{D} \subset \mathcal{P}\), set \(\mu_t = \hat{P}_{t, \tau} \mu\) and \(v_t = \hat{P}_{tl, \tau} v\). For each \(t \in (0, \tau)\), let \(\phi_t\) and \(\psi_t\) be any conjugate \(W_t\)-Kantorovich potentials from \(\mu_t\) to \(v_t\) and vice versa. Then for every \(0 < r < t < s < \tau\)

\[
\frac{1}{2} \frac{\partial r}{\partial t} W^2_t (\mu_r, v_r) \bigg|_{r=t} \leq \mathcal{E}_t (\phi_t, u_t) + \mathcal{E}_t (\psi_t, v_t)
\]

and

\[
\frac{1}{2} \liminf_{\delta \to 0} \frac{1}{\delta} \int_r^s \left[ W^2_t (\mu_{t+\delta}, v_{t+\delta}) - W^2_t (\mu_t, v_t) \right] dt \geq \\
\int_r^s \mathcal{E}_t (\phi_t, u_t) + \mathcal{E}_t (\psi_t, v_t) dt.
\]

Here \(u_t\) and \(v_t\) denote the densities of \(\mu_t\) and \(v_t\), resp., w.r.t. \(m_t\).

**Proof.** We closely follow the argument in the proof of theorem 6.3 in [2]. According to Proposition 2.12 \(u_t, v_t \in \text{Dom}(\mathcal{E})\). Moreover, due to the boundedness of \(X\), the Kantorovich potentials \(\phi_t\) and \(\psi_t\) are Lipschitz and thus also lie in \(\text{Dom}(\mathcal{E})\). Since \(\phi_t\) and \(\psi_t\) are conjugate \(W_t\)-Kantorovich potentials from \(\mu_t\) to \(v_t\) and vice versa, we get

\[
\frac{1}{2} W^2_t (\mu_t, v_t) = \int \phi_t d\mu_t + \int \psi_t dv_t.
\]
whereas
\[ \frac{1}{2} W_t^2(\mu_r, v_r) \geq \int \phi_t d\mu_r + \int \psi_t dv_r \]
for \( r \neq t \). Thus with the help of Lemma [3.8] and Theorem [2.5(ii)]
\[ \frac{1}{2} \limsup_{r \to t} \frac{1}{t-r} \left[ W_t^2(\mu_t, v_t) - W_r^2(\mu_r, v_r) \right] \]
\[ \leq \limsup_{r \to t} \frac{1}{t-r} \left[ \int \phi_t [d\mu_t - d\mu_r] + \int \psi_t [dv_t - dv_r] \right] \]
\[ = \mathcal{E}_t(\phi_t, u_t) + \mathcal{E}_t(\psi_t, v_t). \]
This proves the first claim. With the same notation as we used before, note that
\[ \sup_t \mathcal{E}_t(\phi_t) < \infty \] as well as \( \sup_t \mathcal{E}_t(\psi_t) < \infty \) since each \((X, d_t)\) is bounded
(proposition 2.2 in [2]). We then find again by Lemma [3.8] and Fatou’s lemma that
\[ \frac{1}{2} \liminf_{s \to 0} \frac{1}{\delta} \int_r^s \left[ W_t^2(\mu_{t+\delta}, v_{t+\delta}) - W_r^2(\mu_t, v_t) \right] dt \]
\[ \geq \liminf_{s \to 0} \frac{1}{\delta} \int_r^s \left[ \int \phi_t [d\mu_{t+\delta} - d\mu_t] + \int \psi_t [dv_{t+\delta} - dv_t] \right] dt \]
\[ \geq \int_r^s \mathcal{E}_t(\phi_t, u_t) + \mathcal{E}_t(\psi_t, v_t) dt. \]

**Theorem 4.4.** Assume \((X, d_t, m_t), t \in (0, T)\), is a super-Ricci flow and \((\mu_t)\) and
\((v_t), t \leq \tau\), are dual heat flows started in probability measures \(\mu_t, v_t \in \mathcal{D}\). Then
for a.e. \( t \in (0, T) \)
\[ \partial_t W_t^2(\mu_t, v_t) \geq 0. \]

**Proof.** The assumptions on the densities are preserved by the dual heat flow;
that is, \(\mu_t\) and \(v_t\) will have densities in \(\text{Dom}(\mathcal{E})\) that are bounded from above
and bounded away from 0, uniformly in \(t\). Using the absolute continuity of \(t \mapsto W_t^2(\mu_t, v_t)\), we obtain for all \(r < s\)
\[ W_s^2(\mu_s, v_s) - W_r^2(\mu_r, v_r) \]
\[ \geq \limsup_{\delta \to 0} \int_r^s \frac{1}{\delta} \left[ W_t^2(\mu_{t+\delta}, v_{t+\delta}) - W_t^2(\mu_t, v_t) \right] \]
\[ + W_t^2(\mu_{t+\delta}, v_{t+\delta}) - W_t^2(\mu_t, v_t) \right] dt \]
\[ \geq \liminf_{\delta \to 0} \int_r^s \frac{1}{\delta} (W_t^2(\mu_{t+\delta}, v_{t+\delta}) - W_t^2(\mu_t, v_t)) dt \]
\[ + \liminf_{\delta \to 0} \frac{1}{\delta} \int_r^s (W_t^2(\mu_{t+\delta}, v_{t+\delta}) - W_t^2(\mu_t, v_t)) dt \geq \]
\[ \int_r^s 2(\mathcal{E}_t(u_t, \phi_t) + \mathcal{E}_t(v_t, \psi_t)) \, dt \]
\[ + \liminf_{\delta \to 0} \frac{1}{\delta} \int_r^s (W^2_t(\mu_t, v_t) - W^2_{t-\delta}(\mu_t, v_t)) \, dt \]
\[ \geq \int_r^s 2(\mathcal{E}_t(u_t, \phi_t) + \mathcal{E}_t(v_t, \psi_t)) \, dt \]
\[ - \int_r^s 2(\mathcal{E}_t(u_t, \phi_t) + \mathcal{E}_t(v_t, \psi_t)) \, dt \geq 0, \]

where we used Proposition 4.3 in the third inequality, while the fourth inequality is due to Proposition 4.2 and the definition of super-Ricci flow, i.e.,

\[ -\frac{1}{2} \partial_t W^2_t(\mu_t, v_t) \bigg|_{t=\tau} \leq \partial_a S(\eta_t^{1-}) - \partial_a S(\eta_t^{0+}) \]

for every \( W_t \)-geodesic \((\eta^b_t)_t \in [0, 1] \) connecting \( \mu_t \) and \( v_t \). In the previous argument, we used in the third and fourth inequality that \( \frac{1}{2} [W^2_{t+\delta} - W^2_t] \) is uniformly bounded, which is due to the log-Lipschitz bound on the distances. \( \square \)

**Corollary 4.5.** Assume that \((X, d_t, m_t), t \in (0, T)\), is a super-Ricci flow and that \((\mu_t)\) and \((v_t)\), \( t \leq \tau \), are dual heat flows started in points \( \mu_\tau \) and \( v_\tau \) in \( \mathcal{P} \), resp., for some \( \tau \in (0, T) \). Then for all \( 0 \leq s < t \leq \tau \)

\[ W_s(\mu_s, v_s) \leq W_t(\mu_t, v_t). \]  

**Proof.** For measures \( \mu_\tau, v_\tau \) with densities in \( \text{Dom}(\mathcal{E}) \) that are bounded from above and bounded away from 0, the estimate (4.4) immediately follows from the previous theorem and the fact that the map \( t \mapsto W_t(\mu_t, v_t) \) is absolutely continuous (Lemma 3.7).

The set of such probability measures is dense in \( \mathcal{P} \) (w.r.t. weak topology) and according to Lemma 3.7, \( \hat{P}_{t,s} \) is continuous on \( \mathcal{P} \). Thus the estimate (4.4) carries over to all \( \mu_\tau, v_\tau \in \mathcal{P} \). \( \square \)

**Theorem 4.6 ((I_N) \Rightarrow (II_N)).** Assume that \((X, d_t, m_t), t \in (0, T)\), is a super-N-Ricci flow and that probability measures \( \mu_\tau, v_\tau \) in \( \mathcal{P} \) are given for some \( \tau \in (0, T) \). Then the dual heat flows \( (\mu_t)_{t \leq \tau} \) and \( (v_t)_{t \leq \tau} \) starting in these points satisfy for all \( 0 \leq s < t \leq \tau \)

\[ W^2_s(\mu_s, v_s) \leq W^2_t(\mu_t, v_t) - \frac{2}{N} \int_s^t [S_r(\mu_r) - S_r(v_r)]^2 \, dr. \]  

**Proof.** For measures \( \mu_\tau, v_\tau \) within the subset \( \mathcal{D} \) we follow the proof of the previous theorem line by line and finally use the enforcement of the super-Ricci-flow property to deduce

\[ -\frac{1}{2} \liminf_{\delta \to 0} \frac{1}{\delta} [W^2_{t+\delta}(\mu_{t+\delta}, v_{t+\delta}) - W^2_t(\mu_{t+\delta}, v_{t+\delta})] \leq \partial_a S_t(\eta_t^{1-}) - \partial_a S_t(\eta_t^{0+}) - \frac{1}{N} [S_t(\mu_t) - S_t(v_t)]^2. \]
Together with the other estimates from the proof of the previous theorem, this gives

\[ W_s^2(\mu_s, v_s) - W_t^2(\mu_t, v_t) \leq -\frac{2}{N} \int_s^t [S_r(\mu_r) - S_r(v_r)]^2 \, dr. \]

For general \( \mu_t, v_t \in \mathcal{P} \) we apply the previous result to the pair \( \mu_t, v_t \in \mathcal{D} \) (cf. Lemma [3.7]), which already yields the claim for all \( 0 \leq s < t < \tau \). The claim for \( t = \tau \) now follows by approximation

\[ W_s^2(\mu_s, v_s) \leq W_t^2(\mu_t, v_t) - \frac{2}{N} \int_s^\tau [S_r(\mu_r) - S_r(v_r)]^2 \, dr \]

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 [3.7]. \( \square \)

### 4.2 From Gradient Estimates to Transport Estimates

**Theorem 4.7** \((\text{III}_N \Rightarrow \text{II}_N)\). Assume that \((X, d_t, m_t), t \in (0, T)\), satisfies the Bakry-Ledoux gradient estimate \((\text{III}_N)\) for the primal heat flow. Then the dual heat flow starting in arbitrary points \( \mu^0_t, \mu^1_t \in \mathcal{P}(X) \) satisfies for all \( 0 < s < \tau < T \)

\[ W_s^2(\mu_s^0, \mu_s^1) \leq W_t^2(\mu^0_t, \mu^1_t) - \frac{2}{N} \int_s^\tau [S_t(\mu^0_t) - S_t(\mu^1_t)]^2 \, dt. \quad (4.6) \]

**Proof.**

1. Given \( \tau \in I \) and a regular curve (see Section [3]) \((\mu^a_t)_{a \in [0,1]}\), define for each \( t \leq \tau \) the \( W_t\)-action

\[ A_t(\mu^a_t) = \sup \left\{ \sum_{i=1}^k \frac{1}{a_i - a_{i-1}} W_t^2(\mu^a_t^{a_i - 1}, \mu^a_t^{a_i}) : \right. \]

\[ \left. k \in \mathbb{N}, \ 0 = a_0 < a_1 < \cdots < a_k = 1 \right\} \]

of the curve \( a \mapsto \mu^a_t = \hat{\mu}_{t,t} \mu^a_t \). Let \( t \in (0, \tau) \) be given with \( A_t(\mu^a_t) < \infty \), in other words, such that the curve \( a \mapsto \mu^a_t \) is 2-absolutely continuous. (Obviously, this is true for \( t = \tau \). The subsequent discussion indeed will show that this holds for all \( t \leq \tau \).) Let \((\mu^0_t)_{a \in [0,1]}\) and \((\Phi^0_t)_{a \in [0,1]}\) denote the densities and velocity potentials for the curve \((\mu^a_t), a \in [0,1]\) (see [7] theorem 8.2), or (4.6)–(5.7) in the static space \((X, d_t, m_t)\). Then, in particular,

\[ A_t(\mu^a_t) = \int_0^1 |\dot{\mu}^a_t|_{W_t^1} \, da = \int_0^1 \int_X |\nabla_t \Phi^a_t|^2 \, d\mu^a_t \, da. \]

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

\[ W_s^2(\mu_s^0, \mu_s^1) \leq 2 \left[ \int_X \varphi_s^1 \, d\mu_s^1 - \int_X \varphi_s^0 \, d\mu_s^0 \right] + \epsilon(t - s). \]
and let $(\varphi_{s}^{a})_{a \in [0,1]}$ denote the Hopf-Lax interpolation of $\varphi_{s}^{0}, \varphi_{s}^{1}$ in the static space $(X, d_{s}, m_{s})$.

Then applying the continuity equation (3.6) and the Hamilton-Jacobi equation (3.4) yields

$$\epsilon + \frac{1}{t-s} [A_{t}(\mu_{t}) - W_{s}^{2}(\mu_{s}^{0}, \mu_{s}^{1})]$$

$$\geq \frac{1}{t-s} \int_{0}^{1} |\mu_{t}^{a}|^{2} da - \frac{2}{t-s} \left[ \int_{X} \varphi_{s}^{1} d\mu_{s}^{1} - \int_{X} \varphi_{s}^{0} d\mu_{s}^{0} \right]$$

$$= \frac{1}{t-s} \int_{0}^{1} \left[ \int_{X} |\nabla_{t} \Phi_{t}^{a}|^{2} d\mu_{t}^{a} - 2 \partial_{a} \int_{X} P_{t,s} \varphi_{s}^{a} d\mu_{t}^{a} \right] da$$

$$= \frac{1}{t-s} \int_{0}^{1} \int_{X} |\nabla_{t} \Phi_{t}^{a} - \nabla_{t} P_{t,s} \varphi_{s}^{a}|^{2} - |\nabla_{t} P_{t,s} \varphi_{s}^{a}|^{2} + P_{t,s} |\nabla_{s} \varphi_{s}^{a}|^{2} d\mu_{t}^{a} da$$

$$\geq \frac{1}{t-s} \int_{0}^{1} \int_{X} |\nabla_{t} \Phi_{t}^{a} - \nabla_{t} P_{t,s} \varphi_{s}^{a}|^{2} d\mu_{t}^{a} da$$

$$+ \frac{2}{N(t-s)} \int_{s}^{t} \int_{0}^{1} \int_{X} \left[ P_{t,s} \Delta_{r} P_{r,s} \varphi_{s}^{a} \right]^{2} d\mu_{t}^{a} da dr \geq 0$$

where for the second-to-last inequality we have used the Bakry-Ledoux gradient estimate (3.4).

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

$$W_{s}^{2}(\mu_{s}^{0}, \mu_{s}^{1}) \leq A_{t}(\mu_{t})$$

for any regular curve $(\mu_{t}^{a}), a \in [0,1]$. Given any $\mu_{t}^{0}, \mu_{t}^{1} \in \mathcal{P}(X)$ we can choose regular curves $(\mu_{t}^{a,n}), a \in [0,1]$, for $n \in \mathbb{N}$ such that $A_{t}(\mu_{t,n}) \rightarrow W_{t}^{2}(\mu_{t}^{0}, \mu_{t}^{1})$ and $W_{t}(\mu_{t}^{a,n}, \mu_{t}^{1}) \rightarrow 0$ as well as $W_{t}(\mu_{t}^{0,n}, \mu_{t}^{0}) \rightarrow 0$ for $n \rightarrow \infty$. According to Lemma 3.7, the latter also implies $W_{s}(\mu_{s,n}, \mu_{s}^{0}) \rightarrow 0$ as well as $W_{s}(\mu_{s,n}, \mu_{s}^{1}) \rightarrow 0$ for $n \rightarrow \infty$ where $\mu_{s,n}^{i} := \tilde{P}_{t,s} \mu_{t,n}^{i}$. Together with the previous estimate (applied with $t = \tau$ to the regular curves $(\mu_{\tau,n}^{a}), a \in [0,1]$), we obtain

$$W_{s}^{2}(\mu_{s}^{0}, \mu_{s}^{1}) = \lim_{n \rightarrow \infty} W_{s}^{2}(\mu_{s,n}^{0}, \mu_{s,n}^{1}) \leq \lim_{n \rightarrow \infty} A_{t}(\mu_{t,n}) = W_{t}^{2}(\mu_{t}^{0}, \mu_{t}^{1}).$$

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 $(a_{i}), i = 1, \ldots, k$, yields

$$A_{s}^{i}(\mu_{s}^{i}) \leq A_{\tau}(\mu_{\tau}^{i})$$

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

(2) Moreover, the previous estimates for given $s, t, \epsilon$ can be tightened up by choosing $k \in \mathbb{N}$ and $(a_{i}), i = 1, \ldots, k$, as well as for $i = 1, \ldots, k$ suitable bounded Lipschitz functions $-\varphi_{s}^{0,i}, \varphi_{s}^{1,i}$ that are in $W_{s}$-duality to each other and
which are almost maximizers of the dual representation of $W^2_s(\mu_s^{a_{i-1}}, \mu_s^{a_i})$ such that

$$
\epsilon + \frac{1}{t-s} \left[ A_t(\mu_t) - A_s(\mu_s) \right] \\
\geq \epsilon/2 + \frac{1}{t-s} \left[ A_t(\mu_t) - \sum_{i=1}^{k} \frac{1}{a_i - a_{i-1}} W^2_s(\mu_s^{a_{i-1}}, \mu_s^{a_i}) \right] \\
\geq \frac{1}{t-s} \int_0^1 \left[ \int_X |\nabla_\tau \Phi_\tau^a|^2 \, d\mu_t^a - 2 \partial_a \int_X P_{t,s} \phi_s^{a,k} \, d\mu_t^a \right] \, da \\
= \frac{1}{t-s} \int_0^1 \int_X \left[ |\nabla_\tau \Phi_\tau^a - \nabla_\tau P_{t,s} \phi_s^{a,k}|^2 - |\nabla_\tau P_{t,s} \phi_s^{a,k}|^2 \right] \, d\mu_t^a \, da \\
+ P_{t,s} |\nabla_\tau \phi_s^{a,k}|^2 \, d\mu_t^a \, da \\
\geq \frac{1}{t-s} \int_0^1 \int_X \left[ \nabla_\tau \Phi_\tau^a - \nabla_\tau P_{t,s} \phi_s^{a,k} \right]^2 \, d\mu_t^a \, da \\
+ \frac{2}{N(t-s)} \int_s^t \int_X \left[ P_{t,s} \nabla_\tau \phi_s^{a,k} \right]^2 \, d\mu_t^a \, da \, dr =: \alpha.
$$

The function $\phi_s^{a,k}$ here is obtained for $a \in (a_{i-1}, a_i)$ by Hopf-Lax interpolation of the Lipschitz functions

$$
\phi_s^{a_{i-1}+k} := \frac{1}{a_i - a_{i-1}} \phi_s^{0,i} \quad \text{and} \quad \phi_s^{a_{i-1}-k} := \frac{1}{a_i - a_{i-1}} \phi_s^{1,i}.
$$

Now let us choose $t$ to be a Lebesgue density point of

$$
t \mapsto \int_0^1 E_t(P_{t,s} \phi_s^a, P_{t,s}^* u_t^a) \, da.
$$

Then for $s$ sufficiently close to $t$ the commutator lemma (applied to time points $r$ and $t$) implies that

$$
\left[ \frac{1}{(t-s)} \int_s^t \int_X P_{t,r} \nabla_\tau \phi_s^{a,k} \, d\mu_t^a \, da \, dr \right]^2 \\
\geq \left[ \frac{1}{(t-s)} \int_s^t \int_X \Delta_t P_{t,s} \phi_s^{a,k} \, d\mu_t^a \, da \, dr \right]^2 - \epsilon \cdot N/2.
$$

Let us also briefly remark that the densities $u_t^a$ of the measures $\mu_t^a$ 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 applying Young’s inequality $(a - b)^2 \geq \frac{\delta}{1 + \delta} a^2 - \delta b^2$ where $\delta = N/\epsilon$, we get

\[
(\alpha) = \frac{1}{t - s} \int_0^1 \int_X |\nabla_t \Phi_t^a - \nabla_t P_{t, s} \chi^{a,k}_s|^2 d\mu^a \, da
\]

\[
+ \frac{2}{N} \int_0^1 \int_X |\nabla_t P_{t, s} \chi^{a,k}_s \cdot \nabla_t \log u^a_t | d\mu^a \, da
\]

\[
\geq \frac{2}{N + \epsilon} \int_0^1 \int_X |\nabla_t \Phi_t^a \cdot \nabla_t \log u^a_t | d\mu^a \, da
\]

\[
+ \frac{1}{t - s} - \frac{2}{\epsilon} \int_0^1 |\nabla_t \log u^a_t|^2 d\mu^a \, da
\]

\[
\cdot \int_X |\nabla_t \Phi_t^a - \nabla_t P_{t, s} \chi^{a,k}_s|^2 d\mu^a \, da
\]

\[
\geq \frac{2}{N + \epsilon} \int_0^1 \int_X |\nabla_t \Phi_t^a \cdot \nabla_t \log u^a_t | d\mu^a \, da
\]

provided $s$ is sufficiently close to $t$. Finally, using the continuity equation for the curve $(\mu_t^a), a \in [0, t], (and its velocity potentials $\Phi_t^a$) we obtain

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

Passing to the limit $s \to t$ yields

\[
\epsilon + \partial_{t^-} A_t(\mu_t) \geq \frac{2}{N + \epsilon} |S_t(\mu_t^1) - S_t(\mu_t^0)|^2 - \epsilon
\]

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

\[
(4.7) \quad \partial_{t^-} A_t(\mu_t) \geq \frac{2}{N} |S_t(\mu_t^1) - S_t(\mu_t^0)|^2.
\]

Recall that this holds for a.e. $t \in (0, \tau)$. Moreover, note that $t \mapsto A_t(\mu_t)$ is absolutely continuous. Indeed, by Lemma 3.7 and the log-Lipschitz assumption (3.1)

\[
|W_{t+\epsilon}^2(\mu_{t+\epsilon}, \mu_t) - W_t^2(\mu_t)|
\]

\[
\leq |W_{t+\epsilon}^2(\mu_{t+\epsilon}, \mu_t) - W_t^2(\mu_t, \mu_t)|
\]

\[
+ |W_t^2(\mu_{t+\epsilon}, \mu_t) - W_t^2(\mu_t, \mu_t)| \leq
\]

\[
\frac{2}{N} |S_t(\mu_t^1) - S_t(\mu_t^0)|^2.
\]
\[
\begin{align*}
& \leq 2Le^{2\epsilon} W_t^2(\mu_t^a, \mu_t^b) + \frac{2\sqrt{\epsilon}}{1 - 2\sqrt{\epsilon}} W_t^2(\mu_t^a, \mu_t^b) \\
& + \frac{1}{\sqrt{\epsilon}} W_t^2(\mu_{t+\epsilon}^a, \mu_t^a) + \frac{1}{\sqrt{\epsilon}} W_t^2(\mu_{t+\epsilon}^b, \mu_t^b) \\
& \leq C_0 \sqrt{\epsilon} W_t^2(\mu_t^a, \mu_t^b) + C_1 \sqrt{\epsilon}.
\end{align*}
\]

Thus we may integrate (4.7) from any \( s \in (0, \tau) \) to \( \tau \) to obtain

\[(4.8) \quad A_s(\mu_s^a) \leq A_\tau(\mu_\tau^a) - \frac{2}{N} \int_s^\tau \left[ S_t(\mu_t^0) - S_t(\mu_t^1) \right]^2 dt.
\]

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

\[A_\sigma(\mu_\sigma^a) \to W_\tau^2(\mu_\tau^0, \mu_\tau^1)\]

as \( \sigma \searrow \tau \). Carrying out the previous estimates, finally resulting in (4.8), with \( (\mu_t^a) \), \( a \in [0, 1] \), in the place of \( (\mu_t^a) \), \( a \in [0, 1] \), yields

\[W_s^2(\mu_s^0, \mu_s^1) \leq A_s(\mu_s^a) \leq A_\tau(\mu_\tau^a) - \frac{2}{N} \int_s^\sigma \left[ S_t(\mu_t^0) - S_t(\mu_t^1) \right]^2 dt \]

\[\to W_\tau^2(\mu_\tau^0, \mu_\tau^1) - \frac{2}{N} \int_s^\tau \left[ S_t(\mu_t^0) - S_t(\mu_t^1) \right]^2 dt.
\]

This proves the claim. \( \square \)

**Lemma 4.8.** Assume (III) (with \( N = \infty \)) and let \( (\mu^a) \), \( a \in [0, 1] \), be an arbitrary \( W_t \)-geodesic in \( \mathcal{P}(X) \). Let \( \chi \) be a standard convolution kernel on \( \mathbb{R} \). Then for a.e. \( t < \tau \) and every \( \delta > 0 \) the measures

\[
\mu_t^{a, \delta} := \int_\mathbb{R} (\hat{P}_{t,t} \mu^{a+\delta \chi}) \chi(b) db = \hat{P}_{t,t} \left( \int_\mathbb{R} \mu^{a+\delta \chi} \chi(b) db \right)
\]

constitute a regular curve \( (\mu_t^{a, \delta}) \), \( a \in [0, 1] \) (in the sense of Definition 3.1). Here \( \delta(a) = 0 \) for \( a \in [0, \delta] \), \( \delta(a) = 1 \) for \( a \in [1 - \delta, 1] \), and \( \delta(a) = \frac{a - \delta}{1 - 2\delta} \) for \( a \in [\delta, 1 - \delta] \).

Choosing \( t_n \searrow \tau \) and \( \delta_n \searrow 0 \) yields a sequence of regular curves satisfying (3.10)–(3.13). In addition, for these approximations the endpoints are simply given by the dual heat flow

\[
\mu_{t_n}^{a, \delta} = \hat{P}_{t,t_n} \mu^a
\]

for \( a = 0 \) as well as \( a = 1 \) and for all \( n \).

**Proof.** The reparametrization by means of \( \delta \) 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 $\chi$ guarantees smooth dependence on $a$, i.e., (1) of Definition 3.1 follows from Lemma 3.7. 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 $u_r^a(x)$ for $(a,t,x) \in [0,1] \times (0,t] \times X$ for each $t < \tau$, i.e., (2) of Definition 3.1. Moreover, $u_r^a(x)$ is uniformly bounded away from 0. Thus (3) of Definition 3.1 is equivalent to a uniform bound for the energy $E_t(u^a)$.

Boundedness of $u_r^a$ for $r < \tau$ implies
\[
\int_0^1 \int_0^r \mathcal{E}_t(u_t^a) \, dt \, da \leq \frac{1}{2} \int_0^1 \|u_t^a\|_{L^2(m,r)}^2 \, da < \infty.
\]
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 $\chi$ thus turns the integrable function $a \mapsto \mathcal{E}_t(u_t^{\theta(a)})$ into a bounded function: $\int_{\mathbb{R}} \mathcal{E}_t(u_t^{\theta(a+b)}) \chi(b) \, db \leq C$. Since the energy $u \mapsto \mathcal{E}_t(u)$ is convex, Jensen’s inequality implies
\[
\mathcal{E}_t\left(\int_{\mathbb{R}} u_t^{\theta(a+b)} \chi(b) \, db\right) \leq \int_{\mathbb{R}} \mathcal{E}_t(u_t^{\theta(a+b)}) \chi(b) \, db \leq C.
\]

The action estimate (3.11) 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 reparametrization increases the action by a factor bounded by $\frac{1}{(1-\delta)^2}$.

The entropy estimates (3.12) and (3.13) follow as in the proof of Lemma 3.2.

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

In the next section, we will prove the implication $(\text{II}_N) \Rightarrow (\text{III}_N)$ by composing the results $(\text{II}_N) \Rightarrow (\text{IV}_N)$ and $(\text{IV}_N) \Rightarrow (\text{III}_N)$. Partly, these arguments are quite involved. (And actually, for the last one, we freely make use of the next theorem, Theorem 4.9).

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 [31] for a related result.

**Theorem 4.9** ($(\text{II}) \Leftrightarrow (\text{III})$). For fixed $0 < s < t < T$ the following are equivalent:

$(\text{II})_{t,s}$ For all $\mu, \nu \in \mathcal{P}$,
\[
W_s(\tilde{P}_{t,s} \mu, \tilde{P}_{t,s} \nu) \leq W_t(\mu, \nu).
\]

$(\text{III})_{t,s}$ For all $u \in \text{Dom}(E)$,
\[
\Gamma_t(P_{t,s} u) \leq P_{t,s}(\Gamma_s(u)) \quad \text{m.a.e. on } X.
\]
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 \( (y^a) \), \( a \in [0, 1] \), connecting \( x \) and \( y \), let \( \mu_t^a = \delta_{y^a} \) and \( \mu_t^b = \omega_t \mu_t^a \). The transport estimate \( W_s(\mu_t^a, \mu_t^b) \leq W_t(\mu_t^a, \mu_t^b) \) implies that

\[
|\dot{\mu}_s|_{W_s} \leq |\dot{\mu}_t|_{W_t} = |\dot{y}|_{d_t} = d_t(x, y).
\]

Thus following the argument for [5, theorem 6.4], we obtain

\[
|P_{t,s}u(x) - P_{t,s}u(y)| = \left| \int u \, d\omega_t \delta_x - \int u \, d\omega_t \delta_y \right| \\
\leq \int_0^1 (|\nabla_s u|^2 \, d\mu_t^a)^{1/2} \cdot |\dot{\mu}_s|_{W_s} \, da \\
\leq \int_0^1 (P_{t,s}|\nabla_s u|^2(y^a))^{1/2} \cdot |\dot{y}|_{d_t} \, da \\
\leq d_t(x, y) \cdot \sup \{ P_{t,s}|\nabla_s u|^2(z) : d_t(x, z) + d_t(z, y) = d_t(x, y) \}.
\]

The Hölder continuity of \( z \mapsto P_{t,s}|\nabla_s u|^2(z) \), therefore, allows us to conclude that \( (P_{t,s}|\nabla_s u|^2)^{1/2} \) is an upper gradient for \( P_{t,s}u \). This proves the claim for bounded Lipschitz functions. The extension to \( u \in \text{Dom}(\mathcal{E}) \) follows as in [5].

(III) \( t,s \) \( \Rightarrow \) (II) \( t,s \): See the previous theorem.

\[ \square \]

5 From Transport Estimates to Gradient Estimates
and the Bochner Inequality

As before, for the following 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, \( |\partial_t d_t(x, y)| \leq L \cdot d_t(x, y) \) 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 \( x \).

5.1 The Bochner Inequality

The Time-Derivative of the \( \Gamma \) -Operator

Definition 5.1. Let an interval \( J \subset I \) and \( u \in \mathcal{F}_J \) with \( \Gamma_r(u_r)(x) \leq C \) uniformly in \( (r, x) \in J \times X \) be given. Then we define \( \Gamma_r(u_r)(x) \) as (one of the) weak subsequential limit(s) of

\[
(5.1) \quad \frac{1}{2\delta} [\Gamma_{r+\delta}(u_r) - \Gamma_{r-\delta}(u_r)](x)
\]
in $L^2(J \times X)$ for $\delta \to 0$. That is, for a suitable 0-sequence $(\delta_n)_n$ and all $g \in L^2(J \times X)$

$$
\frac{1}{2\delta_n} \int_J \int_X \left[ \Gamma_{r+\delta_n}(u_r) - \Gamma_{r-\delta_n}(u_r) \right] g_r \, dm_r \, dr \to \int_J \int_X \dot{\Gamma}_r(u_r) g_r \, dm_r \, dr
$$
as $n \to \infty$.

Actually, thanks to the Banach-Alaoglu theorem, such a weak limit always exists since (5.1)—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

$$
\liminf_{\delta \to 0} \frac{1}{2\delta} \int_J \int_X \left[ \Gamma_{r+\delta}(u_r) - \Gamma_{r-\delta}(u_r) \right] g_r \, dm_r \, dr
\leq \int_J \int_X \dot{\Gamma}_r(u_r) g_r \, dm_r \, dr
\leq \limsup_{\delta \to 0} \frac{1}{2\delta} \int_J \int_X \left[ \Gamma_{r+\delta}(u_r) - \Gamma_{r-\delta}(u_r) \right] g_r \, dm_r \, dr.
$$

**Remark 5.2.** All the subsequent statements involving $\dot{\Gamma}_r(u_r)$ 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 1.7 is that each of the properties (I), (II), and (III) will imply (IV) for every choice of the weak subsequential limit $\dot{\Gamma}_r(u_r)$. Conversely, if (IV) is satisfied for some choice of the weak subsequential limit $\dot{\Gamma}_r(u_r)$, then it implies properties (I), (II), and (III). Indeed, the only property of $\dot{\Gamma}_r(u_r)$ that enters the calculations is (5.2).

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

$$
|\dot{\Gamma}_r(u_r)| \leq 2L \cdot \Gamma_r(u_r).
$$

**Lemma 5.3.** For every $u \in \mathcal{F}_J$ with $\sup_{r,x} \Gamma_r(u_r)(x) < \infty$ and every $g \in L^\infty(J \times X)$,

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

In particular,

$$
\liminf_{\delta \to 0} \frac{1}{\delta} \int_J \int_X \left[ \Gamma_{r+\delta}(u_{r+\delta}, u_r) - \Gamma_r(u_r, u_r) \right] g_r \, dm_r \, dr
\leq \int_J \int_X \dot{\Gamma}_r(u_r) g_r \, dm_r \, dr
\leq \limsup_{\delta \to 0} \frac{1}{\delta} \int_J \int_X \left[ \Gamma_{r+\delta}(u_{r+\delta}, u_r) - \Gamma_r(u_r, u_r) \right] g_r \, dm_r \, dr.
$$
Here for the second equality we used an index shift and Lusin’s theorem (to replace 
\( g_r \) \( d m_r \) again by \( g_r \) \( d m_r \)). The last equality follows from the log-Lipschitz 
continuity of \( r \mapsto \tilde{d}_r \), which allows us to estimate

\[
\frac{1}{\delta} \left| \int_J \int_X \left[ \Gamma_{r+\delta}(u_r) - \Gamma_r(u_r) \right] g_r \, d m_r \, d r \right|
\leq 2L \cdot \int_J \int_X \Gamma_r(u_r + \delta - u_r) g_r \, d m_r \, d r
\leq C' \cdot \int_J \tilde{E}_r(u_r + \delta - u_r) d r \to 0
\]
as \( \delta \to 0 \) since \( r \mapsto u_r \), as a map from \( J \) to \( F \), is “nearly continuous” (Lusin’s theorem).

\[\square\]

The Distributional \( \Gamma_2 \)-Operator

**Definition 5.4.** For \( r \in (0, T) \) and \( u \in \text{Dom}(\Delta_r) \) with \( |\nabla_r u| \in L^\infty \), we define the distribution-valued \( \Gamma_2 \)-operator as a **continuous linear operator**

\[ \Gamma_{2,r}(u) : F \cap L^\infty \to \mathbb{R} \]

by

(5.4) \[ \Gamma_{2,r}(u)(g) := \int \left[ -\frac{1}{2} \Gamma_r(\Gamma_r(u), g) + (\Delta_r u)^2 g + \Gamma_r(u, g) \Delta_r u \right] d m_r. \]
Note that
\[ |\Gamma_{2,r}(u)(g)| \leq 2 \|\nabla_r u\|_\infty \cdot \|\nabla_r^2 u\|_2 \cdot \|\nabla_r g\|_2 + \|g\|_\infty \cdot \|\Delta_r u\|_2^2 \\
+ \|\nabla_r u\|_\infty \cdot \|\nabla_r g\|_2 \cdot \|\Delta_r u\|_2^2 \]
thanks to the fact that \( \|\nabla_r^2 u\|_2^2 \leq (1 + K_\kappa) \cdot (\|\Delta_r u\|_2^2 + \|u\|_2^2) \); cf. (3.16).

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., \( u \in \text{Dom}(E_r) \) and \( |\nabla_r u|^2 \in \text{Dom}(\Delta_r) \)), then obviously
\[ \left( \frac{\partial}{\partial r} \right) \left( \frac{\partial}{\partial r} \right) E \left( \int \Delta_r u \cdot g \, dm_r \right) \]
for all \( g \) under consideration where as usual \( E \left( \int \Delta_r u \cdot g \, dm_r \right) \).

On the other hand, if \( g \in \text{Dom}(\Delta_r) \), then in (5.4) we may replace the term \(-\Gamma_r(\Gamma_r(u), g)\) by \( \Gamma_r(u)\Delta_r g \).

**The Bochner Inequality**

**Definition 5.5.**
(i) We say that \((X, d_t, m_t), t \in I, \) \( X \) 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 \( g_t \in L^\infty, u_s \in \text{Lip}(X) \) and for a.e. \( r \in (s, t) \)
\[ \Gamma_{2,r}(u)(g) \geq \frac{1}{2} \int \Gamma_r(u)(g) \, dm_r + \frac{1}{N} \left( \int \Delta_r u \cdot g \, dm_r \right)^2 \]
where \( u_r = P_{r,s} u_s \) and \( g_r = P_{t,r}^* g_t \); cf. (1.11).

(ii) We say that \((X, d_t, m_t), t \in I, \) \( X \) satisfies property (IV \( N \)) if it satisfies the dynamic Bochner inequality with parameter \( N \) as above and in addition the regularity assumption (1.7) is satisfied, i.e., \( u_r \in \text{Lip}(X) \) for all \( r \in (s, t) \) with \( \sup_{r, x} \text{lip}_r u_r(x) < \infty \).

Note that in the case \( N = \infty \) inequality (5.5) simply states that
\[ \Gamma_{2,r}(u) \geq \frac{1}{2} \Gamma_r(u) m_r \]
as an inequality between distributions, tested against nonnegative functions \( g_r \) as above.

**5.2 From Bochner Inequality to Gradient Estimates**

**Theorem 5.6 (IV \( N \) \( \Rightarrow \) III \( N \)).** Suppose that the mm-space \((X, d_t, m_t), t \in I, \) satisfies the dynamic Bochner inequality (5.5) and the regularity assumption from Definition 5.5 ii). Then for a.e. \( x \in X \)
\[ \Gamma_t(P_{t,s}u)(x) - P_{t,s} \Gamma_s(u)(x) \leq -\frac{2}{N} \int_s^t [P_{t,r} \Delta_r u_r(x)]^2 \, dr. \]
PROOF. Given \( s, t \in (0, T) \) as well as \( u \in \text{Lip}(X) \) and \( g \in \mathcal{F} \cap L^\infty \) with \( g \geq 0 \), let \( u_r = P_{r,s}u \) and \( g_r = P^*_{t,r}g \) 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
\[
(5.7) \quad h_\tau - h_\sigma \leq \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau} \left[ h_{r+\delta} - h_r \right] \, dr
\]
\[
\leq \limsup_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau} \int_X \Gamma_{r+\delta}(u_{r+\delta}) \, d(\mu_{r+\delta} - \mu_r) \, dr
\]
\[
+ \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau} \int_X g_r [\Gamma_{r+\delta}(u_{r+\delta} \cdot u_r) - \Gamma_r(u_{r+\delta} \cdot u_r)] \, dm_r \, dr
\]
\[
+ \limsup_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau} \int_X g_r [\Gamma_{r+\delta}(u_{r+\delta} \cdot u_r - u_r) - \Gamma_r(u_{r+\delta} - u_r \cdot u_r)] \, dm_r \, dr
\]
\[=: (I) + (II) + (III') + (III'').\]

Each of the four terms will be considered separately. Since \( r \mapsto \mu_r \) is a solution to the dual heat equation, we obtain
\[
(I) = \limsup_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau} \int_X \Gamma_{r+\delta}(u_{r+\delta}) \cdot \left( - \int_r^{r+\delta} \Delta_q q \, dq \right) \, d\mu_r \, dr
\]
\[
= - \liminf_{\delta \searrow 0} \int_\sigma^{\tau} \int_X \Gamma_r(u_r) \left( \frac{1}{\delta} \int_r^{r+\delta} \Delta_q g q \, dq \right) \, dm_q \, dr
\]
\[
= - \int_\sigma^{\tau} \int_X \Gamma_r(u_r) \cdot \Delta_q g_r \, d\mu_r \, dr
\]
due to 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 2.12) and the function \( r \mapsto \Gamma_r(u_r) \) is in \( L^\infty \) thanks to Definition 5.5(ii).
The second term can easily be estimated in terms of $\hat{G}_r$ according to Lemma 5.3.

\[(\text{II}) = \lim_{\delta \searrow 0} \frac{1}{\delta} \int_{\sigma}^{\tau-\delta} \int_X g_r[\Gamma_{r+\delta}(u_{r+\delta}, u_r) - \Gamma_r(u_{r+\delta}, u_r)] dm_r dr \leq \int_{\sigma}^{\tau} \int_X g_r \hat{G}_r(u_r) dm_r dr.\]

The term (III') is transformed as follows:

\[(\text{III}') = -\lim_{\delta \searrow 0} \frac{1}{\delta} \int_{\sigma}^{\tau-\delta} \int_X (\Gamma_{r+\delta}(g_r, u_{r+\delta}) + g_r \Delta_{r+\delta} u_{r+\delta}) \cdot \left( \int_r^{r+\delta} \Delta_q u_q dq \right) dm_r dr = -\lim_{\delta \searrow 0} \int_{\sigma+\delta}^{\tau} \int_X (\Gamma_r(g_{r-\delta}, u_r) + g_{r-\delta} \Delta_r u_r) \cdot \left( \frac{1}{\delta} \int_{r-\delta}^r \Delta_q u_q dq \right) dm_r dr = -\int_{\sigma}^{\tau} \int_X (\Gamma_r(g_r, u_r) + g_r \Delta_r u_r) \cdot \Delta_r u_r dm_r dr.\]

Here again we used Lebesgue’s density theorem (applied to $r \mapsto \Delta_r u_r$) and the near continuity of $r \mapsto g_r$ as a map from $(s, r)$ into $L^2(X, m)$ and as a 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:

\[(\text{III}'') = -\lim_{\delta \searrow 0} \frac{1}{\delta} \int_{\sigma}^{\tau-\delta} \int_X (\Gamma_r(g_r, u_r) + g_r \Delta_r u_r) \cdot \left( \int_r^{r+\delta} \Delta_q u_q dq \right) dm_r dr = -\int_{\sigma}^{\tau} \int_X (\Gamma_r(g_r, u_r) + g_r \Delta_r u_r) \cdot (\Delta_r u_r) dm_r dr.\]

Summarizing and then using (5.5), we therefore obtain

\[h_\tau - h_\sigma = (\text{I}) + (\text{II}) + (\text{III}') + (\text{III}'') \leq \int_{\sigma}^{\tau} \int_X \left[ -\Gamma_r(u_r) \cdot \Delta_r g_r + g_r \hat{G}_r(u_r) - 2(\Gamma_r(g_r, u_r) + g_r \Delta_r u_r) \Delta_r u_r \right] dm_r dr \leq \]

...
Thus

\( (5.8) \quad \int_X \Gamma_\tau (P_{\tau,\sigma} u) g \, d m_\tau - \int_X P_{\tau,\sigma} \Gamma_\sigma (u) g \, d m_\tau \leq \)

\[ - \frac{2}{N} \int_\sigma^\tau \left[ \int_X P_{\tau,\tau} \Delta_r u_r \, g \, d m_\tau \right]^2 \, d r. \]

(b) Recall that, given \( u \) and \( g \), this holds for a.e. \( \tau \) and a.e. \( \sigma \). 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 \( \sigma \) and \( \tau \) in \( (5.8) \) do not depend on \( g \). Next we may assume that \( \sigma, \tau \in [s, t] \) with \( \sigma < \tau \) is chosen such that \( (5.8) \) with \( N = \infty \) simultaneously holding for all \( u \) from a dense countable set \( C_1 \) in \( \text{Lip}(X) \). Approximating arbitrary \( u \in \text{Lip}(X) \) by \( u_n \in C_1 \) yields

\[ \int_X \Gamma_\tau (P_{\tau,\sigma} u) g \, d m_\tau - \int_X P_{\tau,\sigma} \Gamma_\sigma (u) g \, d m_\tau \leq \]

\[ \liminf_n \int_X \Gamma_\tau (P_{\tau,\sigma} u_n) g \, d m_\tau - \lim_n \int_X P_{\tau,\sigma} \Gamma_\sigma (u_n) g \, d m_\tau \leq 0 \]

due to lower semicontinuity of the weighted energy on \( L^2 \). In other words, we have derived the gradient estimate \( (\text{III}) \) for almost all times \( \sigma \) and \( \tau \).

Thanks to Theorem \( 4.9 \) this implies the transport estimate \( (\text{II}) \) 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 \( 4.9 \) it yields the gradient estimate for given \( s \) and \( t \) and thus our initial assumption \( (5.7) \) is satisfied for the choice \( \sigma = s \) and \( \tau = t \).

(c) Taking (b) into account, we deduce from the lower semicontinuity of \( \int_\sigma^\tau [P_{r,\tau} \Delta_r u_r g \, d m_\tau]^2 \, d r \) in \( \tau, \sigma \) and \( u, g \) w.r.t. \( L^2 \)-convergence that \( (5.8) \) (for given \( N \)) holds with the choices \( \sigma = s \) and \( \tau = t \). Finally, choosing sequences of \( g \)'s that approximate the Dirac distribution at a given \( x \in X \) then implies that for all \( u \in \text{Lip}(X) \)

\[ (5.9) \quad \Gamma_t (P_{t,\sigma} u) (x) - P_{t,\sigma} \Gamma_s (u) (x) \leq \frac{2}{N} \int_s^t \left[ P_{t,\tau} \Delta_r u_r (x) \right]^2 \, d r \]

for a.e. \( x \in X \). This proves the claim for bounded Lipschitz functions. The extension to \( u \in \text{Dom}(\mathcal{E}) \) follows as in [5].
5.3 From Gradient Estimates to Bochner Inequality

In the previous section and the previous subsections of this section, we have proven the implications $(\text{III}_N) \Rightarrow (\text{II}_N)$ and $(\text{IV}_N) \Rightarrow (\text{III}_N)$. Taking the next subsection into account, where we show $(\text{II}_N) \Rightarrow (\text{IV}_N)$, we already have proven that $(\text{III}_N) \Rightarrow (\text{IV}_N)$. In the following, we will present another, more direct proof of this implication.

**Theorem 5.7** ($(\text{III}_N) \Rightarrow (\text{IV}_N)$). Suppose that the mm-space $(X, d_t, m_t)$, $t \in I$, satisfies the gradient estimate (5.6). Then the dynamic Bochner inequality (5.5) holds true as well as the regularity assumption from Definition 5.5(ii).

**Proof.** Assume that the gradient estimate $(\text{III}_N)$ holds true. It immediately implies the regularity assumption (1.7). To derive the dynamic Bochner inequality, let $s, t \in (0, T)$ as well as $u \in \text{Lip}(X)$ and $g \in \mathcal{F} \cap L^\infty$ with $g \geq 0$ be given. Let $u_r = P_{r,s} u$, $g_r = P^*_{t,r} g$ for $r \in [s, t]$ and as before consider the function

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

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

$$h_\tau - h_\sigma \leq \liminf_{\delta \downarrow 0} \frac{1}{\delta} \int_\sigma^{\tau-\delta} [h_{r+\delta} - h_r] dr$$

$$= \liminf_{\delta \downarrow 0} \frac{1}{\delta} \int_\sigma^{\tau-\delta} \int_X \left[ \Gamma_{r+\delta}(u_{r+\delta}) - P_{r+\delta,r} \Gamma_r(u_r) \right] g_{r+\delta} dm_{r+\delta} dr$$

$$\leq -\frac{2}{N} \limsup_{\delta \downarrow 0} \int_\sigma^{\tau-\delta} \int_X \left( P_{r+\delta,q} \Delta_q u_q \right)^2 dq g_{r+\delta} dm_{r+\delta} dr$$

$$\leq -\frac{2}{N} \int_\sigma^{\tau} \liminf_{\delta \downarrow 0} \left( \int_X \frac{1}{\delta} \int_r^{r+\delta} P_{r+\delta,q} \Delta_q u_q dq g_{r+\delta} dm_{r+\delta} dr \right)^2$$

$$= -\frac{2}{N} \int_\sigma^{\tau} \liminf_{\delta \downarrow 0} \left( \int_X \Delta_r u_r \ g_r \ dm_r \right)^2 dr$$

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

$$h_\tau - h_\sigma \geq \limsup_{\delta \downarrow 0} \frac{1}{\delta} \int_{\sigma-\delta}^{\tau} [h_{r+\delta} - h_r] dr \geq$$
\[
\begin{align*}
&\geq \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_{\tau}^{\infty} \int_{X} \Gamma_{r+\delta}(u_{r+\delta}) d(\mu_{r+\delta} - \mu_r) dr \\
&\quad + \limsup_{\delta \searrow 0} \frac{1}{\delta} \int_{\tau}^{\infty} \int_{X} g_{r} \left[ \Gamma_{r+\delta}(u_{r+\delta}, u_{r+\delta}) - \Gamma_{r}(u_{r+\delta}, u_{r}) \right] dm_{r} dr \\
&\quad + \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_{\tau}^{\infty} \int_{X} g_{r} \left[ \Gamma_{r+\delta}(u_{r+\delta}, u_{r+\delta}) - u_{r} \right] dm_{r} dr \\
&\quad + \Gamma_{r}(u_{r+\delta} - u_{r}, u_{r})] dm_{r} dr \\
&=: (I) + (II) + (III') + (III'').
\end{align*}
\]

Each of the four terms can be treated as before, which then yields
\[
\begin{align*}
h_{\tau} - h_{\sigma} &\geq (I) + (II) + (III') + (III'') \\
&\geq \int_{\sigma}^{\tau} \int_{X} \left[ -\Gamma_{r}(u_{r}) \cdot \Delta_{r} g_{r} + g_{r} \hat{r}_{r}(u_{r}) \\
&\quad - 2(\Gamma_{r}(g_{r}, u_{r}) + g_{r} \Delta_{r} u_{r}) \Delta_{r} u_{r} \right] dm_{r} dr \\
&= \int_{\sigma}^{\tau} \left[ -2\Gamma_{2,r}(u_{r})(g_{r}) + \int \hat{r}_{r}(u_{r}) g_{r} m_{r} \right] dr.
\end{align*}
\]

Combining this with the previous upper estimate and varying \(\sigma\) and \(\tau\), we thus have proven the dynamic Bochner inequality
\[
2\Gamma_{2,r}(u_{r})(g_{r}) \geq \int \hat{r}_{r}(u_{r}) g_{r} m_{r} + \frac{2}{N} \left( \int \Delta_{r} u_{r} g_{r} m_{r} \right)^{2}
\]
for a.e. \(r \in (s, t)\). \(\square\)

### 5.4 From Transport Estimates to Bochner Inequality

**Theorem 5.8 ((II\(_N\)) \(\Rightarrow\) (IV\(_N\)))**. Suppose that the mm-space \((X, d_{t}, m_{t}), t \in I\), satisfies the transport estimate \((1.9) = (4.5)\). Then the dynamic Bochner inequality \((1.10) = (5.5)\) with parameter \(N\) holds true as well as the regularity assumption \((1.7)\).

**Proof of Regularity Assumption.** Thanks to Theorem [4.9] we already know that the transport estimate (II\(_N\)) implies the gradient estimate (III\(_N\)) in the case \(N = \infty\). This proves the requested regularity. \(\square\)

**Proof of Dynamic Bochner Inequality.** We follow the argument from [12] with significant modifications due to the 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 \(g_{t} \geq 0\), \(g_{t} \neq 0\), as well as \(u_{s} \in \text{Lip}(X)\) be given and fixed for the following. Without restriction \(\int g_{t} \, dm_{t} = 1\). For \(r \in (s, t)\), let \(u_{r} = P_{r,s} u_{s}\) and \(g_{r} = P_{t,r} g_{t}\). 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 \( E_\tau^g \) on \( L^2(X, g_\tau m_\tau) \) with domain \( \text{Dom}(E_\tau^g) := \text{Dom}(E) \) by

\[
E_\tau^g(u) := \int \Gamma_\tau(u) g_\tau \, dm_\tau \quad \text{for } u \in \text{Dom}(E).
\]

Associated with the closed bilinear form \( (E_\tau^g, \text{Dom}(E_\tau^g)) \) on \( L^2(X, g_\tau m_\tau) \), there is the self-adjoint operator \( \Delta_\tau^g \) and the semigroup \( (H_a^{\tau,g})_{a \geq 0} \), i.e., \( u_a = H_a^{\tau,g} u \) solves

\[
\partial_a u_a = \Delta_\tau^g u_a \quad \text{on } (0, \infty) \times X, \ u_0 = u,
\]

where \( \Delta_\tau^g u = \Delta_\tau u + \Gamma_\tau(\log g_\tau, u) \). For fixed \( \sigma \in (s, \tau) \), we define the path \( (g_\tau^{\sigma,a})_{a \geq 0} \) to be

\[
g_\tau^{\sigma,a} := g_\tau(1 + u_\sigma - H_a^{\tau,g} u_\sigma).
\]

Note that these are probability densities w.r.t. \( m_\tau \). 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 \( a, \sigma, \) and \( \tau \) 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 \( m \)-space at fixed time \( \tau \); thus \( \|H_a^{\tau,g} u_\sigma\|_\infty \leq \|u_\sigma\|_\infty \leq 1/4 \).

Now let us assume that the transport estimate (II_N) holds true and apply it to the probability measures \( g_\tau m_\tau \) and \( g_\tau^{\sigma,a} m_\tau \). Then for all \( s < \sigma < \tau < t \) and all \( a > 0 \),

\[
W^2(\hat{P}_{\tau,s}(g_\tau m_\tau), \hat{P}_{\tau,\sigma}(g_\tau^{\sigma,a} m_\tau))
\]

\[
\leq W^2(g_\tau m_\tau, g_\tau^{\sigma,a} m_\tau)
\]

\[
\leq -\frac{2}{N} \int_\sigma^\tau \left[ S_r(\hat{P}_{\tau,r}(g_\tau m_\tau)) - S_r(\hat{P}_{\tau,r}(g_\tau^{\sigma,a} m_\tau)) \right]^2 \, dr.
\]

Dividing by \( 2a^2 \) and passing to the limit \( a \searrow 0 \), the subsequent Lemmata 5.9, 5.10, and 5.11 allow us to estimate term by term. We thus obtain

\[
-\frac{1}{2} \int P_{\tau,\sigma}^*(\Gamma_\sigma(u_\sigma) g_\tau) \, dm_\tau + \int \Gamma_\tau(P_{\tau,\sigma} u_\sigma, u_\sigma) g_\tau \, dm_\tau
\]

\[
\leq \frac{1}{2(1 - 2\|u_\sigma\|_\infty)} \int \Gamma_\tau(u_\sigma) g_\tau \, dm_\tau
\]

\[
-\frac{1}{N} \int_\sigma^\tau \left[ \int \Gamma_\tau(P_{\tau,r}^*(\log P_{\tau,r}^{\sigma,g}) u_\sigma) g_\tau \, dm_\tau \right]^2 \, dr.
\]
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 5.9, Lemma 5.10, and Lemma 5.11 applied to $\eta u_s$ instead of $u_s$ gives us

$$\frac{-\eta^2}{2} \int P_{\tau,\sigma}(\Gamma_\sigma(u_\sigma)) g_\tau \, dm_\tau + \eta^2 \int \Gamma_\tau(P_{\tau,\sigma} u_\sigma, u_\sigma) g_\tau \, dm_\tau \leq \frac{-\eta^2}{2(1 - 2\eta\|u_\sigma\|_{\infty})} \int \Gamma_\tau(u_\sigma) g_\tau \, dm_\tau$$

$$- \frac{\eta^2}{N} \int_\sigma \left[ \int \Gamma_\tau(P_{\tau,e}(\log P_{\tau,e}^* g_\tau), u_\sigma) g_\tau \, dm_\tau \right]^2 \, dr.$$

By dividing by $\eta^2$ and letting $\eta \to 0$, this inequality becomes

$$- \frac{1}{2} \int P_{\tau,\sigma}(\Gamma_\sigma(u_\sigma)) g_\tau \, dm_\tau + \frac{1}{2} \int \Gamma_\tau(P_{\tau,\sigma} u_\sigma, u_\sigma) g_\tau \, dm_\tau \leq$$

$$\frac{1}{2} \int \Gamma_\tau(u_\sigma) g_\tau \, dm_\tau - \frac{1}{N} \int_\sigma \left[ \int \Gamma_\tau(P_{\tau,e}(\log P_{\tau,e}^* g_\tau), u_\sigma) g_\tau \, dm_\tau \right]^2 \, dr.$$

This can be reformulated as

$$\frac{1}{2} \int \Gamma_\tau(u_\tau) g_\tau \, dm_\tau - \frac{1}{2} \int \Gamma_\tau(u_\sigma) g_\sigma \, dm_\sigma - \frac{1}{2} \int \Gamma_\tau(u_\sigma) g_\tau \, dm_\tau$$

$$- \frac{1}{2} \int \Gamma_\tau(u_\tau) g_\tau \, dm_\tau + \frac{1}{2} \int \Gamma_\tau(u_\tau, u_\sigma) g_\tau \, dm_\tau$$

$$\leq -\frac{1}{N} \int_\sigma \left[ \int \Gamma_\tau(P_{\tau,e}(\log P_{\tau,e}^* g_\tau), u_\sigma) g_\tau \, dm_\tau \right]^2 \, dr.$$

Now let us try to follow the argument from the proof of Theorem 5.7 and consider again the function

$$h_r := \int g_r \Gamma_r(u_r) \, dm_r \quad \text{for} \ r \in (s,t).$$

Recall that we already know from Theorem 4.9 that the transport estimate $(\Pi_N)$ implies the gradient estimate $(\III)$ (without $N$). Thus for all $s < \sigma < \tau < t$

$$\limsup_{\delta \searrow 0} \frac{1}{\delta} \int_{\sigma - \delta}^{\tau} (h_{r+\delta} - h_r) \, dr \leq h_\tau - h_\sigma \leq \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau - \delta} (h_{r+\delta} - h_r) \, dr.$$

Arguing as in the proof of Theorem 5.7, we get

$$h_\tau - h_\sigma \geq \int_{\sigma}^{\tau} \left[ -2 \Gamma_{2,e}(u_r)(g_r) + \int \Gamma_r(u_r) g_r \, dm_r \right] \, dr.$$
On the other hand, applying the previous estimate (5.11) (with \( r + \delta, r, \text{and } q \) in the place of \( \tau, \sigma, \text{and } r \)) we obtain

\[
\liminf_{\delta \to 0} \frac{1}{\delta} \int_{\alpha}^{\tau - \sigma} \left[ -\frac{2}{N} \right. \\
+ \int \Gamma_{r+\delta} (u_{r+\delta} - u_r) g_{r+\delta} \, dm_{r+\delta} \right] \, dr.
\]

We estimate the squared term from below using Young’s inequality:

\[
\left[ \int \Gamma_{r+\delta} (P_{r+\delta,q} (\log P_{r+\delta,q}^* g_{r+\delta}), u_r) g_{r+\delta} \, dm_{r+\delta} \right]^2 \\
\geq \frac{1}{1 + \epsilon} \left[ \int \Gamma_r (P_{r,q} (\log g_q), u_r) g_r \, dm_r \right]^2 \\
- \frac{1}{\epsilon} \left[ \int \Gamma_{r+\delta} (P_{r+\delta,q} (\log P_{r+\delta,q}^* g_{r+\delta}), u_r) g_r \, dm_{r+\delta} \right. \\
\left. - \int \Gamma_r (P_{r,q} (\log g_q), u_r) g_r \, dm_r \right]^2.
\]

where \( \epsilon > 0 \) is arbitrary. Further estimating and using the log-Lipschitz continuity \( r \mapsto \Gamma_r \) yields

\[
\left[ \int \Gamma_{r+\delta} (P_{r+\delta,q} (\log P_{r+\delta,q}^* g_{r+\delta}), u_r) g_{r+\delta} \, dm_{r+\delta} \\
- \int \Gamma_r (P_{r,q} (\log g_q), u_r) g_r \, dm_r \right]^2 \\
\leq 2 \left[ \int \Gamma_{r+\delta} (P_{r+\delta,q} (\log g_q), u_r) g_{r+\delta} \, dm_{r+\delta} \\
- \int \Gamma_r (P_{r+\delta,q} (\log g_q), u_r) g_{r+\delta} \, dm_{r+\delta} \right. \\
\left. - \int \Gamma_r (P_{r,q} (\log g_q), u_r) g_{r+\delta} \, dm_{r+\delta} \right]^2 \\
+ 2 \left[ \int \Gamma_r (P_{r+\delta,q} (\log g_q), u_r) g_{r+\delta} \, dm_{r+\delta} \\
- \int \Gamma_r (P_{r,q} (\log g_q), u_r) g_{r+\delta} \, dm_{r+\delta} \right]^2 \\
+ 2 \left[ \int \Gamma_r (P_{r,q} (\log g_q), u_r) g_{r+\delta} \, dm_{r+\delta} \right. \\
\left. - \int \Gamma_r (P_{r,q} (\log g_q), u_r) g_r \, dm_r \right]^2.
\]
\[
\leq 16L^2\delta^2 \left[ \int \Gamma_{r+\delta}(P_{r+\delta,q}(\log g_q) - u_r) g_{r+\delta} \, dm_{r+\delta} 
+ \int \Gamma_{r+\delta}(P_{r+\delta,q}(\log g_q), u_r) g_{r+\delta} \, dm_{r+\delta} \right]^2 \\
+ 2 \left[ \int \Gamma_{r+\delta}(P_{r+\delta,q}(\log g_q) - P_{r,q}(\log g_q), u_r) g_{r+\delta} \, dm_{r+\delta} \right]^2 \\
+ 2 \left[ \int \Gamma_{r+\delta}(P_{r,q}(\log g_q), u_r) d(g_{r+\delta} - g_r) m_r \right]^2,
\]

which, after integration over \([r, r + \delta]\) and division by \(\delta > 0\), converges to 0 as \(\delta\) goes to 0. Indeed,

\[
\delta \int_r^{r+\delta} \left[ \int \Gamma_{r+\delta}(P_{r+\delta,q}(\log P_{r+\delta,q}^* g_{r+\delta}) - P_{r,q}(\log g_q), u_r) g_{r+\delta} \, dm_{r+\delta} \right]^2 \, dq 
\leq C\delta \left( \int_r^{r+\delta} \int \Gamma_q(\log g_q) dm_q \, dr \right) \mathcal{E}_r(u_r) \xrightarrow{\delta \to 0} 0,
\]

and by Lemma 2.17 and the Lebesgue differentiation theorem we have

\[
\frac{1}{\delta} \int_r^{r+\delta} \left[ \int \Gamma_{r+\delta}(P_{r+\delta,q}(\log g_q) - P_{r,q}(\log g_q), u_r) g_{r+\delta} \, dm_{r+\delta} \right]^2 \, dq \xrightarrow{\delta \to 0} 0,
\]

while

\[
\frac{1}{\delta} \int_r^{r+\delta} \left[ \int \Gamma_{r+\delta}(P_{r,q}(\log g_q), u_r) d(g_{r+\delta} - g_r) m_r \right]^2 \, dq \xrightarrow{\delta \to 0} 0.
\]

Thus, since \(\epsilon\) is arbitrary, from the Lebesgue differentiation theorem we get

\[
\liminf_{\delta \to 0} \frac{1}{\delta} \int_r^{r+\delta} \left[ \int \Gamma_{r+\delta}(P_{r+\delta,q}(\log P_{r+\delta,q}^* g_{r+\delta}) - P_{r,q}(\log g_q), u_r) g_{r+\delta} \, dm_{r+\delta} \right]^2 \, dr 
\geq \left[ \int \Gamma_r(\log g_q, u_r) g_r \, dm_r \right]^2 = \left[ \int \Delta_ru_r g_r \, dm_r \right]^2.
\]

Finally, with Corollary 2.15, the log-Lipschitz continuity of \(r \mapsto \Gamma_r\), Lemma 2.17, and the Lebesgue differentiation theorem applied to \(r \mapsto \Delta_r u_r\) (which is in \(L^2((s, t), \mathcal{H})\) thanks to Theorem 2.12),

\[
\limsup_{\delta \to 0} \frac{1}{\delta} \int_{\sigma}^{\tau-\delta} \int \Gamma_{r+\delta}(u_{r+\delta} - u_r) g_{r+\delta} \, dm_{r+\delta} \, dr 
\leq \limsup_{\delta \to 0} \frac{1}{\delta} \int_{\sigma}^{\tau-\delta} \|g_{r+\delta}\| \infty \int \Gamma_{r+\delta}(u_{r+\delta} - u_r, u_{r+\delta}) dm_{r+\delta} \, dr \leq
\]
\[
\leq \limsup_{\delta \to 0} \frac{1}{\delta} \int_{\delta}^{\tau} e^{L_{r+\delta-t}} \|g_t\|_\infty \left( \int \Gamma_{r+\delta}(u_{r+\delta} - u_r, u_{r+\delta}) \, dt \right) \\
= \limsup_{\delta \to 0} \frac{1}{\delta} \int_{\delta}^{\tau} e^{L_{r+\delta-t}} \|g_t\|_\infty \left( \int \Gamma_{r+\delta}(u_{r+\delta} - u_r, u_r) \, dt \right) \\
= \limsup_{\delta \to 0} \left( \int_{\delta}^{\tau} e^{L_{r+\delta-t}} \|g_t\|_\infty \int \Delta u_{r+\delta} \, dt \int \Delta u_{r+\delta} \, dt \right) \\
= \int_{\delta}^{\tau} e^{L_{r+\delta-t}} \|g_t\|_\infty \left( - \int (\Delta u_{r+\delta})^2 \, dt + \int (\Delta u_{r})^2 \, dt \right) = 0.
\]

Combining the previous estimates we get

\[h_\tau - h_\sigma \leq -\frac{2}{N} \int_{\sigma}^{\tau} \left( \int \Delta u_r g_r \, dt \right)^2 \, dr,
\]
and then

\[-\frac{2}{N} \int_{\sigma}^{\tau} \left( \int \Delta u_r g_r \, dt \right)^2 \, dr \geq \int_{\sigma}^{\tau} \left[ -2 \mathbf{R}_{2,\tau}(u_r) + \int \mathbf{\dot{R}}_\tau(u_r)g_r \, dt \right] \, dr,
\]

which proves the claim. \(\square\)

**Lemma 5.9.** For every \(s < \sigma \leq \tau < t,\)

\[
\liminf_{\alpha \to 0} \frac{W_{2,\alpha}^g(\tilde{P}_{\tau,\sigma}(g_{\tau\sigma} m_\tau), \tilde{P}_{\tau,\sigma}(g_{\tau\sigma} m_\tau))}{2a^2} \geq -\int \frac{1}{2} P_{\tau,\sigma}(u_{\sigma})g_\tau \, dt \, m_\tau + \int \Gamma_\tau(u_\tau, u_\sigma)g_\tau \, dt \, m_\tau.
\]

**Proof.** We denote by \(Q_\alpha^\sigma\) the Hopf-Lax semigroup with respect to the metric \(d_\sigma.\) Note that \(a Q_\alpha^\sigma(\phi) = Q_1^\sigma(a \phi),\) so the Kantorovich duality (3.5) can be written as

\[
\frac{W_{2,\alpha}^g(v_1, v_2)}{2a^2} = \frac{1}{a} \sup_{\phi} \left[ \int Q_\alpha^\sigma \phi \, dv_1 - \int \phi \, dv_2 \right].
\]
We deduce
\[
W^2_\sigma(\widehat{P}_{t,a}(g^\sigma_{t,a} m_\tau), \widehat{P}_{t,a}(g_{t m_\tau}))
\geq \int \frac{Q^\sigma_a u_\sigma P^*_{t,a}(g^\sigma_{t,a}) - u_\sigma P^*_{t,a} g_\tau}{a} d m_\sigma
\geq \int \frac{Q_a u_\sigma - u_\sigma}{a} P^*_{t,a}(g^\sigma_{t,a} - g_\tau) d m_\sigma + \int \frac{Q_a u_\sigma - u_\sigma}{a} P^*_{t,a} g_\tau d m_\sigma
+ \int u_\sigma P^*_{t,a}(g^\sigma_{t,a} - g_\tau) \frac{1}{a} d m_\sigma.
\]

Note that, since \(u_\sigma\) is a Lipschitz function, \(u_\sigma\) is a Lipschitz function as well.

Indeed, from the dual representation of the Kantorovich-Rubinstein distance \(W^1_\sigma\) with respect to the metric \(d_\sigma\), we deduce
\[
|u_\sigma(x) - u_\sigma(y)| = \left| \int u_\sigma(z) d\widehat{P}_{\sigma,z}(\delta_x)(z) - \int u_\sigma(z) d\widehat{P}_{\sigma,z}(\delta_y)(z) \right|
\leq \text{Lip}_\sigma(u_\sigma) W^1_\sigma(\widehat{P}_{\sigma,z}(\delta_x), \widehat{P}_{\sigma,z}(\delta_y))
\leq \text{Lip}_\sigma(u_\sigma) W^1_{\sigma,\delta} \text{Lip}_\sigma(u^\sigma) d_\sigma(x, y),
\]

where the last inequality is a consequence of Theorem 4.19.

Since \(0 \geq (Q^\sigma_a u_\sigma - u_\sigma)/a \geq -2 \text{Lip}(u_\sigma)^2\) and \(g^\sigma_{t,a} \rightarrow g_\tau\) in \(L^2(X)\), the first integral vanishes. For the second integral we use (3.4) and estimate by Fatou’s lemma
\[
\liminf_{\sigma \rightarrow 0} \int \frac{Q_a u_\sigma - u_\sigma}{a} P^*_{t,\sigma} g_\tau d m_\sigma \geq -\frac{1}{2} \int \text{Lip}_\sigma(u_\sigma)^2 P^*_{t,\sigma} g_\tau d m_\sigma.
\]

For the last integral an argument similar to Lemma 3.8 for \(H^\tau_{a,\sigma}\) (compare lemma 4.14 in [6]) yields
\[
\lim_{\sigma \rightarrow 0} \int \frac{P^*_{t,\sigma}(g^\sigma_{t,a} - g_\tau)}{a} d m_\sigma = \int \Gamma_t(P_{t,\sigma} u_\sigma, u_\sigma) g_\tau d m_\tau.
\]

Combining the last two estimates we obtain
\[
\liminf_{\sigma \rightarrow 0} \frac{W^2_{\sigma}(\widehat{P}_{t,a}(g^\sigma_{t,a} m_\tau), \widehat{P}_{t,a}(g_{t m_\tau}))}{2a^2}
\geq -\frac{1}{2} \int \text{Lip}_\sigma(u_\sigma)^2 P^*_{t,\sigma} g_\tau d m_\sigma + \int \Gamma_t(P_{t,\sigma} u_\sigma, u_\sigma) g_\tau d m_\tau
= -\frac{1}{2} \int \text{Lip}_\sigma(u_\sigma)^2 P^*_{t,\sigma} g_\tau d m_\sigma + \int \Gamma_t(P_{t,\sigma} u_\sigma, u_\sigma) g_\tau d m_\tau.
\]
where the last inequality follows from our static \(\text{RCD}(K,N')\) assumption, which implies the Poincaré inequality and doubling property for the static space \((X,d_\sigma,\mu)\), and the fact that \(u_\sigma\) is a Lipschitz function (cf. [14]). □

**Lemma 5.10.** For every \(s < \tau < t\),

\[
\limsup_{a \to 0} \frac{W_2^2(g_\tau^{a,m_\tau},g_\tau^{m_\tau})}{2a^2} \leq \frac{1}{2(1 - 2\|\psi\|_\infty)} \int \Gamma_\tau(u_\sigma) g_\tau \, d\mu.
\]

**Proof.** Let \((Q^\tau_a)_{a \geq 0}\) be the \(d_\tau\) Hopf-Lax semigroup and fix a bounded Lipschitz function \(\phi\). Note that

\[
\partial_a \int Q^\tau_a(\psi) g_\tau^{a,a} \, d\mu \\
\leq -\int \frac{1}{2} \text{lip}_\tau(Q^\tau_a(\psi))^2 g_\tau^{a,a} \, d\mu + \int \Gamma_\tau(Q^\tau_a(\psi), H_\tau^{\tau,g} u_\sigma) g_\tau \, d\mu \\
= \int \left[ -\frac{1}{2} \text{lip}_\tau(Q^\tau_a(\psi))^2 (1 + u_\sigma - H_\tau^{\tau,g} u_\sigma) + \Gamma_\tau(Q^\tau_a(\psi), H_\tau^{\tau,g} u_\sigma) \right] g_\tau \, d\mu,
\]

where the inequality follows from [3, lemma 4.3.4] and dominated convergence. Applying the Cauchy-Schwartz inequality and that \(\Gamma_\tau(\psi) \leq \text{lip}_\tau(\psi) \mu_\tau\)-a.e., we find

\[
\int \Gamma_\tau(Q^\tau_a(\psi), H_\tau^{\tau,g} u_\sigma) g_\tau \, d\mu \leq \sqrt{\mathcal{E}_g(Q^\tau_a(\psi) \mathcal{E}_g(H_\tau^{\tau,g} u_\sigma)}
\]

\[
\leq \int \text{lip}_\tau(Q^\tau_a(\psi))^2 g_\tau \, d\mu \mathcal{E}_g(H_\tau^{\tau,g} u_\sigma)
\]

Then, since \(1 + u_\sigma - H_\tau^{\tau,g} u_\sigma \geq 1 - 2\|u_\sigma\|_\infty\), we obtain using Young’s inequality

\[
\partial_a \int Q^\tau_a(\psi) g_\tau^{a,a} \, d\mu \leq \frac{1}{2(1 - 2\|u_\sigma\|_\infty)} \mathcal{E}_g(H_\tau^{\tau,g} u_\sigma) \leq \frac{1}{2(1 - 2\|u_\sigma\|_\infty)} \mathcal{E}_g(u_\sigma) = \frac{1}{2(1 - 2\|u_\sigma\|_\infty)} \int \Gamma_\tau(u_\sigma) g_\tau \, d\mu.
\]

Integrating over \([0,a]\),

\[
\int Q^\tau_a(\psi) g_\tau^{a,a} \, d\mu - \int \phi g_\tau \, d\mu \leq \frac{a}{2(1 - 2\|u_\sigma\|_\infty)} \int \Gamma_\tau(u_\sigma) g_\tau \, d\mu,
\]

and dividing by \(a > 0\) proves the claim since the Kantorovich duality can be written as

\[
\frac{W_2^2(\nu_1,\nu_2)}{2a^2} = \frac{1}{a} \sup_{\phi} \left[ \int Q^\tau_a(\phi) \, d\nu_1 - \int \phi \, d\nu_2 \right],
\]

and \(\phi\) was an arbitrary bounded Lipschitz function. □
LEMMA 5.11. For every $s < \sigma \leq \tau < t$,
\[
\liminf_{a \to 0} \int_s^\tau \left[ \frac{S_r(\hat{P}_{\tau,r}(g_{\tau,a}^m m_\tau)) - S_r(\hat{P}_{\tau,r}(g_{\tau,m} m_\tau))}{a} \right]^2 \, dr \geq \int_s^\tau \left[ \int \Gamma_r(P_{\tau,r}(\log g_r, u_0) g_{\tau} \, dm_\tau) \right]^2 \, dr.
\]

PROOF. With the same estimates as in [12] we have
\[
\left[ S_r(\hat{P}_{\tau,r}(g_{\tau,a}^m m_\tau)) - S_r(\hat{P}_{\tau,r}(g_{\tau,m} m_\tau)) \right]^2 \\
\geq \frac{1}{(1 + \delta)} \left[ \int (P_{\tau,r}^*(g_{\tau,a}^m) - g_{\tau}) \log g_{\tau} \, dm_r \right]^2 \\
- \frac{1}{\delta} \left[ \int \frac{(P_{\tau,r}^* g_{\tau,a}^m - g_{\tau})^2}{g_{\tau}^2} \, dm_r \right].
\]

Next we apply Jensen’s inequality to the convex function $\alpha : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R} \cup \{+\infty\}$ 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 2.15 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
\[
\int \alpha(P_{\tau,r}^* g_{\tau,a}^m - P_{\tau,r}^* g_{\tau} P_{\tau,r}^* g_{\tau}) \, dm_r \\
\leq \int \frac{\alpha((g_{\tau,a}^m(x) - g_{\tau}(x)) M_{\tau,r}(y), g_{\tau}(x) M_{\tau,r}(y))}{M_{\tau,r}(y)} \\
\cdot p_{\tau,r}(x, y) \, dm_\tau(x) \, dm_r(y)
\]
\[
= \int \alpha((g_{\tau,a}^m(x) - g_{\tau}(x)), g_{\tau}(x)) p_{\tau,r}(x, y) \, dm_\tau(x) \, dm_r(y)
\]
\[
= \int \alpha((g_{\tau,a}^m(x) - g_{\tau}(x)), g_{\tau}(x)) \, dm_\tau(x) = \int g_{\tau}(\psi_{\sigma} - H_{\lambda^*)^2 u_0})^2 \, dm_\tau,
\]
where we applied Jensen’s inequality in the second and third lines, and Fubini and the definition of $g_{\tau,a}^m$ in the last two lines. Dividing by $a$ and taking the lim sup we
end up with
\[
\limsup_{a \to 0} \frac{1}{a} \int \frac{(P^*_{\tau,r} g^a - P^*_{\tau,r} g)^2}{P^*_{\tau,r} g} \, dm_r \\
\leq \limsup_{a \to 0} \frac{1}{a} \int g_\tau (u_\sigma - H^g_{a} u_\sigma)^2 \, dm_\tau \\
\leq \limsup_{a \to 0} 2 \|u_\sigma\|_\infty \int \frac{g_\tau (H^g_{a} u_\sigma - u_\sigma)}{a} \, dm_\tau \\
= -2 \|u_\sigma\|_\infty \int g_\tau \Gamma_\tau (u_\sigma, 1) \, dm_\tau = 0.
\]

The first equality follows from the fact that \(\frac{1}{a}(H^g_{a} u_\sigma - u_\sigma) \to \Delta^g u_\sigma\) weakly in \(\mathcal{F}^*\) (cf. Lemma 3.8 and [6, lemma 4.14]).

Since \(\delta > 0\) is arbitrary it suffices to show
\[
\lim_{a \to 0} \frac{1}{a} \int P^*_{\tau,r} (g(H^g_{a} u_\sigma - u_\sigma)) \log P^*_{\tau,r} g \, dm_r = \int \Gamma_\tau (P^*_{\tau,r} (\log P^*_{\tau,r} g), u_\sigma) g \, dm_\tau.
\]
This, indeed, follows from the fact that \(P^*_{\tau,r} (\log P^*_{\tau,r} g) \in \mathcal{F} = \text{Dom}(\mathcal{E}_\tau) = \text{Dom}(\mathcal{E}^g_\tau)\) (thanks to uniform boundedness of \(P^*_{\tau,r} g\) from above and away from 0) and from the fact that \(\frac{1}{a}(H^g_{a} u_\sigma - u_\sigma) \to \Delta^g u_\sigma\) weakly in \(\mathcal{F}^*\) as \(a \searrow 0\); more precisely (cf. Lemma 3.8)
\[
\frac{1}{a} \int (H^g_{a} u_\sigma - u_\sigma) \phi g_\tau \, dm_\tau \to - \int \Gamma_\tau (u_\sigma, \phi) g_\tau \, dm_\tau
\]
for all \(\phi \in \mathcal{F}\) as \(a \searrow 0\).

6 From Gradient Estimates to Dynamic EVI

In this section we will prove that the dual heat flow is a dynamic backward EVI gradient flow assuming 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 \(u \in \text{Dom}(\mathcal{E})\) and \(0 < s < t < T\)

(6.1) \[\Gamma_t (P_{t,s} u) \leq P_{t,s} (\Gamma_\tau (u)) \quad \text{m.a.e. on } X.\]

For the notion of dynamic backward EVI\(^\pm\) gradient flow we refer to the Appendix.

As in the previous sections, the assumptions from Section 2 will always be in force; in particular, we assume the RCD\(^*\)(\(K, N\)'\()-condition for each static mm-space \((X, d_\tau, m_\tau)\) as well as boundedness and \(L\)-Lipschitz continuity (in \(t\)) for \(\log d_\tau(x, y)\) and (in \(t\) and \(x\)) for \(f_\tau(x)\).
6.1 Dynamic Kantorovich-Wasserstein Distances

For the subsequent discussions let us fix a pair \((s, t) \in I \times I\) and—if not stated otherwise—let \(\vartheta : [0, 1] \to \mathbb{R}\) denote the linear interpolation
\[
\vartheta(a) = (1 - a)s + ta
\]
starting in \(s\) and ending in \(t\).

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 6.1.** For \(s < t\) and a 2-absolutely continuous curve \((\mu^a), a \in [0, 1]\), we define the action
\[
A_{s,t}(\mu) = \lim_{h \to 0} \sup \left\{ \sum_{i=1}^{n} (a_i - a_{i-1})^{-1} W_2^2(\vartheta(a_{i-1}) (\mu^{a_{i-1}}, \mu^{a_i})) \right\}
\]
\[
0 = a_0 < \cdots < a_n = 1, \ a_i - a_{i-1} \leq h\}
\]
For two probability measures \(\mu, \nu \in \mathcal{P}(X)\), we define
\[
W_2^2(\mu, \nu) = \inf \left\{ A_{s,t}(\mu) \mid \mu \in AC^2([0, 1], \mathcal{P}(X)) \text{ with } \mu_0 = \mu, \mu_1 = \nu \right\}.
\]

**Lemma 6.2.** The following holds true.

(i) The action \(\mu \mapsto A_{s,t}(\mu)\) is lower-semicontinuous, i.e., if \(\mu^a \to \mu\) for every \(a \) as \(j \to \infty\), we have
\[
A_{s,t}(\mu) \leq \liminf_{j \to \infty} A_{s,t}(\mu_j).
\]

(ii) For every absolutely continuous curve \(\mu\),
\[
A_{s,t}(\mu) = \liminf_{h \to 0} \left\{ \sum_{i=1}^{n} (a_i - a_{i-1})^{-1} W_2^2(\vartheta(a_{i-1}) (\mu^{a_{i-1}}, \mu^{a_i})) \right\}
\]
\[
0 = a_0 < \cdots < a_n = 1, \ a_i - a_{i-1} \leq h\}
\]

**Proof.** Since \(\mu^a_j \to \mu^a\) for every \(a \in [0, 1]\) in the Wasserstein sense, we have for every partition \(0 = a_0 < \cdots < a_n = 1\)
\[
\sum_{i=1}^{n} (a_i - a_{i-1})^{-1} W_2^2(\vartheta(a_{i-1}) (\mu^{a_{i-1}}, \mu^{a_i})) =
\lim_{j \to \infty} \sum_{i=1}^{n} (a_i - a_{i-1})^{-1} W_2^2(\vartheta(a_{i-1}) (\mu^{a_{i-1}}, \mu^{a_i})).
\]
and hence
\[\sum_{i=1}^{n}(a_i - a_{i-1})^{-1} W_{\theta(a_{i-1})}^2(\mu^{a_{i-1}}, \mu^{a_i}) \leq \liminf_{j \to \infty} A_{s,t}(\mu_j).\]

Taking the supremum over each partition and letting \(h \to 0\) proves
\[A_{s,t}(\mu) \leq \liminf_{j \to \infty} A_{s,t}(\mu_j).\]

We prove the second assertion by contradiction. Assume that there exists a sequence \(h_j \to 0\) and a partition \(0 = a^0_0 < \cdots < a^n_j = 1\) such that
\[a^j_i - a^j_{i-1} \leq h \quad \text{and} \quad \lim_{j \to \infty} \sum_{i=1}^{n}(a^j_i - a^j_{i-1})^{-1} W_{\theta(a^j_{i-1})}^2(\mu^{a^j_{i-1}}, \mu^{a^j_i}) < A_{s,t}(\mu).\]

For every \(j \in \mathbb{N}\) we define the curve \((\mu^a_j), a \in [0, 1]\), by
\[\mu^a_j = \mu^{a_i}_{a^j_{i-1}, a^j_i} \quad \text{if} \quad a \in [a^j_{i-1}, a^j_i),\]
where \((\mu^{a_i}_{a^j_{i-1}, a^j_i})_{a \in [a^j_{i-1}, a^j_i]}\) denotes the \(\theta(a^j_{i-1})\)-geodesic connecting \(\mu^{a^j_{i-1}}\) and \(\mu^{a^j_i}\). Note that for every partition \(\{\bar{a}_i\}_{i=1}^{N}\) with \(\bar{a}_i - \bar{a}_{i-1} \ll h_j\)
\[\sum_{i=1}^{N}(\bar{a}_i - \bar{a}_{i-1})^{-1} W_{\theta(\bar{a}_{i-1})}^2(\mu^a_{\bar{a}_i}, \mu^a_{\bar{a}_{i-1}}) \leq e^{2Lh_j} \sum_{i=1}^{n}(a^j_i - a^j_{i-1})^{-1} W_{\theta(a^j_{i-1})}^2(\mu^{a^j_{i-1}}, \mu^{a^j_i}),\]
since for every \(a^j_{i-1} \leq \bar{a}_{k-1} < \bar{a}_k \leq a^j_i\)
\[W_{\theta(a^j_{i-1})}^2(\mu^a_{\bar{a}_k}, \mu^a_{\bar{a}_{k-1}}) \leq \frac{(\bar{a}_k - \bar{a}_{k-1})^2}{(a^j_i - a^j_{i-1})^2} W_{\theta(a^j_{i-1})}^2(\mu^{a^j_{i-1}}, \mu^{a^j_i}).\]

Hence
\[A_{s,t}(\mu_j) \leq e^{2Lh_j} \sum_{i=1}^{n}(a^j_i - a^j_{i-1})^{-1} W_{\theta(a^j_{i-1})}^2(\mu^{a^j_{i-1}}, \mu^{a^j_i}).\]

This is a contradiction since \(\mu^a_j \to \mu_a\) for every \(a\) and hence
\[\liminf_{j \to \infty} A_{s,t}(\mu_j) \geq A_{s,t}(\mu).\]
\[\square\]

**Proposition 6.3.** For \(s < t \in I\) and \(\mu^0, \mu^1 \in \mathcal{P}\) we have
\[(6.3) \quad W_{s,t}^2(\mu^0, \mu^1) = \inf \left\{ \int_0^1 |\dot{\mu}^a|^2 (s+t-s) d\alpha \right\}\]
where the infimum runs over all 2-absolutely continuous curves \((\mu^a), a \in [0, 1]\), in \(\mathcal{P}\) connecting \(\mu^0\) and \(\mu^1\).

**Proof.** Choose an arbitrary partition \(0 = a_0 < \cdots < a_n = 1\) 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 \((\mu^a)\) and the log-Lipschitz property (3.1), we deduce

\[
\sum_{i=1}^{n} (a_i - a_{i-1})^{-1} W^2_{\theta(a_{i-1})} (\mu^{a_{i-1}}, \mu^{a_i})
\leq \sum_{i=1}^{n} (a_i - a_{i-1})^{-1} \left( \int_{a_i}^{a_{i-1}} |\dot{\mu}^a|_{\theta(a)} da \right)^2
\leq \sum_{i=1}^{n} \int_{a_i}^{a_{i-1}} |\dot{\mu}^a|^2_{\theta(a_{i-1})} da
\leq e^{2Lh} \int_0^1 |\dot{\mu}^a|^2 da.
\]

Taking the supremum over all partitions and letting \(h \to 0\), we obtain

\[
A_{s,t}(\mu) \leq \int_0^1 |\dot{\mu}^a|^2_{\theta(a)} da,
\]

and consequently

\[
W^2_{s,t}(\mu_0, \mu_1) \leq \inf \left\{ \int_0^1 |\dot{\mu}^a|^2_{\theta(a)} da \right\}.
\]

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 \(A_{s,t}(\mu)\). For each \(h > 0\) we consider the partition \(0 = a_0 < \cdots < a_n = 1 < a_{n+1}\) with \(a_i = ih\) and \(nh \leq 1\). We define \(\mu_a^h\) to be the \(W^{\theta(a_{i-1})}\)-geodesic connecting \(\mu_{a_{i-1}}\) with \(\mu_{a_i}\) whenever \(a \in [a_{i-1}, a_i]\). Then we clearly have that \(\mu_a^h \in AC^2([0,1], \mathcal{P}(X))\) and since \(\mu\) is absolutely continuous, for each \(a \in [0,1]\), \(\mu_a^h \to \mu_a\) in \((\mathcal{P}(X), W)\).

Note that \(\dot{\mu}_a^h|_{\theta(a)}\) is a uniformly bounded function in \(L^2([0,1])\):

\[
\int_0^1 |\dot{\mu}_a^h|^2_{\theta(a)} da \leq e^{2Lh} \sum_{i=1}^{n+1} \int_{a_{i-1}}^{a_i} |\dot{\mu}_a^h|^2_{\theta(a_{i-1})} da
\leq e^{2Lh} \sum_{i=1}^{n+1} (a_i - a_{i-1})^{-1} W^2_{\theta(a_{i-1})}(\mu_{a_{i-1}}, \mu_{a_i}) < \infty,
\]

since \(\mu_a^h\) is a piecewise geodesic and \(A_{s,t}(\mu) < \infty\). 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 \(|\dot{\mu}_a^h|_{\theta(a)} \to A\) in \(L^2([0,1])\). Hence from the convergence of
\[ \mu_a^h \to \mu_a \text{ we get} \]

\[ W_{\vartheta(a)}(\mu_a, \mu_{a+h}) = \lim_{h \to 0} W_{\vartheta(a)}(\mu_a^h, \mu_{a+h}) \]

\[ \leq \liminf_{h \to 0} \int_a^{a+h} |\mu_b| \vartheta(a) db \]

\[ \leq \liminf_{h \to 0} e^{\delta(t-s)} \int_a^{a+h} |\mu_b| \vartheta(b) db = e^{\delta(t-s)} \int_a^{a+h} A(b) db, \]

and hence

\[ |\mu_a| \vartheta(a) \leq A(a) \quad \text{for a.e. } a \in [0, 1]. \]

Consequently,

\[ \int_0^1 |\mu_a| \vartheta(a) da \leq \int_0^1 A^2(a) da \]

\[ \leq \liminf_{h \to 0} \int_0^1 |\mu_a^h| \vartheta(a) da \]

\[ \leq \liminf_{h \to 0} e^{2Lh} \sum_{i=1}^{n+1} \int_{a_{i-1}}^{a_i} |\mu_a^h| \vartheta(a_{i-1}) da \]

\[ \leq \liminf_{h \to 0} e^{2Lh} \sum_{i=1}^{n+1} (a_i - a_{i-1})^{-1} W^2_{\vartheta(a_{i-1})}(\mu_{a_{i-1}}, \mu_{a_i}) \]

\[ \leq A_{s,t}(\mu), \]

which proves the claim. \[ \square \]

To conclude this section we define a dynamic “dual distance” inspired by the dual formulation of the Kantorovich distance. We introduce the function space \( HLS_\vartheta \) defined by

\[ HLS_\vartheta := \left\{ \varphi \in \text{Lip}_b([a_0, a_1] \times X) \mid \right. \]

\[ \partial_a \varphi_a \leq -\frac{1}{2} \Gamma_{\vartheta(a)}(\varphi_a)L^1 \times m \text{ a.e. in } (a_0, a_1) \times X \left\}. \right. \]

In particular, for all nonnegative \( \phi \in L^1(X) \) and \( \varphi \in HLS_\vartheta \),

\[ \int \phi \varphi_{a_1} dm - \int \phi \varphi_{a_0} dm \leq -\frac{1}{2} \int_{a_0}^{a_1} \int \phi \Gamma_{\vartheta(a)}(\varphi_a) dm da. \]

**DEFINITION 6.4.** Let \( s < t \) and let \( \vartheta : [a_0, a_1] \to [s, t] \) denote the linear interpolation. Define for two probability measures \( \mu_0, \mu_1 \)

\[ \tilde{W}^2_{\vartheta}(\mu_0, \mu_1) := 2 \sup_{\varphi} \left\{ \int \varphi_{a_1} d\mu_1 - \int \varphi_{a_0} d\mu_0 \right\}, \]
where the supremum runs over all maps $\varphi(a, x) = \varphi_a(x) \in HLS_\theta$.

Note that $\tilde{W}_\theta$ does not necessarily define a distance. It does not even have to be symmetric. The next lemma collects two essential properties of $\tilde{W}_\theta$.

**Lemma 6.5.** The following hold true:

1. $\tilde{W}_\theta$ is lower-semicontinuous with respect to the weak-* topology on $\mathcal{P}(X) \times \mathcal{P}(X)$.
2. For every $\mu_0, \mu_1$

$$W_s^2(\mu_0, \mu_1) \leq e^{2L|s-t|}(a_1 - a_0)\tilde{W}_\theta^2(\mu_0, \mu_1).$$

**Proof.** To show the first assertion, let $\mu_0, \mu_1 \in \mathcal{P}(X)$ and choose $\varphi \in HLS_\theta$ almost optimal, i.e.,

$$\frac{1}{2} \tilde{W}_\theta(\mu_0, \mu_1) \leq \int \varphi_{a_1} \ d\mu_1 - \int \varphi_{a_0} \ d\mu_0 - \varepsilon,$$

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 $\varphi_{a_1}$ and $\varphi_{a_0}$ belong to $C_b(X)$,

$$\frac{1}{2} \tilde{W}_\theta(\mu_0, \mu_1) \leq \lim_{n \to \infty} \left\{ \int \varphi_{a_1} \ d\mu_{a_1} - \int \varphi_{a_0} \ d\mu_{a_0} \right\} - \varepsilon$$

$$\leq \frac{1}{2} \liminf_{n \to \infty} \tilde{W}_\theta(\mu_0^n, \mu_1^n) - \varepsilon.$$

This proves, since $\varepsilon > 0$ was arbitrary, that $\tilde{W}_\theta$ 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 $\varphi \in \text{Lip}_b(X)$. As already mentioned above, the Hopf-Lax semigroup $\varphi_b := Q^s_b(\varphi)$ solves

$$\frac{d}{db} \varphi_b \leq -\frac{1}{2} \Gamma_s(\varphi_b)$$

$$\leq -\frac{1}{2} e^{-2L|s-t|} \Gamma_{(1-b)s+b t}(\varphi_b) \ 1 \times m \text{ a.e. in } (0, 1) \times X.$$

Set $\bar{\varphi}_a := e^{-2L|s-t|}(a_1 - a_0)^{-1} \varphi_{\gamma(a)}$, where $\gamma: [a_0, a_1] \to [0, 1]$ with $\gamma(a) = \frac{a-a_0}{a_1-a_0}$. Then $\bar{\varphi}$ solves

$$\frac{d}{da} \bar{\varphi}_a \leq -\frac{1}{2} \Gamma_{\theta(a)}(\bar{\varphi}_a) \text{ in } (a_0, a_1) \times X$$

and

$$e^{-2L|s-t|}(a_1 - a_0)^{-1} \left( \int \varphi_1 \ d\mu_1 - \int \varphi_0 \ d\mu_0 \right) = \int \varphi_{a_1} \ d\mu_1 - \int \varphi_{a_0} \ d\mu_0.$$
Hence

\[ e^{-2L|s-t|}(a_1 - a_0)^{-1} \left( \int \varphi_1 \, d\mu_1 - \int \varphi_0 \, d\mu_0 \right) \leq \frac{1}{2} \tilde{W}_{\theta}^2(\mu_0, \mu_1). \]

Taking the supremum among all \( \varphi \) the Kantorovich duality for the metric \( W_s \) implies

\[ W_s^2(\mu_0, \mu_1) \leq e^{2L|s-t|}(a_1 - a_0) \tilde{W}_{\theta}^2(\mu_0, \mu_1). \]

**Proposition 6.6.** Let \( \vartheta : [0, 1] \to [s, t] \) be the linear interpolation. Then we have \( \tilde{W}_{\theta} \leq W_{s,t} \).

**Proof.** Fix \( \varphi \in HJS_{\vartheta} \) and \( (\mu), a \in [0, 1] \), a 2- absolutely continuous curve. We subdivide \([0, 1]\) into \( l \) intervals \([(k - 1)/l, k/l]\) of length \( 1/l \). We get an approximation for \((\mu_a)_{|(k-1)/l}, k/l]\) by regular curves \((\rho^a_{n,k})\), \( a \in [(k - 1)/l, k/l] \), on each interval \([(k - 1)/l, k/l]\). Obviously, for each \( k, n \) the map

\[ \left[ \frac{k - 1}{l}, \frac{k}{l} \right] \ni a \mapsto \int \varphi_a \, d\rho^a_{n,k} \]

is absolutely continuous:

\[ \int \varphi_{a+h} \, d\rho_{a+h} - \int \varphi_a \, d\rho_a \leq \text{Lip}(\varphi_{a+h}) W(\rho_{a+h}, \rho_a) + \|\varphi_{a+h} - \varphi_a\|_\infty. \]

Let \( u^a_{k,n} \) be the density of the regular curve \( \rho^a_{k,n} \). Hence for fixed \( k, n \)

\[ \frac{d}{da} \int \varphi_a u^a_{k,n} \, dm \leq \int \varphi_a u^a_{k,n} \, dm - \frac{1}{2} \int u^a_{k,n} \Gamma_{\vartheta}(a)(\varphi_a) \, dm. \]

From Lemma 6.7 we deduce

\[ \int u^a_{k,n} \varphi_a \, dm \leq \frac{1}{2} |\rho^a_{k,n}|^2 \text{Lip}(a_{k-1/l})^2 + \frac{1}{2} \left( \text{Lip}(a_{k-1/l}) \varphi_a \right)^2 \, d\rho^a_{k,n}. \]

Adding these two inequalities, integrating over \([(k-1)/l, k/l]\), and noting that

\[ e^{-L \frac{t-s}{l} i} (\text{Lip}(a_{k-1/l}) \varphi_a) \right)^2 \leq \Gamma_{\vartheta}(a)(\varphi_a) \quad \text{a.e.}, \]

we obtain

\[ \int \varphi_{k/1} u^a_{k/1} \, dm - \int \varphi_{k-1/l} u^a_{k-1/l} \, dm \]

\[ \leq \frac{1}{2} \int_{k-1/l}^{k/l} |\rho^a_{k,n}|^2 \text{Lip}(a_{k-1/l}) \, da \]

\[ + \frac{1}{2} (1 - e^{-L \frac{t-s}{l} i}) \int_{k-1/l}^{k/l} (\text{Lip}(a_{k-1/l}) \varphi_a)^2 \, d\rho^a_{k,n} \, da \]

\[ \leq \frac{1}{2} \int_{k-1/l}^{k/l} |\rho^a_{k,n}|^2 \text{Lip}(a_{k-1/l}) \, da + \frac{C_1}{2l} (1 - e^{-L \frac{t-s}{l} i}). \]
Taking the limit \( n \to \infty \) (and taking the scaling into account) gives

\[
\int \varphi_{k/l} \, d\mu_{k/l} - \int \varphi_{k-1/l} \, d\mu_{k-1/l} \leq \frac{1}{2} \left( W^2_{\vartheta(k/l)}(\mu_{k-1/l}, \mu_{k/l}) + C_1(1 - e^{-L \frac{\vartheta_{k/l}}{l}}) \right).
\]

Summing over each partition and noting that the left-hand side is a telescoping sum yields

\[
\int \varphi_1 \, d\mu_1 - \int \varphi_0 \, d\mu_0 \leq \frac{1}{2} \sum_{k=1}^l \left( W^2_{\vartheta(k/l)}(\mu_{k-1/l}, \mu_{k/l}) + C_1(1 - e^{-L \frac{\vartheta_{k/l}}{l}}) \right).
\]

Letting \( l \to \infty \) we obtain the desired estimate. \( \square \)

**Corollary 6.7.** 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

\[
W_{s,t}(\mu_0, \mu_1) = \widetilde{W}_{\vartheta}(\mu_0, \mu_1).
\]

**Proof.** We know from Proposition 6.6 that \( W_{s,t}(\mu_0, \mu_1) \geq W_{\vartheta}(\mu_0, \mu_1) \). Hence it remains to prove the other inequality.

For this let \( (\varphi_a) \in HLS_{\vartheta} \) and \( (\mu_a) \) be an absolutely continuous curve connecting \( \mu_0 \) and \( \mu_1 \).

Consider the partition \( 0 = a_0 < \cdots < a_n = 1 \) with \( a_i - a_{i-1} \leq h \) for some \( h > 0 \). Set

\[
[a_{i-1}, a_i] \ni a \mapsto \vartheta_i(a) = \frac{a - a_i}{a_i - a_{i-1}} \vartheta(a_{i-1}) + \frac{a - a_{i-1}}{a_i - a_{i-1}} \vartheta(a_i)
\]

and \( \varphi^i_a = \varphi_a|_{[a_{i-1}, a_i]} \). Notice that \( (\varphi^i_a)_a \) is in \( HLS_{\vartheta_i} \). Hence

\[
\widetilde{W}^2_{\vartheta_i}(\mu_{a_{i-1}}, \mu_{a_i}) \leq 2\left\{ \int \varphi_a \, d\mu_a - \int \varphi_{a_{i-1}} \, d\mu_{a_{i-1}} \right\}.
\]

Then summing over the partitions and taking the scalings into account, we end up with

\[
\sum_{i=1}^n (a_i - a_{i-1})^{-1} W^2_{\vartheta(a_{i-1})}(\mu_{a_{i-1}}, \mu_{a_i}) \leq e^{2Lh|s-t|} \sum_{i=1}^n \widetilde{W}^2_{\vartheta_i}(\mu_{a_{i-1}}, \mu_{a_i}) \leq 2e^{2Lh|s-t|} \sum_{i=1}^n \left\{ \int \varphi_a \, d\mu_a - \int \varphi_{a_{i-1}} \, d\mu_{a_{i-1}} \right\} = 2e^{2Lh|s-t|} \left\{ \int \varphi_1 \, d\mu_1 - \int \varphi_0 \, d\mu_0 \right\}.
\]
where we made use of Lemma 6.5(ii) in the first inequality. Taking the supremum over all \( a \) we deduce

\[
\sum_{i=1}^{n} (a_i - a_{i-1})^{-1} W_{\alpha_i(a_{i-1})}^2 (\mu_{a_{i-1}}, \mu_{a_i}) \leq e^{2Lh|s-t|} \tilde{W}_\theta^2 (\mu_0, \mu_1).
\]

We conclude

\[
W_{s,t}^2 (\mu_0, \mu_1) \leq \tilde{W}_\theta^2 (\mu_0, \mu_1).
\]

from taking the supremum in (6.6) over the partition \( 0 = a_0 < \cdots < a_n = 1 \) with \( a_i - a_{i-1} < h \) and subsequently letting \( h \searrow 0 \).

\[\square\]

6.2 Action Estimates

Let us recall the following estimate about the oscillation of \( a \) from \([6, \text{lemma 4.12}]\). For fixed \( t > 0 \), let \( \rho^a \) be a 2- absolutely continuous curve in \( \mathcal{P} \) with \( \rho^a = u^a dm_t \) and \( u \in C^{1}((0,1), L^1(X,m_t)) \). Then for any Lipschitz function \( \varphi \) we have

\[
\left| \int u^a \varphi \, dm_t \right| \leq \frac{1}{2} \left| \rho^a \right|^2_t + \frac{1}{2} \int \Gamma_t (\varphi) d\rho^a.
\]

Actually, we have inequality (6.7) for each \( \varphi \in \text{Dom}(\mathcal{E}) \) since we assume that each \( (X,d,\mu) \) is a static \( \text{RCD}(K,\infty) \), implying that Lipschitz functions are dense in the domain of the quadratic form \( \mathcal{E} \) with respect to the norm \( \sqrt{\|\varphi\|^2 + \mathcal{E}(\varphi)} \) \([5, \text{prop. 4.10}]\).

Moreover, we will use the following result about difference quotients and concatenations of functions in \( \mathcal{F}_{(s,t)} \).

**Lemma 6.8.** Let \( 0 < s < T \).

1. Let \( u \in \mathcal{F}_{(s,t)} \). Then for almost every \( a \in (s,t) \)

\[
\frac{1}{h} (u_{a+h} - u_a) \to \partial_a u_a \quad \text{weak-* in } \mathcal{F}^*,
\]

i.e., for every \( v \in \mathcal{F} \) and for almost every \( a \in (s,t) \),

\[
\int \frac{1}{h} (u_{a+h} - u_a) v \, dm_\cdot \to (\partial_a u_a, v).
\]

2. For \( u \in \mathcal{F}_{(s,t)} \) and \( \tilde{\vartheta} \in C^{1}([0,1]) \) the linear interpolation from \( s \) to \( t \), we have that \( (u \circ \tilde{\vartheta}) \in \mathcal{F}_{(0,1)} \) with distributional derivative

\[
\partial_a (u \circ \tilde{\vartheta})(a) = (t-s) \partial_a u_{\tilde{\vartheta}(a)}.
\]

**Proof.** From corollary 5.6. in \([36]\) it follows for \( u \in \mathcal{F}_{(s,t)} \) and \( v \in \mathcal{F} \)

\[
\int u_{a+h} v \, dm_\cdot - \int u_a v \, dm_\cdot = \int_a^{a+h} \langle \partial_b u_b, v \rangle db.
\]
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 $a \in (s, t)$

$$
\lim_{h \to 0} \frac{1}{h} \int u_{a+h} v \, dm - \int u_a v \, dm = \lim_{h \to 0} \frac{1}{h} \int_{a}^{a+h} \langle \partial_b u_b, v \rangle \, db = \langle \partial_a u_a, v \rangle.
$$

This proves the first assertion.

To show the second, recall that we can approximate each $u \in \mathcal{F}(s, t)$ by smooth functions $(u^n) \subset C^\infty([s, t] \to \mathcal{F})$ by virtue of [36, lemma 5.3]. So for each $n \in \mathbb{N}$ and for each smooth, compactly supported test function $\psi : (0, 1) \to \mathcal{F}$, we have that

$$
\int_0^1 \int (u^n \circ \vartheta)(a) \partial_a \psi_a \, dm \, da = - \int_0^1 \int \vartheta(a) \partial_a u^n_{\vartheta(a)} \psi_a \, dm \, da.
$$

Note that the term on the left-hand side converges to

$$
\int_0^1 \int (u \circ \vartheta)(a) \partial_a \psi_a \, dm \, da \quad \text{as } n \to \infty
$$

since

$$
\left| \int_0^1 \int (u^n \circ \vartheta - u \circ \vartheta) \partial_a \psi_a \, dm \, da \right| \leq (t-s)^{-1} \int_s^t \|u^n_a - u_a\|_{\mathcal{F}} \|\partial_a \psi_{\vartheta^{-1}(a)}\|_{\mathcal{F}} \, da,
$$

where we applied integration by substitution. Similarly for the right-hand side

$$
\left| \int_0^1 \vartheta(a) \langle \partial_a u^n_{\vartheta(a)} - \partial_a u_{\vartheta(a)}, \psi_a \rangle \, dm \, da \right| \leq \int_s^t \|\partial_a u^n_a - \partial_a u_a\|_{\mathcal{F}^*} \|\psi_{\vartheta^{-1}(a)}\|_{\mathcal{F}} \, da,
$$

and consequently as $n \to \infty$

$$
\int_0^1 \int (u \circ \vartheta)(a) \partial_a \psi_a \, dm \, da = - \int_0^1 (t-s) \langle \partial_a u_{\vartheta(a)}, \psi_a \rangle \, da,
$$

which is the assertion. \[\square\]

For the lemmas below let $(\rho_a), a \in [0, 1], be a regular curve and let $\vartheta : [0, 1] \to [0, \infty)$

$$
\vartheta(a) := (1-a)s + at \quad \text{where } s < t.
$$

Set $\rho_{a, \vartheta} := \tilde{P}_{t, \vartheta(a)}(\rho_a) = u_a, \vartheta m \vartheta(a)$.

**Lemma 6.9.** The curve $(u_{a, \vartheta}), a \in [0, 1], belongs to Lip([0, 1], \mathcal{F}^*) with

$$
\|u_{a, \vartheta}\|_{L^2([0, 1] \to \mathcal{F})}
$$
and distributional derivative

\[ \partial_a u_{a,\vartheta} \in L^\infty([0,1] \rightarrow \mathcal{F}^*) \]

satisfying

\[ \partial_a u_{a,\vartheta} = -(t-s)\Delta_{\vartheta(a)} u_{a,\vartheta} + \partial_a f_{\vartheta(a)} u_{a,\vartheta} - P_{t,\vartheta(a)}^*(\hat{u}_a). \]

**Proof.** First we show that \((u_{a,\vartheta})\) is in \(L^2([0,1] \rightarrow \mathcal{F})\). For this recall that, since \((\rho_\vartheta)\) is regular, \(u_a \leq R\) and \(E_t(\sqrt{u_a}) \leq E\) for all \(a \in [0,1]\), and hence by Lemma \[2.15\] we get

\[
\int_0^1 \left\| u_{a,\vartheta} \right\|_{L^2(m_\vartheta(a))}^2 da \leq e^{L(t-s)} \int_0^1 \left\| u_a \right\|_{L^2(m_a)}^2 da \\
\leq Re^{L(t-s)} \int_0^1 \left\| u_a \right\|_{L^1(m_a)} da = Re^{L(t-s)},
\]

and by Theorem \[2.12\]

\[
\int_0^1 E_{\vartheta(a)}(u_{a,\vartheta}) da \leq e^{3L(t-s)} \left[ \int E_t(u_a) + \left\| u_a \right\|_{L^2(m_a)}^2 \right] da \\
\leq e^{3L(t-s)} \sqrt{R} \left[ \int_0^1 2E_t(\sqrt{u_a}) da + R \right] \\
\leq e^{3L(t-s)} \sqrt{R}(2E + R).
\]

This shows that \((u_{a,\vartheta})\) is in \(L^2([0,1] \rightarrow \mathcal{F})\).

Next we show that \((u_{a,\vartheta})\) is contained in \(\text{Lip}([0,1], \mathcal{F}^*)\). For this let \(\psi \in \mathcal{F}\). Then, for almost every \(a_0, a_1 \in (0,1)\), since \(P_{t,\vartheta(a)}^* u_{a_0} \in \mathcal{F}_{(0,1)}\), we obtain with Lemma \[5.8\]

\[
\int \psi u_{a_1,\vartheta} \, dm_\vartheta - \int \psi u_{a_0,\vartheta} \, dm_\vartheta \\
= \int \psi (P_{t,\vartheta(a)}^* u_{a_0} - P_{t,\vartheta(a_1)}^* u_{a_0}) \, dm_\vartheta + \int \psi P_{t,\vartheta(a_1)} (u_{a_1} - u_{a_0}) \, dm_\vartheta \\
= (t-s) \int_{a_0}^{a_1} E_{\vartheta(a)}^\vartheta (P_{t,\vartheta(a)}^* u_{a_0}, \psi) \, da \\
+ (t-s) \int_{a_0}^{a_1} f_{\vartheta(a)} P_{t,\vartheta(a)}^* u_{a_0} \psi \, dm_\vartheta \, da \\
+ \int P_{t,\vartheta(a_1)} (\psi f_{\vartheta(a_1)})(u_{a_1} - u_{a_0}) \, dm_t \leq
\]
Due to our assumptions on \( f \) we have that
\[
\text{Lip}(f_{\vartheta(a)}) \leq C, \quad \|f_{\vartheta(a)}\|_{\infty} \leq L, \quad \|f_t\|_{\infty} \leq C.
\]
while the energy estimate Theorem 2.12 and Corollary 2.15 yield
\[
\mathcal{E}_{\vartheta(a)}(P_{t,\vartheta(a)}^* u_{a_0}) \leq e^{3L(t-s)} \mathcal{E}_t(u_{a_0}) \leq e^{3L(t-s)} \mathcal{E}_t(u_{a_0}) + \|u_{a_0}\|_{L^2(\Omega)}^2.
\]
\[
\|P_{t,\vartheta(a)}^* u_{a_0}\|_{L^2(\Omega)} \leq e^{L(t-s)/2} \|u_{a_0}\|_{L^2(\Omega)}.
\]
Note that the last two expressions are bounded since \( u \) is a regular curve. Moreover, from (2.2), the gradient estimate (6.1), and Corollary 2.15, we find
\[
\mathcal{E}_\varphi(P_{t,\vartheta(a_1)}(\psi e^{f_{\vartheta(a_1)}})) \leq C e^{L(t-s)} \text{Lip}(e^{f_{\vartheta(a_1)}})^2 \mathcal{E}_{\vartheta(a_1)}(\psi).
\]
Applying (2.2) once more we find that there exists a constant \( \lambda \) such that
\[
(6.8) \quad \int \psi u_{a_1,\vartheta} \, dm_\varphi - \int \psi u_{a_0,\vartheta} \, dm_\varphi \leq (a_1 - a_0)\lambda \|\psi\|_{\mathcal{F}},
\]
and thus
\[
\|u_{a_1} - u_{a_0}\|_{\mathcal{F}} \leq \lambda.
\]
Note also that (6.8) 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
\[
\int \psi u_{a_1,\vartheta} \, dm_\varphi - \int \psi u_{a_0,\vartheta} \, dm_\varphi = \int_{a_0}^{a_1} (\partial_a u_{a,\vartheta}, \psi)_{\mathcal{F}^*,\mathcal{F}} \, da.
\]
Fix \( \psi \in \text{Lip}_b(X) \). By a similar calculation as above it ultimately follows that

\[
\begin{align*}
\lim_{h \to 0} & \frac{1}{h} \left( \int \psi u_{a+h, \theta} \, dm_\circ - \int \psi u_{a, \theta} \, dm_\circ \right) \\
= & \ (t-s) \mathcal{E}_{\partial(a)}^\circ \left( P^*_{t, \theta(a)} u_{a, \theta} \psi \right) + (t-s) \int \dot{f}_\theta(a) P^*_{t, \theta(a)} u_a \psi \, dm_\circ \\
+ & \lim_{h \to 0} \int P_{t, \theta(a+h)} \left( \psi e^{f_{\theta(a+h)}} \right) \frac{(u_{a+h} - u_a)}{h} \, dm_t \\
\end{align*}
\]

almost everywhere. To determine the last integral, since \( u \in C^1([0,1], L^1(X)) \) and \( \psi \in \text{Lip}_b(X) \), we note that

\[
\begin{align*}
\lim_{h \to 0} & \int P_{t, \theta(a+h)} \left( \psi e^{f_{\theta(a+h)}} \right) \frac{(u_{a+h} - u_a)}{h} \, dm_t \\
= & \int P_{t, \theta(a)} \left( \psi e^{f_{\theta(a)}} \right) u_a \, dm_t \\
= & \int \left( \psi e^{f_{\theta(a)}} \right) P^*_{t, \theta(a)} u_a \, dm_{\theta(a)} = \left[ P^*_{t, \theta(a)} u_a, \psi \right]_{\mathcal{F}_*, \mathcal{F}}. \\
\end{align*}
\]

From the Lipschitz continuity of \( (u_{a, \theta}, \psi) \) we deduce that for almost every \( a \in [0,1] \)

\[
(\partial_a u_{a, \theta}, \psi)_{\mathcal{F}_*, \mathcal{F}} = \left\{ -(t-s) \Delta_{\theta(a)} u_{a, \theta} + \partial_a f_{\theta(a)} u_{a, \theta} - P^*_{t, \theta(a)} (\dot{u}_a) \psi \right\}_{\mathcal{F}_*, \mathcal{F}}.
\]

We conclude the proof by approximating \( \psi \in \mathcal{F} \) with bounded Lipschitz functions. \( \square \)

**Lemma 6.10.** For any map \( \varphi \in HLS_\theta \), the map \( a \mapsto \int \varphi_a \, d\rho_{a, \theta} \) is absolutely continuous and

\[
\begin{align*}
\int \varphi_1 \, d\rho_{1, \theta} - \int \varphi_0 \, d\rho_{0, \theta} & \leq \int_0^1 \left[ -\frac{1}{2} \int \Gamma_{\partial(a)}(\varphi_a) \, d\rho_{a, \theta} \\
& \quad + \int P_{t, \theta(a)}(\varphi_a) \partial_a u_a \, dm_t \\
& \quad + (t-s) \int \Gamma_{\partial(a)}(\varphi_a, u_{a, \theta}) \, dm_{\theta(a)} \right] \, da.
\end{align*}
\]

**Proof.** Let us begin by showing that \( a \mapsto \rho_{a, \theta} \) is 2-absolutely continuous. Indeed, letting \( a_0 < a_1 \), we have with the equivalence of the gradient estimate (6.1) and the Wasserstein contraction (4.9)

\[
W_{\partial(a_0)}(\rho_{a_0, \theta}, \rho_{a_1, \theta}) \leq W_{\theta(a_0)}(\hat{P}_{t, \theta(a_0)} \rho_{a_0}, \hat{P}_{t, \theta(a_0)} \rho_{a_1}) \\
+ W_{\theta(a_0)}(\hat{P}_{t, \theta(a_0)} \rho_{a_1}, \hat{P}_{t, \theta(a_1)} \rho_{a_1}) \\
\leq W_t(\rho_{a_0}, \rho_{a_1}) + W_{\theta(a_0)}(\hat{P}_{t, \theta(a_0)} \rho_{a_1}, \hat{P}_{t, \theta(a_1)} \rho_{a_1}).
\]
By virtue of Lemma 3.7(iv) we have that $\rho_a = \tilde{\rho}_{t,\vartheta(a)}\rho_a = \tilde{u}_a m_{\vartheta(a)}$ 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

$$
\int \varphi_a \, d\rho_{a,\vartheta} = \int \varphi_a \, d\rho_{a,1,\vartheta} - \int \varphi_a \, d\rho_{a,0,\vartheta}
$$

$$
= \int (\varphi_{a,1} - \varphi_{a,0}) \, d\rho_{a,1,\vartheta} + \int \varphi_{a,0} \, d\rho_{a,1,\vartheta} - \int \varphi_{a,0} \, d\rho_{a,0,\vartheta}
$$

$$
\leq \|\varphi_a - \varphi_{a,0}\|_{\infty} + \text{Lip}(\varphi_{a,0}) W(\rho_{a,1,\vartheta}, \rho_{a,0,\vartheta}).
$$

To compute its derivative we consider difference quotients. Since $\vartheta \in \text{Lip}([0, 1], L^\infty(X))$ is in $HLS\vartheta$, and $u_{a+h,\vartheta} \rightarrow u_{a,\vartheta}$ in $L^1(X)$, we have

$$
\lim_{h \rightarrow 0} h^{-1} \int (\varphi_{a+h} - \varphi_a) \, d\rho_{a+h,\vartheta} \leq -\frac{1}{2} \int |\nabla \vartheta_{a}\varphi_a|^2 \, d\rho_{a,\vartheta}.
$$

Now we need to determine

$$
\lim_{h \rightarrow 0} \frac{1}{h} \left( \int \varphi_a e^{-f_{\vartheta(a)}} (u_{a+h,\vartheta} - u_{a,\vartheta}) \, dm_{\vartheta}
\right.
$$

$$
+ \left. \int \varphi_a u_{a+h,\vartheta} \, d(m_{\vartheta(a+h)} - m_{\vartheta(a)}) \right) .
$$

The expression on the right-hand side clearly converges to

$$
\hat{\vartheta}(a) \int \varphi_a \hat{f}_{\vartheta(a)} u_{a,\vartheta} \, dm_{\vartheta(a)},
$$

while from Lemma 6.9 we deduce

$$
\lim_{h \rightarrow 0} \int e^{-f_{\vartheta(a)}} \varphi_a \frac{1}{h} (u_{a+h,\vartheta} - u_{a,\vartheta}) \, dm_{\vartheta}
$$

$$
= (\partial_a u_{a,\vartheta}, e^{-f_{\vartheta(a)}} \varphi_a)_{\mathcal{F}, \mathcal{F}^*},
$$

$$
= (t - s) \left( \int \hat{f}_{\vartheta(a)} u_{a,\vartheta} \varphi_a e^{-f_{\vartheta(a)}} \, dm_{\vartheta} + \mathcal{E}_{\vartheta(a)} (u_{a,\vartheta}, \varphi_a e^{-f_{\vartheta(a)}}) \right)
$$

$$
(6.11) \quad = (t - s) \left( \int \hat{f}_{\vartheta(a)} u_{a,\vartheta} \varphi_a \, dm_{\vartheta(a)} + \int \Gamma_{\vartheta(a)} (u_{a,\vartheta}, \varphi_a) \, dm_{\vartheta(a)} \right).
$$

Then from the absolute continuity of $a \mapsto \int \varphi_a \, d\rho_{a,\vartheta}$ together with (6.9), (6.10), and (6.11), we obtain

$$
\int \varphi_1 \, d\rho_{1,\vartheta} - \int \varphi_0 \, d\rho_{0,\vartheta}
$$

$$
= \int_0^1 \partial_a \int \varphi_a \, d\rho_{a,\vartheta} \, da \leq
$$
We regularize the entropy functional by truncating the singularities of the logarithm. Define \( e_\varepsilon : [0, \infty) \) by setting \( e_\varepsilon'(r) = \log(\varepsilon + r) + 1 \) and \( e_\varepsilon(0) = 0 \). Then \( e_\varepsilon \) is still a convex function and \( e_\varepsilon' \in \text{Lip}_b([0, R]) \). For any \( t \) and \( \rho = u_m t \in \mathcal{P}(X) \) we define

\[
S_\varepsilon^t(\rho) = \int e_\varepsilon(u) dm_t.
\]

Note that for any \( \rho \in \text{Dom}(S) \) we clearly have \( S_\varepsilon^t(\rho) \to S(\rho) \) as \( \varepsilon \to 0 \).

As in [6] we introduce

\[
p_\varepsilon(r) := e_\varepsilon'(r^2) - \log \varepsilon.
\]

**Lemma 6.11.** With the same notation as in Lemma 6.10 we find that for any \( \varepsilon > 0 \)

\[
S_\varepsilon^t(\rho_{1, \vartheta}) - S_\varepsilon^t(\rho_{0, \vartheta})
\geq \int_0^1 \left[ -\frac{1}{2} \int |\nabla_{\vartheta(a)} \mathcal{Q}_a|^2 \, d\rho_{a, \vartheta} + \int P_{t, \vartheta(a)} \mathcal{Q}_a \, dm_t \right. \\
- (t-s) \int \mathcal{Q}_a \mathcal{G}_{\vartheta(a)} u_{a, \vartheta} \, dm_{\vartheta(a)} + (t-s) \int \mathcal{G}_{\vartheta(a)} u_{a, \vartheta} \mathcal{Q}_a \, dm_{\vartheta(a)} \\
+ (t-s) \int \mathcal{G}_{\vartheta(a)}(u_{a, \vartheta}, \mathcal{Q}_a) \, dm_{\vartheta(a)} \right] \, da.
\]

**Proof.** From the convexity of \( e_\varepsilon \) we get that for every \( a_0, a_1 \in [0, 1] \) by virtue of Lemma 6.9

\[
S_\varepsilon^t(\rho_{a_1, \vartheta}) - S_\varepsilon^t(\rho_{a_0, \vartheta})
\geq \int e_\varepsilon(u_{a_1, \vartheta}) + e_\varepsilon(u_{a_0, \vartheta}) \, dm_t \\
+ \int e_\varepsilon(u_{a_1, \vartheta})(e^{-f_{\vartheta(a_1)}} - e^{-f_{\vartheta(a_0)}}) \, dm_t \\
+ \int e_\varepsilon(u_{a_0, \vartheta})(e^{-f_{\vartheta(a_1)}} - e^{-f_{\vartheta(a_0)}}) \, dm_t \\
+ \int e_\varepsilon(u_{a_1, \vartheta})(e^{-f_{\vartheta(a_1)}} - e^{-f_{\vartheta(a_0)}}) \, dm_t =
\]
\[ \begin{aligned}
&= \int_{a_0}^{a_1} \left( (\partial_a u_{a,\theta} \cdot e^{-f_{\theta(a)}} e'_e(u_{a_0,\theta})) \
&\quad - \int e\varepsilon(u_{a_1,\theta}) \hat{\vartheta}(a) \dot{\vartheta}(a) e^{-f_{\theta(a)}} dm_a \right) da \\
&= \int_{a_0}^{a_1} \left( -\dot{\vartheta}(a) \Delta_{\vartheta(a)} u_{a,\theta} + \dot{\vartheta}(a) \dot{\vartheta}(a) u_{a,\theta} \\
&\quad + P^*_t,\vartheta(a) (\dot{u}_a) \cdot e^{-f_{\theta(a)}} e'_e(u_{a_0,\theta})) \
&\quad - \int e\varepsilon(u_{a_1,\theta}) \dot{\vartheta}(a) \dot{\vartheta}(a) e^{-f_{\theta(a)}} dm_a \right) da \\
&= \int_{a_0}^{a_1} \left( -\dot{\vartheta}(a) \Delta_{\vartheta(a)} u_{a,\theta} \cdot e^{-f_{\theta(a)}} e'_e(u_{a_0,\theta})) \
&\quad + \int \dot{\vartheta}(a) \dot{\vartheta}(a) u_{a,\theta} e^{-f_{\theta(a)}} e'_e(u_{a_0,\theta}) dm_a \\
&\quad + \int P^*_t,\vartheta(a) (\dot{u}_a) e^{-f_{\theta(a)}} e'_e(u_{a_0,\theta}) dm_a \\
&\quad - \int e\varepsilon(u_{a_1,\theta}) \dot{\vartheta}(a) \dot{\vartheta}(a) e^{-f_{\theta(a)}} dm_a \right) da.
\end{aligned} \]

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

\[ S^*_t(\rho_{1,\theta}) - S^*_t(\rho_{0,\theta}) \]

\[ = \sum_{i=1}^{n} \left( S^*_{\vartheta(a_i)}(\rho_{a_i,\theta}) - S^*_{\vartheta(a_{i-1})}(\rho_{a_{i-1},\theta}) \right) \]

\[ \geq \sum_{i=1}^{n} \int_{a_{i-1}}^{a_i} \left( -\dot{\vartheta}(a) \Delta_{\vartheta(a)} u_{a,\theta} e^{-f_{\theta(a)}} e'_e(u_{a_{i-1},\theta})) \
&\quad + \int \dot{\vartheta}(a) \dot{\vartheta}(a) u_{a,\theta} e^{-f_{\theta(a)}} e'_e(u_{a_{i-1},\theta}) dm_a \\
&\quad + \int P^*_t,\vartheta(a) (\dot{u}_a) e^{-f_{\theta(a)}} e'_e(u_{a_{i-1},\theta}) dm_a \\
&\quad - \int e\varepsilon(u_{a_{i-1},\theta}) \dot{\vartheta}(a) \dot{\vartheta}(a) e^{-f_{\theta(a)}} dm_a \right) da \\
\]

\[ = \int_{0}^{1} \left( -\dot{\vartheta}(a) \Delta_{\vartheta(a)} u_{a,\theta} \cdot \frac{h}{a} \right) + \int \dot{\vartheta}(a) \dot{\vartheta}(a) u_{a,\theta} \frac{h}{a} dm_a \\
&\quad + \int P^*_t,\vartheta(a) (\dot{u}_a) \frac{h}{a} dm_a - \int \omega\frac{h}{a} \dot{\vartheta}(a) \dot{\vartheta}(a) e^{-f_{\theta(a)}} dm_a \right) da.
\]
where

\[
\zeta_a^h = e^{-f_\theta(a_{i-1})}e'_\varepsilon(u_{a_{i-1}, \theta}) \quad \text{for } a \in (a_{i-1}, a_i]
\]

\[
\omega_a^h = e_\varepsilon(u_{a_i, \theta}) \quad \text{for } a \in (a_{i-1}, a_i].
\]

Letting \( h \to 0 \) we obtain

\[
\zeta_a^h \to e^{-f_\theta(a)}e'_\varepsilon(u_{a, \theta}) \quad \text{in } L^1(X) \text{ for a.e. } a \in (0, 1)
\]

\[
\omega_a^h \to e_\varepsilon(u_{a, \theta}) \quad \text{in } L^1(X) \text{ for a.e. } a \in (0, 1),
\]

and thus from dominated convergence

\[
S^\varepsilon_1 \rho_1, \theta \rangle - S^\varepsilon_1 \rho_0, \theta \rangle
\]

\[
\geq \limsup_{h \to 0} \left[ \int_0^1 (-\dot{\theta}(a)\Delta_\theta(a)u_{a, \theta}, \zeta_a^h) + \int \dot{\theta}(a)f_\theta(a)u_{a, \theta}s_a^h \, dm_\circ \right.
\]

\[
+ \int P_{\varepsilon}(\theta)(\hat{u}_a)s_a^h \, dm_\circ - \int \omega_a^h \dot{\theta}(a)f_\theta(a)e^{-f_\theta(a)} \, dm_\circ \, da \right]
\]

\[
\geq \limsup_{h \to 0} \left[ \int_0^1 (-\dot{\theta}(a)\Delta_\theta(a)u_{a, \theta}, \zeta_a^h) \, da \right]
\]

\[
+ \int_0^1 \left( \int \dot{\theta}(a)f_\theta(a)u_{a, \theta} e^{-f_\theta(a)}e'_\varepsilon(u_{a, \theta}) \, dm_\circ 
\]

\[
+ \int P_{\varepsilon}(\theta)(\hat{u}_a) e^{-f_\theta(a)}e'_\varepsilon(u_{a, \theta}) \, dm_\circ
\]

\[
- \int e_\varepsilon(u_{a, \theta}) \dot{\theta}(a)f_\theta(a)e^{-f_\theta(a)} \, dm_\circ \, da.
\]

To see that \( \langle \Delta_\theta(a)u_{a, \theta}, \zeta_a^h \rangle \to \langle \Delta_\theta(a)u_{a, \theta}, e^{-f_\theta(a)}e'_\varepsilon(u_{a, \theta}) \rangle \), recall that from Theorem 2.12 it suffices to show that

\[
\zeta_a^h \to e^{-f_\theta(a)}e'_\varepsilon(u_{a, \theta}) \quad \text{in } L^2(X).
\]

This is a consequence of the boundedness of \( u_{a, \theta} \) and \( f_\theta(a) \). Then again by dominated convergence we have

\[
S^\varepsilon_1 \rho_1, \theta \rangle - S^\varepsilon_1 \rho_0, \theta \rangle
\]

\[
\geq \int_0^1 \left[ \dot{\theta}(a)e_\varepsilon(\theta)(u_{a, \theta}, e^{-f_\theta(a)}e'_\varepsilon(u_{a, \theta})) 
\right.
\]

\[
+ \int \dot{\theta}(a)f_\theta(a)u_{a, \theta}e^{-f_\theta(a)}e'_\varepsilon(u_{a, \theta}) \, dm_\circ
\]

\[
+ \int P_{\varepsilon}(\theta)(\hat{u}_a)e^{-f_\theta(a)}e'_\varepsilon(u_{a, \theta}) \, dm_\circ
\]

\[
- \int e_\varepsilon(u_{a, \theta})\dot{\theta}(a)f_\theta(a)e^{-f_\theta(a)} \, dm_\circ \, da =
\]
\[= \int_0^1 \left[ \hat{\theta}(a) \mathcal{E}_\theta(u_{a, \hat{q}}, \epsilon'_\delta(u_{a, \hat{q}})) + \int \hat{\theta}(a) f_{\theta(a)} u_{a, \hat{q}} \epsilon'_\delta(u_{a, \hat{q}}) dm_{\theta(a)} \\
+ \int P_{\theta(a)^*} \epsilon'_\delta(u_{a, \hat{q}}) dm_{\theta(a)} - \int \epsilon'_\delta(u_{a, \hat{q}}) f_{\theta(a)} dm_{\theta(a)} \right] da. \]

### 6.3 The Dynamic EVI Property

**Proposition 6.12.** Let \( \rho^a = u^a m_t \) be a regular curve. Then setting \( \rho_{a, \theta} = \hat{P}_{\theta(a)^*} \rho^a \), we have

\[
(6.12) \quad \frac{1}{2} \int \dot{\rho}_a^2 (\rho_1, \theta) - (t - s)(S_t(\rho_1, \theta) - S_s(\rho_0, \theta)) \leq \frac{1}{2} \int_0^1 |\dot{\rho}_a|^2 \, da - (t - s)^2 \int_0^1 \dot{f}_{\theta(a)} \, d\rho_{a, \theta} \, da.
\]

**Proof.** Applying Lemma [6.10] and Lemma [6.11] we find

\[
\int \varphi_1 \, d\rho_{1, \theta} - \int \varphi_0 \, d\rho_{0, \theta} - (t - s) (S_t^\epsilon(\rho_1, \theta) - S_s^\epsilon(\rho_0, \theta)) \leq \int_0^1 \left[ \int \dot{\varphi}_a P_{\theta(a)}(\varphi_a - (t - s) \epsilon'_\delta(u_{a, \hat{q}})) \, dm_t \right. \\
- \frac{1}{2} \int \Gamma_{\theta(a)}(\varphi_a) \, d\rho_{a, \theta} + (t - s) \int \Gamma_{\theta(a)}(\varphi_a, u_{a, \hat{q}}) \, dm_{\theta(a)} \\
- 4(t - s)^2 \int \epsilon''(u_{a, \theta}) \Gamma_{\theta(a)}(\sqrt{u_{a, \theta}}) \, d\rho_{a, \theta} \\
- (t - s)^2 \int (\epsilon_\delta(u_{a, \theta}) - \epsilon'_\delta(u_{a, \hat{q}}) u_{a, \theta}) \dot{f}_{\theta(a)} \, dm_{\theta(a)} \] da.

Then since

\[
4t \epsilon''(r) \geq 4t^2 (\epsilon'_\delta(r))^2 = r (p'_\delta(\sqrt{r}))^2,
\]

we can estimate

\[-4u_{a, \theta} \epsilon''(u_{a, \theta}) \Gamma_{\theta(a)}(\sqrt{u_{a, \theta}}) \leq -u_{a, \theta} (p'_\delta(\sqrt{u_{a, \theta}}))^2 \Gamma_{\theta(a)}(\sqrt{u_{a, \theta}}) \]

and while, with \( q_\delta(r) := \sqrt{r}(2 - \sqrt{r} p'_\delta(\sqrt{r})) \),

\[
\Gamma_{\theta(a)}(u_{a, \theta}, \varphi_a) = 2 \sqrt{u_{a, \theta}} \Gamma_{\theta(a)}(\sqrt{u_{a, \theta}}, \varphi_a) \\
= u_{a, \theta} \Gamma_{\theta(a)}(\sqrt{u_{a, \theta}}, \varphi_a) + q_\delta(u_{a, \theta}) \Gamma_{\theta(a)}(\sqrt{u_{a, \theta}}, \varphi_a)
\]
we find

\[
\int \varphi_1 \, d\rho_{1,\theta} - \int \varphi_0 \, d\rho_{0,\theta} - (t-s)(S_t^x(\rho_{1,\theta}) - S_t^x(\rho_{0,\theta}))
\]

\[
\leq \int_0^1 \left[ \int \dot{u}_a P_{t,\theta(a)}(\varphi_a - (t-s)e'_s(u_{a,\theta})) \, dm_t - \frac{1}{2} \int \Gamma_{\theta(a)}(\varphi_a) \, d\rho_{a,\theta} \\
+ (t-s) \int \Gamma_{\theta(a)}(\varphi_a, p_e(\sqrt{u_{a,\theta}})) \, d\rho_{a,\theta} \\
- (t-s)^2 \int \Gamma_{\theta(a)}(p_e(\sqrt{u_{a,\theta}})) \, d\rho_{a,\theta} \\
+ (t-s) \int q_e(u_{a,\theta}) \Gamma_{\theta(a)}(\sqrt{u_{a,\theta}}, \varphi_a) \, dm_{\theta(a)} \\
- (t-s)^2 \int (e_e(u_{a,\theta}) - e'_s(u_{a,\theta})) \, d\rho_{a,\theta} \right] \, da. 
\]

(6.14)

Hence, by means of (6.7), the gradient estimate (6.1), and the Young inequality \(2xy \leq \delta x^2 + y^2 / \delta\), this yields

\[
\int \varphi_1 \, d\rho_{1,\theta} - \int \varphi_0 \, d\rho_{0,\theta} - (t-s)(S_t^x(\rho_{1,\theta}) - S_t^x(\rho_{0,\theta}))
\]

\[
\leq \int_0^1 \left[ \frac{1}{2} |\dot{\rho}_a|^2 + \frac{1}{2} \int P_{t,\theta(a)}(\varphi_a - (t-s)\varphi'_e(u_{a,\theta})) \, d\rho_a \\
- \frac{1}{2} \int P_{t,\theta(a)}(\varphi_a - (t-s)p_e(\sqrt{u_{a,\theta}})) \, d\rho_a \\
+ (t-s) \int q_e(u_{a,\theta}) \Gamma_{\theta(a)}(\sqrt{u_{a,\theta}}, \varphi_a) \, dm_{\theta(a)} \\
- (t-s)^2 \int (e_e(u_{a,\theta}) - e'_s(u_{a,\theta})) \, d\rho_{a,\theta} \right] \, da
\]

\[
\leq \int_0^1 \left[ \frac{1}{2} |\dot{\rho}_a|^2 + (t-s) \int |q_e(u_{a,\theta})| \| \Gamma_{\theta(a)}(\sqrt{u_{a,\theta}}, \varphi_a) \| \, dm_{\theta(a)} \\
- (t-s)^2 \int (e_e(u_{a,\theta}) - e'_s(u_{a,\theta})) \, d\rho_{a,\theta} \right] \, da
\]

\[
\leq \int_0^1 \left[ \frac{1}{2} |\dot{\rho}_a|^2 + \frac{(t-s)}{2\delta} \int (q_e(u_{a,\theta}))^2 \Gamma_{\theta(a)}(\varphi_a) \, dm_{\theta(a)} \\
+ (t-s) \delta \frac{1}{2} \int \Gamma_{\theta(a)}(\sqrt{u_{a,\theta}}) \, dm_{\theta(a)} \\
- (t-s)^2 \int (e_e(u_{a,\theta}) - e'_s(u_{a,\theta})) \, d\rho_{a,\theta} \right] \, da.
\]
We first pass to the limit $\varepsilon \to 0$,
\[
\lim_{\varepsilon \to 0} q_\varepsilon^2(r) = 0, \quad q_\varepsilon^2(r) = 4r \left(1 - \frac{r}{\varepsilon + r}\right)^2 \leq 4r,
\]
\[
\lim_{\varepsilon \to 0} (e_\varepsilon(r) - re_\varepsilon'(r)) = -r.
\]
\[
|e_\varepsilon(r) - re_\varepsilon'(r)| \leq 2(\varepsilon + r)|\log(\varepsilon + r)| + r + \varepsilon \log \varepsilon
\]
\[
\leq 2\sqrt{\varepsilon + r} + r + \varepsilon \log \varepsilon,
\]
and then $\delta \to 0$,
\[
\int \varphi_1 d\rho_{1,\theta} - \int \varphi_0 d\rho_{0,\theta} - (t-s)(S_\tau(\rho_{1,\theta}) - S_\tau(\rho_{0,\theta})) \leq \int_0^1 \left[\frac{1}{2} |\tilde{\rho}_a'|^2 + (t-s)^2 \int f_{\partial(a)} d\rho_{a,\theta} \right] da.
\]
Taking the supremum over $\varphi$ we obtain the desired estimate \[6.12\].

**THEOREM 6.13.** Assume that the gradient estimate holds true for the time-dependent metric measure space $(X, d_\tau, m_\tau)$, $\tau \in (0, T)$. Then for every $\tau \in (0, T]$ the dual heat flow emanating from $\mu$ solves
\[
S_{\tau}(\mu_s) - S_{\tau}(\sigma) \leq \frac{1}{2(t-s)} (W_{\tau}(\mu_s, \sigma) - W_{\tau}(\mu_s, \sigma)) - (t-s) \int_0^1 \int f_{\partial(a)} d\rho_{a,\theta} da
\]
for all $s \in (0, \tau)$ and all $\sigma, \mu \in \text{Dom}(S)$. Here $(\rho_\alpha)$, $\alpha \in [0, 1]$, denotes the $W_t$-geodesic connecting $\rho_0 = \mu_1$, $\rho_1 = \sigma$, and $\rho_{a,\theta} = \tilde{P}_{t,\theta}(\rho_\alpha)$.

In particular, $\mu_t$ is a dynamic upward $\text{EVIT}^+$-gradient flow; i.e., for every $t \in (0, \tau)$ and every $\sigma \in \text{Dom}(S)$, we have
\[
\frac{1}{2} \frac{d}{ds} W_{s,\tau}(\mu_s, \sigma) \bigg|_{s=t} \geq S_{\tau}(\mu_t) - S_{\tau}(\sigma).
\]

**PROOF.** Let $(\rho_\alpha)$, $\alpha \in [0, 1]$, be a $W_t$-geodesic connecting $\mu$ and $\sigma$, which exists and is unique. We approximate the geodesic $(\rho_\alpha)$, $\alpha \in [0, 1]$, by regular curves $(\rho^n_\alpha)$, $\alpha \in [0, 1]$. Proposition \[6.12\] states that for each $(\rho^n_\alpha)$, $\alpha \in [0, 1]$
\[
\frac{1}{2} \tilde{W}_\theta^2(\rho^n_{1,\theta}, \rho^n_{0,\theta}) - (t-s)(S_t(\rho^n_{1,\theta}) - S_t(\rho^n_{0,\theta})) \leq \frac{1}{2} \int_0^1 |\tilde{\rho}_a'|^2 da - (t-s)^2 \int_0^1 \int f_{\partial(a)} d\rho_{a,\theta} da.
\]
Since for every $\alpha \in [0, 1]$, $\rho^n_\alpha$ converges to $\rho_\alpha$ in duality with bounded continuous functions, $\rho^n_{a,\theta}$ converges to $\rho_{a,\theta}$ in duality with bounded continuous functions as well. By virtue of Lemma \[6.5\] we obtain
\[
\liminf_{n \to \infty} \tilde{W}_\theta^2(\rho^n_{1,\theta}, \rho^n_{0,\theta}) \geq \tilde{W}_\theta^2(\rho_{1,\theta}, \rho_{0,\theta}).
\]
Note that \((\rho^n_\alpha)\) also converges to \(\rho_\alpha\) in duality with \(L^\infty\) functions, since Lemma 3.2 provides \(\sup_n S_t(\rho^n_\alpha) < \infty\). The same argument applies then to \(\rho^n_{\alpha,\theta}\). Hence
\[
\lim_{n \to \infty} \int f_\theta(a) d\rho^n_{\alpha,\theta} = \int f_\theta(a) d\rho_{\alpha,\theta}.
\]
Then we end up with
\[
\frac{1}{2} W^2_\theta(\mu_\alpha, \sigma) - (t - s)(S_t(\sigma) - S_s(\mu_\alpha)) \leq \frac{1}{2} W^2_t(\mu_\alpha, \sigma) - (t - s)^2 \int_0^1 \int f_\theta(a) d\rho_{\alpha,\theta} da.
\]
Applying Corollary 6.7 we obtain
\[
(t - s)(S_s(\mu_\alpha) - S_t(\sigma)) \leq \frac{1}{2} W^2_t(\mu_\alpha, \sigma) - \frac{1}{2} W^2_{s,t}(\mu_\alpha, \sigma) - (t - s)^2 \int_0^1 \int f_\theta(a) d\rho_{\alpha,\theta} da.
\]
Dividing by \(t - s\) and letting \(s \not\to t\), we find
\[
S_t(\mu_\alpha) - S_t(\sigma) \leq \liminf_{s \not\to t} \frac{1}{2(t - s)} (W^2_t(\mu_\alpha, \sigma) - W^2_{s,t}(\mu_\alpha, \sigma)) \\
= \frac{1}{2} \partial_s W^2_{s,t}(\mu_\alpha, \sigma) \big|_{s = t}.
\]

### 6.4 Summarizing

The precise integrated version (6.15) of the EVI\(^-\)-property also implies a relaxed version of the EVI\(^+\)-property, which then in turn allows us to prove uniqueness of dynamic EVI flows for the entropy.

**Corollary 6.14.** The gradient estimate (III) implies the EVI\(^+\)(\(-2L, \infty\))-property. More precisely, for every \(\mu \in \text{Dom}(S)\) and every \(\tau \leq T\) the dual heat flow \(\mu_t := \hat{P}_{t,\tau}\mu\) emanating from \(\mu\) satisfies
\[
\frac{1}{2} \partial_s W^2_{s,t}(\mu_\alpha, \sigma) \big|_{s = t} \geq S_t(\mu_\alpha) - S_t(\sigma) - LW^2_t(\mu_\alpha, \sigma)
\]
for all \(t < \tau\) and all \(\sigma \in \mathcal{P}(X)\).

**Proof.** Let \(\mu_t := \hat{P}_{t,\tau}\mu\) for \(t < \tau\) and consider (6.15) for fixed \(s < t\). Then as \(t \searrow s\)
\[
S_s(\mu_\alpha) - S_s(\sigma) = \lim_{t \searrow s} (S_s(\mu_\alpha) - S_t(\sigma))
\leq \lim_{t \searrow s} \frac{1}{2(t - s)} [W^2_t(\mu_\alpha, \sigma) - W^2_{s,t}(\mu_\alpha, \sigma)] \leq
\]
\[
\left(\lim_{t \to s} \frac{1}{2(t-s)} \left[ W_{t,s}^2(\mu_t, \sigma) - W_s^2(\mu_s, \sigma) \right] + \frac{L}{2} \left[ W_t^2(\mu_t, \sigma) + W_s^2(\mu_s, \sigma) \right] \right)
= \frac{1}{2} \partial_t^* W_{t,s}^2(\mu_t, \sigma)_{t=s^+} + L W_s^2(\mu_s, \sigma)
\]
where the last estimate follows from (A.3). \hfill \square

**Corollary 6.15.** 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 from some \(\mu \in \mathcal{P}\). Then
\[
\mu_t = \hat{P}_{t,t} \mu
\]
for all \(t \in (\sigma, \tau)\). That is, the dual heat flow is the unique dynamic backward EVI\(^{-}\)-flow for the Boltzmann entropy.

**Proof.** The assertion follows from applying Corollary A.8 together with Corollary 6.14 and Theorem 6.13. \hfill \square

**Theorem 6.16.** The gradient estimate (III\(_N\)) implies the dynamic \(N\)-convexity of the Boltzmann entropy (I\(_N\)).

**Proof.** According to Theorem 4.7 and Theorem 6.13 the gradient estimate (III\(_N\)) implies both
\begin{itemize}
  \item the transport estimate (II\(_N\)) and
  \item the EVI\(^{-}\)(0, \(\infty\))-property.
\end{itemize}
According to Theorem A.11 and Remark A.12 both properties together imply dynamic \(N\)-convexity. \hfill \square

**Appendix: EVI and Dynamic Convexity**

**A.1 Time-Dependent Geodesic Spaces**

For this section, our basic setting will be a space \(X\) equipped with a 1-parameter family of complete geodesic metrics \(d_t\) where \(t \in I \subset \mathbb{R}\) is a bounded open interval, say for convenience \(I = (0, T)\). (More generally, one might allow \(d_t\) to be pseudometrics 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 be a constant \(L \in \mathbb{R}\) (log-Lipschitz bound) such that
\[
\log \frac{d_t(x, y)}{d_s(x, y)} \leq L \cdot |t - s|
\]
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 A.1. Given $s, t \in I$ and $x, y \in X$ we let

\[(A.2)\]

$$d_{s,t}(x, y) := \inf \left\{ \int_0^1 |\dot{\gamma}^a|^2_{s+a(t-s)} da \right\}^{1/2}$$

where the infimum runs over all absolutely continuous curves $(\gamma^a), a \in [0, 1]$, in $X$ connecting $x$ and $y$.

PROPOSITION A.2.

(i) The infimum in the above formula is attained. Every 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 $\gamma$ 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 $x, y \in X$

$$1 - e^{-Lt\cdot|t-s|} \leq \frac{d_{s,t}(x, y)}{d_s(x, y)} \leq \frac{e^{Lt\cdot|t-s|} - 1}{L|t-s|}.$$  

Thus, in particular,

\[(A.3)\]

$$\left| \partial_t d_{s,t}(x, y) \right|_{t=s} \leq \frac{L}{2} d_s(x, y).$$

(iv) For all $s < t \in I$ and $x, y \in X$

\[(A.4)\]

$$d_{s,t}(x, y) = \lim_{\delta \to 0} \inf_{(t_i, x_i), \delta} \left\{ \sum_{i=1}^k \frac{t - s}{t_i - t_{i-1}} d_{t_i}^2(x_i, x_{i-1}) \right\}^{1/2}$$

where the infimum runs over all $k \in \mathbb{N}$, all partitions $(t_i)_{i=0,...,k}$ of $[s, t]$ with $t_0 = s$, $t_k = t$ and $|t_i - t_{i-1}| \leq \delta$, as well as over all $x_i \in X$ with $x_0 = x, x_k = y$.

PROOF.

(i) For each absolutely continuous curve $(\gamma^a), a \in [0, 1]$,

$$\left( \int_0^1 |\dot{\gamma}^a|^2_{s+a(t-s)} da \right)^{1/2} \geq \int_0^1 |\dot{\gamma}^a|_{s+a(t-s)} da$$

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 “≥” inequality. Conversely, the action of the concatenation of any minimizer of $d_{s,r}(x, z)$ and $d_{r,t}(z, y)$ is bounded by the scaled actions of the two minimizers. This proves the “≤” inequality.
(iii) The log-Lipschitz continuity of the distance implies that for each absolutely continuous curve
\[ e^{-La[t-s]} \int_0^1 |\dot{\gamma}^a_s|_s da \leq \int_0^1 |\dot{\gamma}^a_{s+a(t-s)}| da \leq e^{La[t-s]} \int_0^1 |\dot{\gamma}^a_s|_s da. \]
(iv) See Section 6.1 for the argument in the case of \( W_{s,t} \). \( \square \)

A.2 EVI Formulation of Gradient Flows

For the subsequent discussion, a lower semibounded 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, \( \text{Dom}(V) = \{ x \in X : V_t(x) < \infty \} \) is independent of \( x \)) and such that for each \( t \in I \) the function \( x \mapsto V_t(x) \) is \( \kappa \)-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 \( \text{Dom}(V) \) are unique.

The results obtained here will be applied to the following settings:
- the Boltzmann entropy \( S_t \) on the time-dependent geodesic space \( (\mathcal{P}, W_t) \), \( t \in I \), as well as
- the Dirichlet energy \( E_t \) 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 A.3.** Given a left-open interval \( J \subset I \), an absolutely continuous curve \((x_t)_{t \in J}\) will be called a *dynamic backward EVI\(^-\)-gradient flow* for \( V \) if for all \( t \in J \) and all \( z \in \text{Dom}(V_t) \)
\[ \frac{1}{2} \partial_s d^2_{s,t}(x_s, z) \big|_{s=t-} \geq V_t(x_t) - V_t(z) \]  
where \( d_{s,t} \) is defined in Definition [A.1].

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 from \( x' \in X \) if
\[ \lim_{t \searrow \sup J} x_t = x'. \]

Being a dynamic backward EVI\(^\pm\)-gradient flow for \( V \) obviously implies that \( x_t \in \text{Dom}(V_t) \) for all \( t < \tau \).

**Remark.** Note that these definitions are slightly different from one presented in [51]. If \( d_s \) depends smoothly on \( s \), then
\[ \partial_s d^2_{s,t}(x_s, z) \big|_{s=t-} = \partial_s d^2_t(x_s, z) \big|_{s=t-} + \partial_s d^2_{s,t}(x_t, z) \big|_{s=t-} \]
and always \( \frac{d^2}{ds^2} d^2_{s,t}(x_t, z) \big|_{s=t} \geq b_t^0(y) \) for any \( d_t \)-geodesic \( \gamma \) connecting \( x_t \) and \( z \).

Often, we ask for an improved notion of dynamic backward EVI gradient flows, involving parameters \( N \in (0, \infty] \) (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 A.4.** 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 \( z \in \text{Dom}(V_t) \) and all \( t \in (\sigma, \tau) \)

\[
\left( A.6 \right) \quad \frac{1}{2} \frac{d^2}{ds^2} d^2_{s,t}(x_t, z) \big|_{s=t} - \frac{K}{2} \cdot d_t^2(x_t, z) \geq V_t(x_t) - V_t(z) + \frac{1}{N} \int_0^1 \left( \partial_a V_t(\gamma^a) \right)^2 (1 - a) da
\]

where \( \gamma \) denotes the \( d_t \)-geodesic connecting \( x_t \) and \( z \).

Analogously, we define dynamic backward EVI\((K, N)\)-gradient flows for \( V \).

In the case \( K = 0 \), dynamic backward EVI\((K, N)\)-gradient flows will simply be 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 a positive time direction, it follows the upward gradient of the entropy.

On the other hand, in the 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 A.5.** 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 \( z \in \text{Dom}(V_t) \) and all \( t \in (\sigma, \tau) \)

\[
\left( A.7 \right) \quad -\frac{1}{2} \frac{d^2}{ds^2} d^2_{s,t}(x_t, z) \big|_{s=t} - \frac{K}{2} \cdot d_t^2(x_t, z) \geq V_t(x_t) - V_t(z) + \frac{1}{N} \int_0^1 \left( \partial_a V_t(\gamma^a) \right)^2 (1 - a) da
\]

where \( \gamma \) denotes the \( d_t \)-geodesic connecting \( x_t \) and \( z \).

We say that a forward gradient flow emanates from 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 A.6. For each dynamic backward EVI\(\hat{\Phi}(K, \infty)\)-gradient flow \((x_t)_t \in (\sigma, \tau)\) for \(V\)
\[
\int_{\sigma}^{\tau} V_t(x_t)\, dt < \infty.
\]

PROOF. Choose \(z \in \text{Dom}(V)\), apply the EVI\((K, \infty)\)-property at time \(t\), and then integrate w.r.t. time \(t\):
\[
\int_{\sigma}^{\tau} V_t(x_t)\, dt \leq \int_{\sigma}^{\tau} \left[ V_t(z) + \frac{1}{2} \partial_s d^2_{s,t}(x_s, z) \right] \, dt
\]
\[
\leq (C_0 V_t(z) + C_1)(\tau - \sigma)
\]
\[
+ \frac{1}{2} \int_{\sigma}^{\tau} \left[ \partial_t d^2_t(x_t, z) + (L - K)d^2_t(x_t, z) \right] \, dt
\]
\[
= (C_0 V_t(z) + C_1)(\tau - \sigma) + \frac{1}{2} d^2_\tau(x_\tau, z)
\]
\[
- \frac{1}{2} d^2_\sigma(x_\sigma, z) + \frac{L - K}{2} \int_{\sigma}^{\tau} d^2_t(x_t, z)\, dt.
\]

Obviously, the right-hand side is finite, which thus proves the claim.

\[\square\]

A.3 Contraction Estimates

THEOREM A.7. Given two curves \((x_t)_t \in (\sigma, \tau)\) and \((y_t)_t \in (\sigma, \tau)\), one of which is a dynamic backward EVI\((K, N)\)-gradient flow for \(V\) and the other a dynamic backward EVI\(\hat{\Phi}(K, N)\)-gradient flow for \(V\), then for all \(\sigma < s < t < \tau\)
\[
d^2_s(x_s, y_s) \leq e^{-2K(t-s)} \cdot d^2_t(x_t, y_t)
\]

(A.8)
\[
- \frac{2}{N} \int_{s}^{t} e^{-2K(r-s)} \cdot |V_r(x_r) - V_r(y_r)|^2 \, dr.
\]

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\(\hat{\Phi}\)-gradient flow for \(V\).

This implies that \(r \mapsto d_r(x_r, y_r)\) is absolutely continuous since
\[
|d_t(x_t, y_t) - d_s(x_s, y_s)| \leq d_s(x_s, x_t) + d_s(y_s, y_t) + L(t-s)d_t(x_t, y_t).
\]

Thus by the very definition of EVI flows
\[
d^2_t(x_t, y_t) - d^2_s(x_s, y_s)
\]
\[
= \lim_{\delta \searrow 0} \left[ \frac{1}{\delta} \int_{t-\delta}^{t} d^2_r(x_r, y_r)\, dr - \frac{1}{\delta} \int_{s+\delta}^{s+\delta} d^2_r(x_r, y_r)\, dr \right]
\]
\[
= \lim_{\delta \searrow 0} \frac{1}{\delta} \int_{s+\delta}^{t} \left[ d^2_r(x_r, y_r) - d^2_{r-\delta}(x_{r-\delta}, y_{r-\delta}) \right] \, dr \geq
\]
\[
\begin{align*}
&\geq \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_s^{t} \left[ d_r^2(x_r, y_r) - d_{r, r-\delta}^2(x_r, y_{r-\delta}) \right] dr \\
&+ \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_s^{t} \left[ d_{r, r-\delta}^2(x_r, y_{r-\delta}) - d_{r-\delta}^2(x_{r-\delta}, y_{r-\delta}) \right] dr \\
&= \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_s^{t} \left[ d_r^2(x_r, y_r) - d_{r, r-\delta}^2(x_r, y_{r-\delta}) \right] dr \\
&+ \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_s^{t-\delta} \left[ d_{r+\delta, r}^2(x_{r+\delta}, y_r) - d_r^2(x_r, y_r) \right] dr \\
&\overset{(*)}{\geq} \int_s^{t} \liminf_{\delta \searrow 0} \frac{1}{\delta} \left[ d_r^2(x_r, y_r) - d_{r, r-\delta}^2(x_r, y_{r-\delta}) \right] dr \\
&+ \int_s^{t} \liminf_{\delta \searrow 0} \frac{1}{\delta} \left[ d_{r+\delta, r}^2(x_{r+\delta}, y_r) - d_r^2(x_r, y_r) \right] dr \\
\geq 2 \int_s^{t} \left[ \frac{K}{2} d_r^2(x_r, y_r) + V_r(y_r) - V_r(x_r) + \frac{1}{N} \int_0^1 (\partial_a V_r(y_r^a))^2 a da \right] dr \\
&+ 2 \int_s^{t} \left[ \frac{K}{2} d_r^2(x_r, y_r) + V_r(x_r) - V_r(y_r) \right. \\
&\left. \quad + \frac{1}{N} \int_0^1 (\partial_a V_r(y_r^a))^2 (1-a) da \right] dr \\
&= 2K \int_s^{t} d_r^2(x_r, y_r) dr + \frac{2}{N} \int_s^{t} \int_0^1 (\partial_a V_r(y_r^a))^2 da dr \\
&\geq 2K \int_s^{t} d_r^2(x_r, y_r) dr + \frac{2}{N} \int_s^{t} |V_r(x_r) - V_r(y_r)|^2 dr.
\end{align*}
\]

Dividing by \( t - s \) and passing to the limit \( t - s \searrow 0 \) yields

\[
\partial_t d_r^2(x_t, y_t) \geq 2K d_r^2(x_t, y_t) + \frac{2}{N} |V_t(x_t) - V_t(y_t)|^2
\]

for a.e. \( t \). The claim now follows via variation of constants.

It remains to justify the interchange of \( \liminf_{\delta \searrow 0} \) and \( \int \cdots dr \) in \((*)\), which requires quite some effort. Recall from Proposition A.2 that

\[
\left| \frac{d_{s,d}^2(x, y)}{d_r^2(x, y)} - 1 \right| \leq 2L \cdot |t - s| \quad \text{for all } x, y, s, t \text{ with } |t - s| \leq \frac{1}{L}.
\]

Thus we can estimate

\[
-\frac{1}{\delta} \left[ d_r^2(x_r, y_r) - d_{r, r-\delta}^2(x_r, y_{r-\delta}) \right] \\
\leq -\frac{1}{\delta} \left[ d_r^2(x_r, y_r) - d_{r-\delta}^2(x_{r-\delta}, y_{r-\delta}) \right] + o_1 =
\]
A.4 Dynamic Convexity

for any $s < t$ (A.9)

d flow can be extended to a flow terminating in any $x$ \( x \in X \)

d gradient flow terminates in \( x \) \( x \in X \) and satisfies

\[
\frac{1}{\delta} \int_{r-\delta}^r \partial_s d_s^2(x_s, y_s) ds + o_1 \leq 0
\]
\[
\leq -\frac{1}{\delta} \int_{r-\delta}^r \partial_t d^2_{s,t}(x_r, y_r) \big|_{t=s} ds + o_1 + o_2
\]
\[
\leq \frac{2}{\delta} \int_{r-\delta}^r [V_s(x_r) - V_s(y_s)] ds + o_1 + o_2 + o_3
\]
\[
\leq 2C_0 \cdot V_r(x_r) + 2C_1 + C + o_1 + o_2 + o_3
\]

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 let \( o_1(r, \delta) = 2L d_r^2(x_r, y_{r-\delta}) \), \( o_2(r, \delta) = 2L \frac{1}{\delta} \int_{r-\delta}^r d_r^2(x_r, y_r) d\sigma \), and \( o_3(r) = K d_r^2(x_r, y_r) \). 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

\[
-\frac{1}{\delta} \left[ d_{r+\delta,r}(x_{r+\delta}, y_r) - d_r^2(x_r, y_r) \right]
\]
\[
\leq -\frac{1}{\delta} \int_{r}^{r+\delta} \partial_s d_s^2(x_s, y_r) ds + o'_1
\]
\[
\leq 2C_0 \cdot V_r(y_r) + 2C_1 + C + o'_1 + o'_2 + o'_3.
\]

In both cases, the final expression is integrable w.r.t. \( r \in [s, t] \) according to Lemma A.6 since by assumption \( V_t(x_t) < \infty \) as well as \( V_t(y_t) < \infty \).

**Corollary A.8.** Assume that \( (x_t)_{t \in (s, \tau)} \) is a dynamic backward \( \text{EVI}(K, N) \)-gradient flow for \( V \) and that \( (y_t)_{t \in (s, \tau)} \) is a dynamic backward \( \text{EVI}^+(K, N) \)- or \( \text{EVI}^-(K, N) \)-gradient flow for \( V \) emanating from the same point \( x_\tau = y_\tau \). Then

\[
x_t = y_t \quad \text{for all } t \leq \tau.
\]

**Corollary A.9.** Assume that for given \( \tau \), a dynamic upward \( \text{EVI}(K, \infty) \)-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

\[
d_s(x_s, y_s) \leq e^{-K(t-s)} \cdot d_t(x_t, y_t)
\]

for any \( s < t \leq \tau \).

A.4 Dynamic Convexity

Let us recall the notion of dynamic convexity as introduced in [51].

**Definition A.10.** 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 \((y^a)\),
$a \in [0, 1]$, with $\gamma^0, \gamma^1 \in \text{Dom}(V_t)$,

\[
\begin{align*}
  \frac{\partial}{\partial_a} V_t(\gamma_{1-a}^a) - \frac{\partial}{\partial_a} V_t(\gamma^0) & \geq -\frac{1}{2} \partial_s d_a^2(\gamma^0, \gamma^1) + \frac{K}{2} d_a^2(\gamma^0, \gamma^1) \\
  & + \frac{1}{N} \left| V_t(\gamma^0) - V_t(\gamma^1) \right|^2 .
\end{align*}
\]

(A.10)

**Theorem A.11.** Assume that for each $t \in I$ and each $x' \in \text{Dom}(V_t)$ there exists a dynamic backward EVI$(K, N)$-gradient flow $(x_s)_{s \in [\sigma, t]}$ for $V$ emanating from $x'$ and such that $\lim_{s \nearrow t} V_s(x_s) = V_t(x_t)$. Then $V$ is strongly dynamically $(K, N)$-convex.

**Remark A.12.** To be more precise, we request the inequality (A.5) at the point $t$ and the inequality (A.6) at all times before $t$.

**Proof.** Fix $t \in I$ and a $d_t$-geodesic $(\gamma^a)$, $a \in [0, 1]$, with $\gamma^0, \gamma^1 \in \text{Dom}(V_t)$. The a priori assumption of $\kappa$-convexity implies $\gamma^a \in \text{Dom}(V_t)$ for all $a \in [0, 1]$. For each $a$, let $(\gamma_s)_{s \leq t}$ denote the EVI$_N$-gradient flow for $V$ emanating from $\gamma^a = \gamma^a_t$. Then for all $a \in (0, 1/2)$

\[
V_t(\gamma^a) - V_t(\gamma^0) \leq \frac{1}{2} \partial_s d_s, d_s \left( \gamma_s^a, \gamma^0 \right) \bigg|_{s=t^-} \\
\leq \frac{1}{2} \partial_s d_s \left( \gamma_s^a, \gamma^0 \right) \bigg|_{s=t^-} + a^2 L d_a^2(\gamma^0, \gamma^1)
\]

(due to the log-Lipschitz continuity of $s \mapsto d_s$) and

\[
V_t(\gamma^{1-a}) - V_t(\gamma^1) \leq \frac{1}{2} \partial_s d_s, d_s \left( \gamma_s^{1-a}, \gamma^1 \right) \bigg|_{s=t^-} \\
\leq \frac{1}{2} \partial_s d_s \left( \gamma_s^{1-a}, \gamma^1 \right) \bigg|_{s=t^-} + a^2 L d_a^2(\gamma^0, \gamma^1).
\]

Moreover, Theorem [A.7] implies

\[
0 \leq \liminf_{s \nearrow t} \frac{1}{t-s} \left[ \frac{1}{2} d_t^2(\gamma^a, \gamma^{1-a}) - \frac{1}{2} d_s^2(\gamma_s^a, \gamma_s^{1-a}) - K d_t^2(\gamma^a, \gamma^{1-a}) \\
- \frac{1}{N} \int_s^t \left| V_r(\gamma^a) - V_r(\gamma_r^{1-a}) \right|^2 dr \right]
\]

\[
= \frac{1}{2} \partial_s d_s^2(\gamma_s^a, \gamma_s^{1-a}) \bigg|_{s=t^-} - K d_t^2(\gamma^a, \gamma^{1-a}) - \frac{1}{N} \left| V_t(\gamma^a) - V_t(\gamma^{1-a}) \right|^2 .
\]

(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

$$\frac{1}{a} \left[ V_t(y^a) - V_t(y^0) + V_t(y^{1-a}) - V_t(y^1) \right]$$

$$\leq \liminf_{s \to t} \frac{1}{2(t-s)} \left( \frac{1}{a} d^2_s(y^0, y^a) + \frac{1}{1-2a} d^2_t(y^a, y^{1-a}) + \frac{1}{a} d^2_t(y^{1-a}, y^1) \right)$$

$$- \frac{1}{N(1-2a)} \left[ |V_t(y^a) - V_t(y^{1-a})|^2 \right]$$

$$+ 2aL d^2_t(y^0, y^1) - \frac{K}{1-2a} d^2_t(y^a, y^{1-a})$$

$$- \frac{1}{N(1-2a)} \left[ d^2_s(y^0, y^1) - d^2_s(y^0, y^1) \right]$$

$$- \left[ (1-2a)K - 2aL \right] \cdot d^2_s(y^0, y^1) - \frac{1}{N(1-2a)} \left| V_t(y^a) - V_t(y^{1-a}) \right|^2.$$  

In the limit $a \to 0$ this yields the claim. \qed

**Acknowledgments.** The authors gratefully acknowledge support from the German Research Foundation through the Hausdorff Center for Mathematics and the Collaborative Research Center 1060 as well as support by the European Union through the ERC-AdG “RicciBounds.” They also thank the MSRI for hospitality in spring 2016 and related support by the National Science Foundation under Grant No. DMS-1440140.

**Bibliography**

[1] Ambrosio, L.; Erbar, M.; Savaré, G. Optimal transport, Cheeger energies and contractivity of dynamic transport distances in extended spaces. *Nonlinear Anal.* 137 (2016), 77–134. doi:10.1016/j.na.2015.12.006

[2] Ambrosio, L.; Gigli, N.; Mondino, A.; Rajala, T. Riemannian Ricci curvature lower bounds in metric measure spaces with $\sigma$-finite measure. *Trans. Amer. Math. Soc.* 367 (2015), no. 7, 4661–4701. doi:10.1090/S0002-9947-2015-06111-X

[3] Ambrosio, L.; Gigli, N.; Savaré, G. Gradient flows in metric spaces and in the space of probability measures. Lectures in Mathematics ETH Zürich. Birkhäuser, Basel, 2005.

[4] Ambrosio, L.; Gigli, N.; Savaré, G. Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below. *Invent. Math.* 195 (2014), no. 2, 289–391. doi:10.1007/s00222-013-0515-1

[5] Ambrosio, L.; Gigli, N.; Savaré, G. Metric measure spaces with Riemannian Ricci curvature bounded from below. *Duke Math. J.* 163 (2014), no. 7, 1405–1490. doi:10.1215/00127094-2681605
[6] Ambrosio, L.; Gigli, N.; Savaré, G. Bakry-Émery curvature-dimension condition and Riemannian Ricci curvature bounds. *Ann. Probab.* **43** (2015), no. 1, 339–404. [doi:10.1214/14-AOP907]

[7] Ambrosio, L.; Mondino, A.; Savaré, G. Nonlinear diffusion equations and curvature conditions in metric spaces. Preprint, 2015. [arXiv:1509.07273 [math.AP]]

[8] Arnaudon, M.; Coulibaly, K. A.; Thalmaier, A. Horizontal diffusion in $C^1$ path space. *Séminaire de Probabilités XLIII*, 73–94. Lecture Notes in Mathematics, 2006. Springer, Berlin, 2011. [doi:10.1007/978-3-642-15217-7_2]

[9] Arnaudon, M.; Coulibaly, K. A.; Thalmaier, A. Brownian motion with respect to a metric depending on time: definition, existence and applications to Ricci flow. *C. R. Math. Acad. Sci. Paris* **346** (2008), no. 13–14, 773–778. [doi:10.1016/j.crma.2008.05.004]

[10] Bacher, K.; Sturm, K.-T. Ricci bounds for Euclidean and spherical cones. *Singular phenomena and scaling in mathematical models*, 3–23. Springer, Cham, 2014. [doi:10.1007/978-3-319-00786-1_1]

[11] Bakry, D.; Gentil, I.; Ledoux, M. *Analysis and geometry of Markov diffusion operators*. Grundlehren der mathematischen Wissenschaften, 348. Springer, Cham, 2014. [doi:10.1007/978-3-319-00227-9]

[12] Bolley, F.; Gentil, I.; Guillin, A.; Kuwada, K. Equivalence between dimensional contractions in Wasserstein distance and the curvature-dimension condition. Preprint, 2015. [arXiv:1510.07793 [math.PR]]

[13] Cao, H.-D.; Zhu, X.-P. A complete proof of the Poincaré and geometrization conjectures—application of the Hamilton-Perelman theory of the Ricci flow. *Asian J. Math.* **10** (2006), no. 2, 165–492. [doi:10.4310/AJM.2006.v10.n2.a2]

[14] Cheeger, J. Differentiability of Lipschitz functions on metric measure spaces. *Geom. Funct. Anal.* **9** (1999), no. 3, 428–517. [doi:10.1007/s000390050004]

[15] Chen, Z.-Q.; Fukushima, M. *Symmetric Markov processes, time change, and boundary theory*. London Mathematical Society Monographs Series, 35. Princeton University Press, Princeton, N.J., 2012.

[16] Coulibaly-Pasquier, K. A. Brownian motion with respect to time-changing Riemannian metrics, applications to Ricci flow. *Ann. Inst. Henri Poincaré Probab. Stat.* **47** (2011), no. 2, 515–538. [doi:10.1214/10-AIHP364]

[17] Erbar, M.; Kuwada, K.; Sturm, K.-T. On the equivalence of the entropic curvature-dimension condition and Bochner’s inequality on metric measure spaces. *Invent. Math.* **201** (2015), no. 3, 993–1071. [doi:10.1007/s00222-014-0563-7]

[18] Erbar, M.; Kuwada, K.; Sturm, K.-T. Rigidity for Ricci flat metric cones. In preparation.

[19] Georgiou, T.; Sander, C.; Tannenbaum, A.; Zhu, L. Ricci curvature and robustness of cancer networks. Preprint, 2015. [arXiv:1502.04512 [q-bio.MN]]

[20] Gigli, N. Nonsmooth differential geometry-An approach tailored for spaces with Ricci curvature bounded from below. Preprint, 2014. [arXiv:1407.0809 [math.DG]]

[21] Hajłasz, P. Sobolev spaces on an arbitrary metric space. *Potential Anal.* **5** (1996), no. 4, 403–415. [doi:10.1007/BF00275475]

[22] Hamilton, R. S. Three-manifolds with positive Ricci curvature. *J. Differential Geom.* **17** (1982), no. 2, 255–306.

[23] Hamilton, R. S. The formation of singularities in the Ricci flow. *Surveys in differential geometry, Vol. II* (Cambridge, MA, 1993), 7–136. International Press, Cambridge, Mass., 1995.

[24] Haslhofer, R.; Naber, A. Weak solutions for the Ricci flow I. Preprint, 2015. [arXiv:1504.00911 [math.DG]]

[25] Ketterer, C. Cones over metric measure spaces and the maximal diameter theorem. *J. Math. Pures Appl.* (9) **103** (2015), no. 5, 1228–1275. [doi:10.1016/j.matpur.2014.10.011]

[26] Kleiner, B.; Lott, J. Notes on Perelman’s papers. *Geom. Topol.* **12** (2008), no. 5, 2587–2855. [doi:10.2140/gt.2008.12.2587]

[27] Kleiner, B.; Lott, J. Singular Ricci flows I. *Acta Math.* **219** (2017), no. 1, 65–134.
28. Kopfer, E. Gradient flow for the Boltzmann entropy and Cheeger’s energy on time-dependent metric measure spaces. *Calc. Var. Partial Differential Equations* **57** (2018), no. 1, Art. 20, 40 pp.

29. Kopfer, E. Super-Ricci flows and improved gradient and transport estimates. Preprint, 2017. [arXiv:1704.04177 [math.PR]]

30. Kopfer, E.; Sturm, K.-T. Super-Ricci flows and functional inequalities. In preparation.

31. Kuwada, K. Space-time Wasserstein controls and Bakry-Ledoux type gradient estimates. *Calc. Var. Partial Differential Equations* **54** (2015), no. 1, 127–161. [DOI:10.1007/s00526-014-0781-2]

32. Kuwada, K.; Philipowski, R. Coupling of Brownian motions and Perelman’s $\mathcal{L}$-functional. *J. Funct. Anal.* **260** (2011), no. 9, 2742–2766. [DOI:10.1016/j.jfa.2011.01.017]

33. Kuwada, K.; Philipowski, R. Non-explosion of diffusion processes on manifolds with time-dependent metric. *Math. Z.* **268** (2011), no. 3-4, 979–991. [DOI:10.1007/s00209-010-0704-7]

34. Lakzian, S.; Munn, M. Super Ricci flow for disjoint unions. Preprint, 2012. [arXiv:1211.2792 [math.DG]]

35. Lin, S.; Li, X.-D. The $W$-entropy formula for the Witten Laplacian on manifolds with time dependent metrics and potentials. *Pacific J. Math.* **278** (2015), no. 1, 173–199. [DOI:10.2140/pjm.2015.278.173]

36. Lierl, J.; Saloff-Coste, L. Parabolic Harnack inequality for time-dependent non-symmetric Dirichlet forms. Preprint, 2012. [arXiv:1205.6493 [math.PR]]

37. Lions, J.-L.; Magenes, E. *Non-homogeneous boundary value problems and applications. Vol. I.* Die Grundlehren der mathematischen Wissenschaften, 181. Springer, New York–Heidelberg, 1972.

38. Lott, J.; Villani, C. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math. (2)* **169** (2009), no. 3, 903–991. [DOI:10.4007/annals.2009.169.903]

39. McCann, R. J.; Topping, P. M. Ricci flow, entropy and optimal transportation. *Amer. J. Math.* **132** (2010), no. 3, 711–730. [DOI:10.1353/ajm.0.0110]

40. Morgan, J.; Tian, G. *Ricci flow and the Poincaré conjecture.* Clay Mathematics Monographs, 3. American Mathematical Society, Providence, R.I.; Clay Mathematics Institute, Cambridge, Mass., 2007.

41. Ollivier, Y. Ricci curvature of Markov chains on metric spaces. *J. Funct. Anal.* **256** (2009), no. 3, 810–864. [DOI:10.1016/j.jfa.2008.11.001]

42. Perelman, G. The entropy formula for the Ricci flow and its geometric applications. Preprint, 2002. [arXiv:math/0211159 [math.DG]]

43. Perelman, G. Finite extinction time for the solutions to the Ricci flow on certain three-manifolds. Preprint, 2003. [arXiv:math/0307245 [math.DG]]

44. Perelman, G. Ricci flow with surgery on three-manifolds. Preprint, 2003. [arXiv:math/0303109 [math.DG]]

45. Renardy, M.; Rogers, R. C. *An introduction to partial differential equations.* Second edition. Texts in Applied Mathematics, 13. Springer, New York, 2004.

46. Sandhu, R.; Georgiou, T.; Tannenbaum, A. Market fragility, systemic risk, and Ricci curvature. Preprint, 2015. [arXiv:1505.05182 [q-fin.RM]]

47. Shanmugalingam, N. Newtonian spaces: an extension of Sobolev spaces to metric measure spaces. *Rev. Mat. Iberoamericana* **16** (2000), no. 2, 243–279. [DOI:10.4171/RMI/275]

48. Sturm, K.-T. Analysis on local Dirichlet spaces. II. Upper Gaussian estimates for the fundamental solutions of parabolic equations. *Osaka J. Math.* **32** (1995), no. 2, 275–312.

49. Sturm, K. T. Diffusion processes and heat kernels on metric spaces. *Ann. Probab.* **26** (1998), no. 1, 1–55. [DOI:10.1214/aop/1022855410]

50. Sturm, K.-T. On the geometry of metric measure spaces. I. *Acta Math.* **169** (2002), no. 1, 65–131.

51. Sturm, K.-T. Super Ricci flows for metric measure spaces. I. Preprint, 2016. [arXiv:1603.02193 [math.DG]]
[52] Sturm, K.-T. Remarks about synthetic upper Ricci bounds for metric measure spaces. Preprint, 2017. [arXiv:1711.01707] [math.DG]

[53] Topping, P. Ricci flow: the foundations via optimal transportation. Optimal transportation, 72–99. London Mathematical Society Lecture Note Series, 413. Cambridge University Press, Cambridge, 2014. [doi:10.1017/CBO9781107297296.006]

EVA KOPFER
Universität Bonn
Endenicher Allee 60
53115 Bonn
GERMANY
E-mail: eva.kopfer@iam.uni-bonn.de

KARL-THEODOR STURM
Universität Bonn
Endenicher Allee 60
53115 Bonn
GERMANY
E-mail: sturm@uni-bonn.de

Received January 2017.