HEAT FLOW ON TIME-DEPENDENT METRIC MEASURE SPACES AND SUPER-RICCI FLOWS

EVA KOPFER, KARL-THEODOR STURM

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 which 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)
- ▷ 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.

Contents

The authors gratefully acknowledges support by 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.

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
(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 Riemannian manifolds. It is intimately linked to other important objects like Dirichlet 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 informations on the underlying space, like curvature, genus, index etc.

Within the last decades, the heat flow was also successfully studied on more general spaces, in particular, on metric measure spaces [14, 21, 47, 49]. The foundational work of Ambrosio, Gigli and Savaré [4, 5, 6] clarified the picture, allowed to unify various of the previous approaches, and made clear that for each metric measure space (X, d, m) with $\int \exp\big(-Cd^2(x,z)\big)dm(x) < \infty$ (for some C, z) there exists a unique solution to the heat equation, most conveniently defined as gradient flow in $L^2(X, m)$ for the Dirichlet energy ('Cheeger energy') $\mathcal{E}(u) = \int_X |\nabla u|^2\,dm$.

Synthetic lower Ricci bounds. The heat flow on Riemannian manifolds – and more generally on metric measure 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 $S(um) = \int u \log u\, dm$ 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 $|\nabla P_t u|^2 \le P_t |\nabla u|^2$
- ▷ the existence of coupled pairs of Brownian motions with $d(X_t, Y_t) \le d(X_0, Y_0)$
- ▷ the transport estimate $W\big((P_t u)m, (P_t v)m\big) \le W\big(um, vm\big)$
- ▷ the convexity of the Boltzmann entropy S on the geodesic space $(\mathcal{P}(X), W)$.

Indeed, in the Lott-Stum-Villani approach to synthetic lower Ricci bounds [50, 38] 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 perfectly make sense in discrete settings, cf. Ollivier [41], Tannenbaum et al. [19], Sandhu et al. [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 question 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 $\mathrm{CD}(0, \infty)$-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, non-degenerate g_t – question (a) allows for an easy, affirmative answer. Surprisingly enough, Brownian motion was constructed only recently [8, 16]. Question (b) was unsolved so far. McCann/Topping 2010 [39], Arnaudon/Coulibaly/Thalmaier [9], and Haslhofer/Naber [24] proved that the first three questions in (c) have one common answer:

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

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 $(\mathcal{P}(X), W_t)$.

The current paper, regarded as 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 $\mathcal{E}_t$ in $L^2(X, m_t)$, the latter as the backward gradient flow for the Boltzmann entropy S_t in $(\mathcal{P}(X), W_t)$. 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 which led to the formulation of the so-called curvature-dimension condition $\mathrm{CD}(K, N)$. The time-dependent counterpart to the $\mathrm{CD}(K, N)$-condition will be so-called *super-(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 worth 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 thus will 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 $\mathrm{Ric}_{g_t} = -\frac{1}{2}\partial_t g_t$ for all $t \in I$. It is called *super-Ricci flow* if instead only $\mathrm{Ric}_{g_t} \geq -\frac{1}{2}\partial_t g_t$ 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 'super-solutions' 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, 43], see also [13, 26, 40], Ricci flow has attracted 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 spacetimes) is defined and analyzed in detail, allowing also for Ricci flows through singularities.

Compared to Ricci flows, super-Ricci flows allow for a 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 $(M, g_t, e^{-f_t} dvol_{g_t})_t$, for instance, is a super-n-Ricci flow if and only if g_t satisfies (1) and if f_t is constant for each t, see Theorem 2.9 in [51].

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. Work in progress deals with synthetic upper Ricci bounds [52] which – in combination with the former – then also will 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 which 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, Definition 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 $\mathrm{Ric} = -Kg$ for some $K > 0$ depicted in Figure 1. Under the evolution of a Ricci flow the curvature of the surface where $\mathrm{Ric} = -Kg$ will increase, while the curvature of the "rims" ($\mathrm{Ric} = +\infty$) will decrease. In this sense the region of negative curvature will inflate, while the edges will smooth out. Under the evolution of a super-Ricci flow the surface inflates as well but it may keep the edges – or it may start to smoothen them at any later time or with smaller speed.

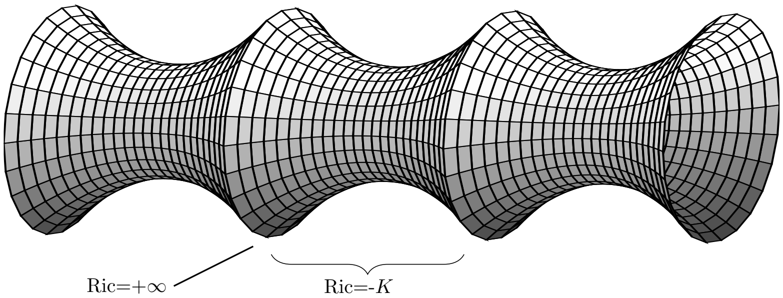

FIGURE 1. Surface of revolution of a piecewise hyperbolic space

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

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

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 put

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

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 $t \le t_*$ will be identified with a 1-point space (and m_t with the Dirac mass in this point), see also Firgure 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 be no super-N-Ricci flow for $N < n$.

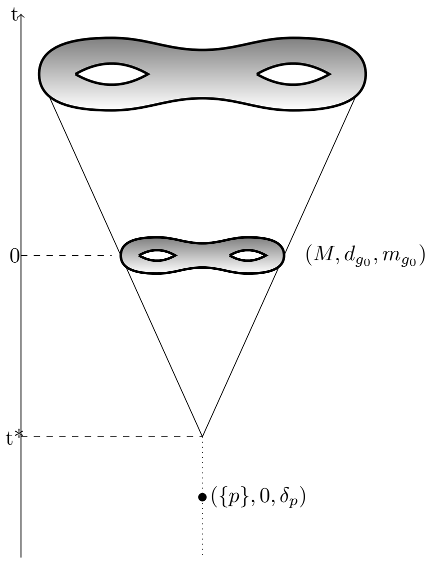

FIGURE 2. Point exploding to a hyperbolic quotient

More generally, consider $(\overline{M}, \overline{g}_t) = (M' \times M, g' \otimes g_t)$ 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 which coincides with the given mm-space for $t \le t_*$ 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 $\mathcal{S} := M \times \{0\}$ and $\mathcal{N} := M \times \{\pi\}$ to a point, the 'south' and the 'north pole', respectively. The resulting space is called *spherical 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 where $(x, s), (x', s') \in M \times [0, \pi]$. Since M is a RCD*(3, 4) space, the cone of it is a RCD*(4, 5) space [25].

The punctured cone $\Sigma_0 := \Sigma(M) \setminus \{\mathcal{S}, \mathcal{N}\}$ is an incomplete 5-dimensional Riemannian manifold. Let g_0 denote the metric tensor of Σ_0. The curvature of the punctured cone can be calculated explicitly and is given by $\mathrm{Ric}(g_0) = 4g_0$. Then $g(t) := (1 - 8t)g_0$ defines a solution to the Ricci flow $\mathrm{Ric}(g_t) = -\frac{1}{2}\partial_t g_t$ with $g(0) = g_0$, which collapses to a point at time $T = \frac{1}{8}$.

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 $\mu_0, \mu_1 \in$ *Dom*(S_t) on $\Sigma(M)$ be given. Let $(\mu_a)_{a\in[0,1]}$ be a W_t-geodesic connecting μ_0, μ_1. Then, $\mu_a = (e_a)_*\nu$, where ν is an optimal path measure, i.e. a probability measure on the d_t-geodesics $\Gamma(\Sigma(M))$ of $\Sigma(M)$ such that $(e_0, e_1)_*\nu$ is an optimal coupling of $(e_0)_*\nu = \mu_0, (e_1)_*\nu = \mu_1$, where $e_a \colon \Gamma(\Sigma(M)) \to \Sigma(M)$ denotes the evaluation map. According to Theorem 3.3 in [10] every optimal path measure ν will give no mass to d_t-geodesics through the poles. Hence we can omit the d_t-geodesics through the poles without changing the W_t-geodesics. Since the punctured cone $(\Sigma_0, g_t)_{t\in I}$ is a Ricci flow, and in particular a super-Ricci flow in the sense of Definition 2.4 in [51], the metric measure space $(\Sigma(M), d_{\Sigma(M)}(t), \hat{m}_t)_{t\in I}$ is a super-Ricci flow as well.

Let us emphasize that for each $t \in [0, 1/8)$ the sectional curvature of the punctured spherical cone Σ_0 is neither bounded from below nor from above. Indeed, for $x, y \in S^2(1/\sqrt{3})$ and $0 < r < \pi$ an orthonormal basis of the tangent space $T_{(x,y,r)}\Sigma_0$ is given by $\{\hat{u}_1, \hat{u}_2, \hat{v}_1, \hat{v}_2, \hat{w}\}$ where $\hat{u}_i = \frac{1}{\sin r}(u_i, 0, 0)$, $\hat{v}_i = \frac{1}{\sin r}(0, v_i, 0)$, $\hat{w} = (0, 0, 1)$ and u_1, u_2 is an orthonormal basis of $T_x(S^2(1/\sqrt{3}))$ and v_1, v_2 is an orthonormal basis of $T_y(S^2(1/\sqrt{3}))$. Then for the sectional curvature we find

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

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

Let us 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.**

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

$$\left|\log \frac{d_t(x, y)}{d_s(x, y)}\right| \le L \cdot |t - s| \tag{2}$$

('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* Δ_t on $L^2(X, m_t)$ characterized by $-\int_X \Delta_t u\, v\, dm_t = \mathcal{E}_t(u, v)$ where the Dirichlet energy

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

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

Heat equation. Our first important result concerns existence and uniqueness for solutions 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 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

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

with $u_s = h$;

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

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

is the unique solution to the adjoint heat equation

$$\partial_s v_s = -\Delta_s v_s + \dot f_s \cdot v_s \qquad on\ (0, t) \times X$$

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

Many properties which 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 $Dom(\Delta_t)$ 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 $\tau \le T$ and all $w \in Dom(\mathcal{E})$*

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

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

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

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

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

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

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

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

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

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

Characterization of super-Ricci flows. In [51], the second author has introduced and analyzed the notion of *super-Ricci flows* for time-dependent metric measure $(X, d_t, m_t)_{t\in I}$. The defining property of the latter is the so-called dynamic convexity of the Boltzmann entropy $S : I \times \mathcal{P} \to (-\infty, \infty]$ with

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

and $S_t(\mu) = \infty$ if $\mu \not\ll m_t$. Here $\mathcal{P} = \mathcal{P}(X)$ will denote the space of probability measures on X, equipped with time-dependent Kantorovich-Wasserstein distances W_t induced by d_t, $t \in I$. 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: $\mathrm{Ric}_{g_t} + \frac{1}{2}\partial_t g_t \geq 0$.

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, Savaré [6].

Theorem 1.7. *The following assertions are equivalent:*

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

$$\partial_a^+ S_t(\mu^a)\big|_{a=1-} - \partial_a^- S_t(\mu^a)\big|_{a=0+} \geq -\frac{1}{2}\partial_t^- W_{t-}^2(\mu^0,\mu^1) \tag{3}$$

('dynamic convexity').

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

$$W_s(\hat{P}_{t,s}\mu, \hat{P}_{t,s}\nu) \leq W_t(\mu,\nu) \tag{4}$$

('transport estimate').

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

$$\big|\nabla_t(P_{t,s}u)\big|^2 \leq P_{t,s}\big(|\nabla_s u|^2\big) \tag{5}$$

('gradient estimate').

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

$$\mathbf{\Gamma}_{2,r}(u_r)(g_r) \geq \frac{1}{2}\int \overset{\bullet}{\Gamma}_r (u_r) g_r dm_r \tag{6}$$

('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:

$$u_r \in \mathrm{Lip}(X) \textit{ for all } r \in (s,t) \textit{ with } \sup_{r,x} \mathrm{lip}_r u_r(x) < \infty. \tag{7}$$

Here and in the sequel

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

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

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

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

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

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

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

Characterization of super-N-Ricci flows. For static metric measure spaces, it turned out that many powerful applications of synthetic lower bounds on the Ricci curvature are available only in combination with some synthetic upper bound on the dimension. This led to the so-called curvature-dimension condition 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 $\mu^0,\mu^1 \in Dom(S)$*

$$\partial_a^+ S_t(\mu^a)\big|_{a=1-} - \partial_a^- S_t(\mu^a)\big|_{a=0+} \geq -\frac{1}{2}\partial_t^- W_{t-}^2(\mu^0,\mu^1) + \frac{1}{N}\big|S_t(\mu^0) - S_t(\mu^1)\big|^2. \tag{8}$$

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

$$W_s^2(\hat{P}_{t,s}\mu, \hat{P}_{t,s}\nu) \leq W_t^2(\mu,\nu) - \frac{2}{N}\int_s^t \Big[S_r(\hat{P}_{t,r}\mu) - S_r(\hat{P}_{t,r}\nu)\Big]^2 dr. \tag{9}$$

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

$$\big|\nabla_t(P_{t,s}u)\big|^2 \leq P_{t,s}\big(|\nabla_s(u)|^2\big) - \frac{2}{N}\int_s^t \Big(P_{t,r}\Delta_r P_{r,s}u\Big)^2 dr. \tag{10}$$

(IV$_N$) *For all $0 < s < t < T$ and for all $u_s, g_t \in \mathcal{F}$ with $g_t \geq 0$, $g_t \in L^\infty$, $u_s \in \mathrm{Lip}(X)$ the regularity assumption (7) is satisfied and for a.e. $r \in (s,t)$*

$$\mathbf{\Gamma}_{2,r}(u_r)(g_r) \geq \frac{1}{2}\int \overset{\bullet}{\Gamma}_r(u_r) g_r dm_r + \frac{1}{N}\Big(\int \Delta_r u_r g_r dm_r\Big)^2 \tag{11}$$

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

Remark 1.10. a. In **(I$_N$)**, the requested property for a.e. t will imply that it holds true for all $t \in (0,T)$.

b. The transport estimate $(\mathbf{II}_N)$ implies the 'stronger' property

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

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

c. Under slightly more restrictive assumptions on (X,d_t,m_t) – namely, C^1-dependence of $t\mapsto\log d_t$ instead of Lipschitz continuity – in subsequent work of the first author [29] a refined version of the dynamic Bochner inequality $(\mathbf{IV}_N)$ will be deduced with estimate (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, resp.

d. Note that the regularity assumption (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 (2) and the RCD$(-K,\infty)$-condition for the static mm-spaces (X,d_t,m_t).

Super-(K,N)-Ricci flows. A more general version of the previous Theorem will deal with the equivalences to *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 time and space.

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

$$\begin{aligned}\partial_a^+S_t(\mu^a)\big|_{a=1-}-\partial_a^-S_t(\mu^a)\big|_{a=0+}\ \ \ge\ & -\frac12\partial_t^-W_{t-}^2(\mu^0,\mu^1)+\frac1N\big|S_t(\mu^0)-S_t(\mu^1)\big|^2\\ &+KW_t^2(\mu^0,\mu^1).\end{aligned}\tag{12}$$

Then for each $C\in\mathbb R$ the time-dependent mm-space $(X,\tilde d_t,\tilde m_t)_{t\in\tilde I}$ is a super-N-Ricci flow if we put

$$\tilde d_t=e^{-K\tau(t)}d_{\tau(t)},\qquad \tilde m_t=m_{\tau(t)},\qquad \tau(t)=\frac{-1}{2K}\log(C-2Kt)$$

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

Proof. Put $\tilde d=e^{-K\tau(t)}d_{\tau(t)}$. Then every $\tilde W_t$-geodesic will be a $W_{\tau(t)}$-geodesic. Therefore, the transformation $d\mapsto\tilde d$ will not change the term $\frac1N\big|S_t(\mu^0)-S_t(\mu^1)\big|^2$ nor the term $\partial_a^+S_t(\mu^a)\big|_{a=1-}-\partial_a^-S_t(\mu^a)\big|_{a=0+}$ in (12). Moreover,

$$\begin{aligned}\frac12\partial_t^-\tilde W_{t-}^2(\mu^0,\mu^1)&=e^{-2K\tau(t)}\Big[-K\partial_t\tau(t)\cdot W_{\tau(t)}+\big(\partial_t^-W_.\big)\big(\tau(t)-\big)\cdot\partial_t\tau(t)\Big]\cdot W_{\tau(t)}\\ &=e^{-2K\tau(t)}\cdot\partial_t\tau(t)\cdot\Big[-K\cdot W_.^2+\frac12\partial_t^-W_.^2\Big]\big(\tau(t)-\big)\\ &=\Big[-K\cdot W_.^2+\frac12\partial_t^-W_.^2\Big]\big(\tau(t)-\big).\end{aligned}$$

Thus (12) implies

$$\begin{aligned}\frac12\partial_t^-\tilde W_{t-}^2(\mu^0,\mu^1)&=\Big[-K\cdot W_.^2+\frac12\partial_t^-W_.^2\Big]\big(\tau(t)-\big)\\ &\ge-\partial_a^+S_{\tau(t)}(\mu^a)\big|_{a=1-}+\partial_a^-S_{\tau(t)}(\mu^a)\big|_{a=0+}+\frac1N\big|S_{\tau(t)}(\mu^0)-S_{\tau(t)}(\mu^1)\big|^2\end{aligned}$$

which proves the dynamic N-convexity of $\tilde S$ and thus the super-N-Ricci flow property of $(X,\tilde d_t,\tilde m_t)_{t\in\tilde I}$. □

Corollary 1.12. *For each $N\in(0,\infty)$ and $K\in\mathbb R$ the following are equivalent:*

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

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

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

$$e^{-2Ks}W_s^2(\hat{P}_{t,s}\mu, \hat{P}_{t,s}\nu) \leq e^{-2Kt}W_t^2(\mu,\nu) - \frac{2}{N}\int_s^t e^{-2Kr}\Big[S_r(\hat{P}_{t,r}\mu) - S_r(\hat{P}_{t,r}\nu)\Big]^2 dr. \tag{14}$$

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

$$e^{2Kt}\big|\nabla_t(P_{t,s}u)\big|^2 \leq e^{2Ks}P_{t,s}\big(|\nabla_s(u)|^2\big) - \frac{2}{N}\int_s^t e^{2Kr}\Big(P_{t,r}\Delta_r P_{r,s}u\Big)^2 dr. \tag{15}$$

(IV$_{K,N}$) *For all $0 < s < t < T$ and for all $u_s, g_t \in \mathcal{F}$ with $g_t \geq 0$, $g_t \in L^\infty$, $u_s \in \mathrm{Lip}(X)$ the regularity assumption (7) is satisfied and for a.e. $r \in (s,t)$*

$$\mathbf{\Gamma}_{2,r}(u_r)(g_r) \geq \frac{1}{2}\int \overset{\bullet}{\Gamma}_r(u_r)g_r dm_r + K\int \Gamma_r(u_r)g_r dm_r + \frac{1}{N}\Big(\int \Delta_r u_r g_r dm_r\Big)^2 \tag{16}$$

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 $(X, \tilde{d}_t, \tilde{m}_t)_{t\in\tilde{I}}$ with $\tilde{d}_t = e^{-K\tau(t)}d_{\tau(t)}$, $\tilde{m}_t = m_{\tau(t)}$ and $\tilde{I} = \{\tau(t) : t \in I, 2Kt < C\}$ where $\tau(t) = \frac{-1}{2K}\log(C - 2Kt)$. Then

$$\tilde{W}_t^2 = e^{-2K\tau}W_\tau^2, \quad \tilde{\Gamma}_t = e^{2K\tau}\Gamma_\tau, \quad \tilde{\Delta}_t = e^{2K\tau}\Delta_\tau, \quad \Gamma_{2,t} = e^{2K\tau}\Gamma_{2,\tau}, \quad \dot{\tau}_t = e^{2K\tau}.$$

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

$$e^{-2K\tau}W_\tau^2 - e^{-2K\sigma}W_\sigma^2 = \tilde{W}_t^2 - \tilde{W}_s^2$$

for $\tau = \tau(t)$ and $\sigma = \tau(s)$ and

$$\frac{2}{N}\int_s^t \Big[\tilde{S}_r(\hat{\tilde{P}}_{t,r}\mu) - S_r(\hat{\tilde{P}}_{t,r}\nu)\Big]^2 dr \quad = \quad \frac{2}{N}\int_\sigma^\tau e^{-2Kr}\Big[S_r(\hat{P}_{t,r}\mu) - S_r(\hat{P}_{t,r}\nu)\Big]^2 dr.$$

□

Discussion of standing assumptions. Let us briefly comment on the assumptions which 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.

Remark 1.13. Given a time-dependent mm-space $(X, d_t, m_t)_{t\in I}$ which satisfies all the assumptions mentioned in the beginning of this chapter but no Riemannian curvature-dimension condition is requested. Instead of that, each static mm-space (X, d_t, m_t) is merely assumed to be infinitesimally Hilbertian and S_t is requested to be absolutely continuous along W_t-geodesics.

Then assertion **(I$_N$)** of the Main Theorem 1.9 implies that for a.e. $t \in I$ the static space

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

Proof. **(I$_N$)** together with the log-Lipschitz bound (2) implies that along all W_t-geodesics

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

In combination with the absolute continuity of $a \mapsto S_t(\mu^a)$ this yields the RCD$^*(-L,N)$-condition, cf. [51]. □

Next, we will discuss the assumption (2) concerning log-Lipschtiz continuity of $t \mapsto d_t$.

Remark 1.14. Let $(M,g_t)_t$ be a time-dependent Riemannian manifold and let $(X,d_t,m_t)_t$ be the induced time-dependent mm-space.

(i) Then for any $L_1,L_2\in[-\infty,\infty]$

$$L_1\le \frac{1}{t-s}\log\frac{d_t}{d_s}\le L_2 \quad\Longleftrightarrow\quad L_1 g_t\le \frac12\partial_t g_t\le L_2 g_t.$$

Moreover, if $(M,g_t)_t$ evolves as Ricci flow then the previous assertions are equivalent to

$$-L_2 g_t\le \mathrm{Ric}_{g_t}\le -L_1 g_t. \tag{17}$$

If $(M,g_t)_t$ is a super-Ricci flow then instead we merely have the implications

$$\frac{1}{t-s}\log\frac{d_t}{d_s}\le L_2 \quad\Longrightarrow\quad -L_2 g_t\le \mathrm{Ric}_{g_t}$$

and

$$L_1\le \frac{1}{t-s}\log\frac{d_t}{d_s} \quad\Longleftarrow\quad \mathrm{Ric}_{g_t}\le -L_1 g_t.$$

The *proof* is obvious. Similar assertions holds 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 $f_t-f_s=\int_s^t \mathrm{scal}_{g_r}dr$. Thus

$$L_1\le \frac{1}{t-s}\log\frac{dm_t}{dm_s}\le L_2 \quad\Longleftrightarrow\quad -L_2\le \mathrm{scal}_{g_t}\le -L_1.$$

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) The condition (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 which characterize solutions to $\mathrm{Ric}+\frac12\partial_t g_t\ge 0$ 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 which 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 [8], 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 $(\mathbf{III}_N)\Rightarrow(\mathbf{II}_N)$. Even though we benefited very much from the powerful, detailed calculus on mm-spaces developed in [5, 6, 4] 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 $(\mathbf{II}_N)\Rightarrow(\mathbf{III}_N)$, 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 which among others will address the following challenges:

- time-discrete gradient flow scheme à Jordan-Kinderlehrer-Otto for the heat equation and its dual as gradient flows of energy and entropy, resp. [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 ∂_t as a short hand notation for $\frac{d}{dt}$. Moreover, we put $\partial_t^+ u(t) = \limsup_{s\to t} \frac{1}{t-s}(u(t)-u(s))$ and $\partial_t^- u(t) = \liminf_{s\to t} \frac{1}{t-s}(u(t)-u(s))$.

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

1.4. Sketch of the Argumentation for the Main Result.

The structure of the proof of Theorem 1.9 is as follows. In Chapter 4, we present the implications $(\mathbf{I}_N) \Longrightarrow (\mathbf{II}_N)$ and $(\mathbf{III}_N) \Longrightarrow (\mathbf{II}_N)$ as well as the converse of the latter in the case $N=\infty$. Chapter 5 is devoted to the proof of the equivalence $(\mathbf{III}_N) \Longleftrightarrow (\mathbf{IV}_N)$ as well as to the proof of the implication $(\mathbf{II}_N) \Longrightarrow (\mathbf{IV}_N)$.

In Chapter 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 (Chapter 7) which is devoted to the study of dynamical EVI-flows in a general framework. Here in particular, it will also be shown that $(\mathbf{III}_N)$ **&** $\mathbf{EVI}^- \Longrightarrow (\mathbf{I}_N)$. □

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

$(\mathbf{I}_N) \Longrightarrow (\mathbf{II}_N)$. Given two solutions to the dual heat flow $(\mu_r)_r$ and $(\nu_r)_r$, for fixed t we connect the measures $\mu_t = u m_t$ and $\nu_t = v m_t$ by a W_t-geodesic $(\eta^a)_{a\in[0,1]}$ and we choose a pair of functions ϕ, ψ in duality w.r.t. $\frac{1}{2}W_t^2$ and optimal for the pair μ_t, ν_t ('Kantorovich potentials'), see Figure 3. (Note that in the smooth Riemannian setting the W_t-geodesic and the Kantorovich potentials are linked through the relation $\eta^a = \big(\exp(-a\,\nabla\phi)\big)_*\mu_t = \big(\exp(-(1-a)\,\nabla\psi)\big)_*\nu_t$.)

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

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

Additing up these estimates yields $\frac{1}{2}\partial_r^- W_t^2(\mu_r,\nu_r)|_{r=t+} + \frac{1}{2}\partial_r^- W_r^2(\mu_t,\nu_t)\big|_{r=t-} \geq \frac{1}{N}\big[S_t(\mu_t) - S_t(\nu_t)\big]^2$. A careful time shift argument allows to replace the left hand side by $\frac{1}{2}\partial_{t+}^- W_t^2(\mu_t,\nu_t)$ which then proves the claim.

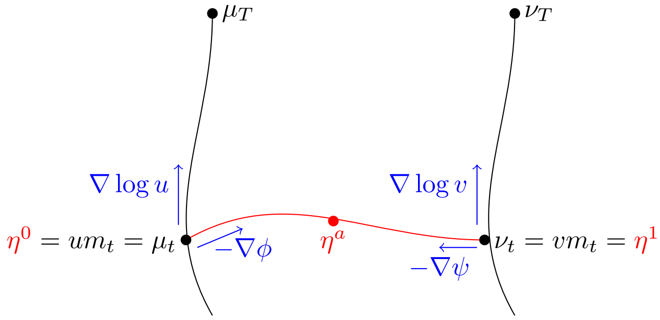

FIGURE 3.

(II$_N$) $\Longrightarrow$ (IV$_N$). Given a Lipschitz function u and a probability density g (w.r.t. m_τ) put $g_r = P^*_{\tau,r} g$, $u_r = P_{r,\sigma} u$ and $h_r := \int g_r \Gamma_r(u_r) dm_r$ for $0 < \sigma < r < \tau < T$.

By duality we already know that the transport estimate **(II$_N$)** implies the infinite-dimensional gradient estimate **(III)** which helps us to deduce that

$$h_\tau - h_\sigma \geq \int_\sigma^\tau \Big[-2\mathbf{\Gamma}_{2,r}(u_r)(g_r) + \int \overset{\bullet}{\Gamma}_r(u_r)\, g_r\, m_r \Big]\, dr.$$

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

$$g_\tau^{\sigma,a} := g_\tau\Big(1 - a[\Delta_\tau u_\sigma + \Gamma_\tau(\log g_\tau, u_\sigma)]\Big)$$

for small $a > 0$. 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 **(II$_N$)** applied to the probability measures $g_\tau m_\tau$ and $g_\tau^{\sigma,a} m_\tau$ gives us for all $a > 0$

$$\begin{aligned} W_\sigma^2(\hat{P}_{\tau,s}(g_\tau m_\tau), \hat{P}_{\tau,\sigma}(g_\tau^{\sigma,a} m_\tau)) \quad &- \quad W_\tau^2(g_\tau m_\tau, g_\tau^{\sigma,a} m_\tau) \\ &\leq \quad -\frac{2}{N}\int_\sigma^\tau [S_r(\hat{P}_{\tau,r}(g_\tau m_\tau)) - S_r(\hat{P}_{\tau,r}(g_\tau^{\sigma,a} m_\tau))]^2 dr. \end{aligned}$$

In the limit $a \searrow 0$ we eventually end up with

$$h_\tau - h_\sigma \leq -\frac{2}{N}\int_\sigma^\tau \Big(\int \Delta_r u_r\, g_r dm_r \Big)^2 dr.$$

Together with the previous lower estimate for $h_\tau - h_\sigma$ this proves the claim.

(IV$_N$) $\Longleftrightarrow$ (III$_N$). This is – modulo regularity issues – a simple, well-known (cf. [51], Theorem 5.5) differentiation-integration argument for the function

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

(III$_N$) $\Longrightarrow$ (II$_N$). Given any 'regular' curve $(\mu_\tau^a)_{a\in[0,1]}$ and $\tau \in I$ we will study the evolution of this curve under the dual heat flow. More precisely, we analyze the growth of the action

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

of the curve $(\mu_t^a)_{a\in[0,1]}$ for $t < \tau$ where $\mu_t^a = \hat{P}_{\tau,t}\mu_\tau^a = u_t^a m_t$ and $(\Phi_t^a)_{a\in[0,1]}$ denotes the velocity potentials in the static space (X, d_t, m_t). For $s < t$ we approximate the action $\mathcal{A}_s\big(\dot{\mu_s}\big)$ by

$$\sum_{i=k} \frac{1}{a_i - a_{i-1}} W_s^2\big(\mu_s^{a_{i-1}}, \mu_s^{a_i}\big),$$

the latter in terms of W_s-Kantorovich potentials, and finally by means of the interpolating Hopf-Lax semigroup. Applying the Bakry-Ledoux gradient estimate $(\mathbf{III}_N)$ then allows to estimate

$$\begin{aligned} 2\varepsilon + \frac{1}{t-s}\Big[\mathcal{A}_t(\dot{\mu_t}) - \mathcal{A}_s(\dot{\mu_s})\Big] &\geq \frac{2}{N+\varepsilon}\Big|\int_0^1 \int_X \nabla_t \Phi_t^a \cdot \nabla_t \log u_t^a \, d\mu_t^a da\Big|^2 \\ &= \frac{2}{N+\varepsilon}\Big|S_t(\mu_t^1) - S_t(\mu_t^0)\Big|^2 \end{aligned}$$

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 τ yields

$$\mathcal{A}_s(\dot{\mu_s}) \leq \mathcal{A}_\tau(\dot{\mu_\tau}) - \frac{2}{N}\int_s^\tau \big[S_t(\mu_t^0) - S_t(\mu_t^1)\big]^2 dt.$$

This indeed proves the claim since

$$W_\tau^2(\mu^0, \mu^1) = \inf\Big\{\mathcal{A}_\tau(\dot{\mu_\tau}) : \ (\mu_\tau^a)_{a\in[0,1]} \text{ regular curve connecting } \mu^0, \mu^1\Big\}$$

for any μ^0, μ^1 and τ whereas $W_s^2(\mu_s^0, \mu_s^1) \leq \mathcal{A}_s(\dot{\mu_s})$ for all $s < \tau$.

$(\mathbf{III}_N) \implies (\mathbf{I}_N)$. 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

$$W_t^2(\mu_t^0, \mu_t^1) = \frac{1}{a}W_t^2(\mu_t^0, \mu_t^a) + \frac{1}{1-2a}W_t^2(\mu_t^a, \mu_t^{1-a}) + \frac{1}{a}W_t^2(\mu_t^{1-a}, \mu_t^1) \tag{18}$$

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

$$W_s^2(\mu_t^0, \mu_t^1) \leq \frac{1}{a}W_s^2(\mu_t^0, \mu_s^a) + \frac{1}{1-2a}W_s^2(\mu_s^a, \mu_s^{1-a}) + \frac{1}{a}W_s^2(\mu_s^{1-a}, \mu_t^1). \tag{19}$$

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

$$\liminf_{s\nearrow t} \frac{1}{t-s}\frac{1}{1-2a}\Big[W_t^2(\mu_t^a, \mu_t^{1-a}) - W_s^2(\mu_s^a, \mu_s^{1-a})\Big] \geq \frac{2}{N}\frac{1}{1-2a}\Big[S_t(\mu_t^a) - S_t(\mu_t^{1-a})\Big]^2.$$

The EVI-property to be discussed below will allow to estimate

$$\liminf_{s\nearrow t} \frac{1}{t-s}\frac{1}{a}\Big[W_t^2(\mu_t^0, \mu_t^a) - W_s^2(\mu_t^0, \mu_s^a)\Big] \geq \frac{2}{a}\Big[S_t(\mu_t^a) - S_t(\mu_t^0)\Big] - LaW_t^2(\mu_t^0, \mu_t^1),$$

as well as

$$\liminf_{s\nearrow t} \frac{1}{t-s}\frac{1}{a}\Big[W_t^2(\mu_t^{1-a}, \mu_t^1) - W_s^2(\mu_s^{1-a}, \mu_t^1)\Big] \geq \frac{2}{a}\Big[S_t(\mu_t^{1-a}) - S_t(\mu_t^1)\Big] - LaW_t^2(\mu_t^0, \mu_t^1).$$

Using (18) together with (19) and adding up the last three inequalities we obtain after letting $a \searrow 0$ (see also Figure 4):

$$\begin{aligned} \liminf_{s\nearrow t} \frac{1}{t-s}\Big[W_t^2(\mu_t^0, \mu_t^1) - W_s^2(\mu_s^0, \mu_s^1)\Big] &\geq \frac{2}{N}\Big[S_t(\mu_t^0) - S_t(\mu_t^1)\Big]^2 \\ &\quad + 2\partial_a^- S_t(\mu_t^a)\big|_{a=0+} - 2\partial_a^+ S_t(\mu_t^a)\big|_{a=1-}. \end{aligned}$$

In order to prove the EVI-property, we follow the approach by [6] and [17] respectively 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 $\tilde{W}_{s,t}$ 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 σ and $\mu_t := \hat{P}_{\tau,t}\mu$ where $\mu, \sigma \in \mathcal{P}(X)$ are fixed. We then apply the dual heat flow $\rho_{a,\vartheta} := \hat{P}_{t,s+a(t-s)}\rho^a$ to the regular curve, cf. Figure 5, and eventually show using $(\mathbf{III})$ that

$$\frac{1}{2}\tilde{W}_{s,t}^2(\rho_{1,\vartheta}, \rho_{0,\vartheta}) - (t-s)(S_t(\rho_{1,\vartheta}) - S_s(\rho_{0,\vartheta})) \leq \int_0^1 \Big[\frac{1}{2}|\dot{\rho}^a|_t^2 + (t-s)^2 \int \dot{f}_{\vartheta(a)} d\rho_{a,\vartheta}\Big] da.$$

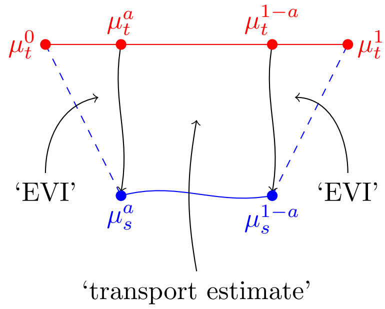

FIGURE 4.

Then, by approximation, we obtain

$$\frac{1}{2}\tilde{W}^2_{s,t}(\mu_s,\sigma) - (t-s)(S_t(\sigma) - S_s(\mu_s)) \leq \frac{1}{2}W^2_t(\mu_t,\sigma) - (t-s)^2 \int_0^1 \int \dot{f}_{\vartheta(a)} d\rho_{a,\vartheta} da.$$

In contrast to the static case we obtain the additional error term $(t-s)^2 \int_0^1 \int \dot{f}_{\vartheta(a)} d\rho_{a,\vartheta} da$ which however vanishes after dividing by $t-s$ and letting $s \nearrow t$. Thus

$$S_t(\mu_t) - S_t(\sigma) \leq \liminf_{s\nearrow t} \frac{1}{2(t-s)}\Big(W^2_t(\mu_t,\sigma) - \tilde{W}^2_{s,t}(\mu_s,\sigma)\Big) = \frac{1}{2}\partial_s^- W^2_{s,t}(\mu_s,\sigma)_{|s=t-}.$$

Note that the log-Lipschitz continuity of the distance allows to estimate the last term from above by

$$\frac{1}{2}\partial_s^- W^2_t(\mu_s,\sigma)_{|s=t-} + \frac{L}{2}W^2(\mu_t,\sigma).$$

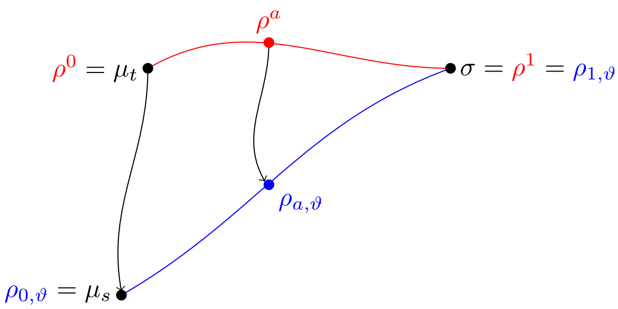

FIGURE 5.

2. The Heat Equation for Time-dependent Dirichlet Forms

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

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

$$\mathcal{E}_\diamond(\xi \circ u) \le \mathcal{E}_\diamond(u) \quad \text{for all } \xi\colon \mathbb{R} \to \mathbb{R} \text{ 1-Lipschitz such that } \xi(0) = 0.$$

Here and in the sequel, the same symbol will be used for a bilinear form and the quadratic form associated with it, i.e. $\mathcal{E}_\diamond(u) = \mathcal{E}_\diamond(u,u)$. The Dirichlet form $\mathcal{E}_\diamond$ is called strongly local if $\mathcal{E}_\diamond(u,v) = 0$ whenever $(u+c)v = 0$ $m_\diamond$-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 $0 \le s < \tau \le T$ the Hilbert spaces

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

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

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

$$\Gamma_\diamond(f_t)(x) \le C. \tag{20}$$

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

- symmetric, positive semidefinite bilinear forms Γ_t on $\mathcal{F}$, each of which has the diffusion property
$$\Gamma_t(\Psi(u_1,\dots,u_k), v) = \sum_{i=1}^k \Psi_i(u_1,\dots,u_k)\Gamma_t(u_i,v)$$
$\forall k \in \mathbb{N}, \forall v, u_1,\dots,u_k \in \mathcal{F} \cap L^\infty(X, m_\diamond), \forall \Psi \in \mathcal{C}^1(\mathbb{R}^k)$ with $\Psi(0) = 0$, [11],
- and all of them being uniformly comparable ('uniformly elliptic') w.r.t. the reference form $\Gamma_\diamond$ on $\mathcal{F}$, i.e. $\exists C$ s.t. $\forall t \in (0,T), \forall u \in \mathcal{F}, \forall x \in X$
$$\frac{1}{C}\,\Gamma_\diamond(u)(x) \le \Gamma_t(u)(x) \le C\,\Gamma_\diamond(u)(x). \tag{21}$$

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

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

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

Definition 2.1. *A function u is called solution to the heat equation*

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

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

$$-\int_s^\tau \mathcal{E}_t(u_t,w_t)dt = \int_s^\tau \langle \partial_t u_t, w_t e^{-f_t}\rangle_{\mathcal{F}^*,\mathcal{F}}\, dt \tag{22}$$

where $\langle\cdot,\cdot\rangle_{\mathcal{F}^*,\mathcal{F}} = \langle\cdot,\cdot\rangle$ denotes the dual pairing. Note that thanks to (20)*, $w \in L^2\big((s,\tau)\to\mathcal{F}\big)$ if and only if $we^{-f} \in L^2\big((s,\tau)\to\mathcal{F}\big)$.*

Since $u_t \in Dom(A_t)$ (and thus $\partial_t u_t \in L^2$) for almost every t by virtue of Theorem 2.12 we may equivalently rewrite the right hand side of the above equation as

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

which allows for a more intuitive, alternative formulation of (22) as follows:

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

Theorem 2.2. *For all $0\le s<\tau\le T$ and each $h\in\mathcal{H}$ there exists a unique solution $u\in\mathcal{F}_{(s,\tau)}$ of the heat equation on $(s,\tau)\times X$ with $u_s=h$ (or equivalently with $\lim_{t\searrow s} u_t = h$).*

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

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

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

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

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

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

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

if C is chosen such that $|\Gamma_t(u,v)|\le C\cdot\Gamma_\diamond(u)^{1/2}\cdot\Gamma_\diamond(v)^{1/2}$ for all u,v and t. Thus we may apply the general existence result for solutions to time-dependent operator equations $\partial_t u = A_t u$ on a fixed Hilbert space $\mathcal{H}$. For this, we refer to [37], Chapter III, Theorem 4.1 and Remark 4.3, see also [45], Theorem 10.3. (Note, however, that the latter assumes a continuity of $t\mapsto A_t$ in operator norm which is not really necessary.) □

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

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

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

$$\partial_t P_{t,s}h = A_t P_{t,s}h. \tag{23}$$

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

2.2. **The Adjoint Heat Equation.**

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

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

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

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

Theorem 2.5. *Assume* (20) *and*

$$|f_t(x) - f_s(x)| \le L\,|t-s|. \tag{24}$$

(i) *Given $0 \le \sigma < t \le T$, 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 $(P^*_{t,s})_{s \le t}$ of linear operators on $\mathcal{H}$ satisfying the 'adjoint propagator property'

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

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

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

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

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

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

□

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

Lemma 2.6. *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: $Dom(A_t) = Dom(A^*_t)$*

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

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

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

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

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

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

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

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

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

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

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

Proposition 2.8. (i) *For all solutions $u \ge 0$ to the heat equation and all $s < t$*

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

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

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

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

Proof. In both cases, straightforward calculations yield

$$e^{Lt} \partial_t \left[e^{-Lt} \int u_t \log u_t\, dm_t \right] \ \le \ \int (\log u_t + 1) \partial_t u_t\, dm_t = -\int \Gamma_t(\log u_t)\, u_t\, dm_t \le 0$$

and

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

□

2.3. **Energy Estimates.** Throughout this section, assume (20) as well as (24) and in addition

$$|\Gamma_t(u) - \Gamma_s(u)| \le 2L \cdot \int_s^t \Gamma_r(u) dr \tag{25}$$

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

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

$$\int_s^\tau \mathcal{E}_t(u_t)\, dt \le \frac{1}{2} \|u_s\|^2_{L^2(m_s)}. \tag{26}$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

□

Lemma 2.10. *For all $\delta > 0$ and all $u, v \in \mathcal{F}$ the map $t \mapsto \tilde{\mathcal{E}}^\delta_t(u,v)$ is absolutely continuous with*

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

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

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

Here we also used the fact that $\tilde{\mathcal{E}}_t(u^\delta_t, u^\delta_t) \nearrow \tilde{\mathcal{E}}_t(u_t, u_t)$ 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 $\delta > 0$*

$$\int_s^\tau \int_X \left| (I - \delta\tilde{A}_t)^{-1/2}\tilde{A}_t u_t \right|^2 dm_\diamond\, dt \le C \cdot \left[\mathcal{E}_s(u_s) + \|u_s\|^2_{L^2(m_s)} \right]. \tag{27}$$

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

$$\int_s^\tau \int_X \left| \tilde{A}_t u_t \right|^2 dm_\diamond\, dt \le C \cdot \left[\mathcal{E}_s(u_s) + \|u_s\|^2_{L^2(m_s)} \right]. \tag{28}$$

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

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

Here

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

according to the previous Lemma and

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

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

Theorem 2.12. *For all $0 < s < \tau < T$ and for all solutions $u \in \mathcal{F}_{(s,T)}$ to the heat equation*

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

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

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

More precisely,

$$e^{-3L\tau}\mathcal{E}_\tau(u_\tau) + 2\int_s^\tau e^{-3Lt}\int_X \big|A_t u_t\big|^2 dm_t\,dt \leq e^{-3Ls}\cdot\mathcal{E}_s(u_s). \tag{29}$$

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

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

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

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

(ii): The log-Lipschitz bound (25) states $|\partial_t\Gamma_t(.)| \leq 2L\cdot\Gamma_t(.)$. Together with (24) this implies $\partial_s\mathcal{E}_s(u_t)\big|_{s=t} \leq 3L\cdot\mathcal{E}_t(u_t)$. Therefore,

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

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

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

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

and thus

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

□

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

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

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

Proposition 2.14. *The following holds true.*

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

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

More generally, for any $M \geq 0$

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

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

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

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

More generally, for any $M \geq 0$

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

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

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

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

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 $t > s$*

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

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

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

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

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

Therefore, for a.e. t

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

where we used (24) in the last inequality. Thus $u_s \le M$ will imply $u_t \le M$ for all $t > s$.

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

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

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

Applying Gronwall's inequality yields

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

which proves the claim.

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

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

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

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

(iv) Choose $w=(u^2-g)_+$. Then, again by the chain rule and since u and g are solutions to the heat equation, we find for a.e. t

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

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

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

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

□

As a direct consequence we obtain the following corollary.

Corollary 2.15. *For all $s < t$*

(i) $\|P_{t,s}\|_{L^\infty(m_s)\to L^\infty(m_t)} \le 1, \qquad \|P^*_{t,s}\|_{L^1(m_t)\to L^1(m_s)} \le 1,$
(ii) $\|P_{t,s}\|_{L^1(m_s)\to L^1(m_t)} \le e^{L(t-s)}, \qquad \|P^*_{t,s}\|_{L^\infty(m_t)\to L^\infty(m_s)} \le e^{L(t-s)},$
(iii) $\|P_{t,s}\|_{L^2(m_s)\to L^2(m_t)} \le e^{L(t-s)/2}, \qquad \|P^*_{t,s}\|_{L^2(m_t)\to L^2(m_s)} \le e^{L(t-s)/2}.$

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 7.

Theorem 2.16. (i) *Then the heat flow is a* dynamic forward EVI$(-L/2,\infty)$-flow for $\frac{1}{2}\times$ the Dirichlet energy *on $L^2(X,m_t)_{t\in I}$, see Appendix. More precisely, for all solutions $(u_t)_{t\in(s,\tau)}$ to the heat equation, for all $\tau \le T$ and all $w \in Dom(\mathcal{E})$*

$$-\frac{1}{2}\partial_s^+ \big\|u_s - w\big\|^2_{s,t}\Big|_{s=t} + \frac{L}{4}\cdot\big\|u_t - w\big\|_t^2 \ \ge \frac{1}{2}\mathcal{E}_t(u_t) - \frac{1}{2}\mathcal{E}_t(w) \tag{30}$$

where $\|.\|_{s,t}$ is defined according to Definition 7.1 with $d_t(v,w) = \big\|v-w\big\|_t = (\int |v-w|^2 dm_t)^{1/2}$.

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

Proof. (i) Assumption (24) implies $\partial_t\big\|v\big\|_t^2 \le L\,\big\|v\big\|_t^2$ as well as (following the argumentation from Proposition 7.2)

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

for all v and t. Therefore, we can estimate

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

(ii) Uniqueness and the growth estimate immediately follow from the EVI-property. Indeed, the distance $\big\|.\big\|_t$ 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: $\partial_t\big\|v\big\|_t^2 \le L\,\big\|v\big\|_t^2$ and the energy satisfies the growth bound $\mathcal{E}_s \le C_0\,\mathcal{E}_t$. □

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

$$\lim_{s\nearrow t} P^*_{t,s} g = g \quad \text{in } (Dom(\mathcal{E}), \sqrt{\mathcal{E}}),$$

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

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

$$\mathcal{E}_t(g) \le \liminf_{s\nearrow t} \mathcal{E}_t(P^*_{t,s}g).$$

On the other hand from Theorem 2.12(iii)

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

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

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

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

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

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

The second assertion follows along the same lines replacing Theorem 2.12(iii) by Theorem 2.12(ii). □

2.4. **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

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

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

Lemma 2.18. *For all $\sigma < \tau$, all solutions $u \in \mathcal{F}_{(\sigma,\tau)}$ to the heat equation, and all solutions $v \in \mathcal{F}_{(\sigma,\tau)}$ to the adjoint heat equation*

$$|\mathcal{E}_t(u_t, v_t) - \mathcal{E}_s(u_s, v_s)| \le C(u_s, v_t) \cdot |t - s|^{1/2} \tag{31}$$

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

$$C(u_s, v_t) = C \cdot \Big[\mathcal{E}_s(u_s) + \mathcal{E}_t(v_t) + \|v_t\|^2_{L^2(m_t)}\Big] \tag{32}$$

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

In other words, the commutator lemma states

$$\left| \int_X \big[A_t(P_{t,s}u_s) - P_{t,s}(A_s u_s)\big]\, v_t\, dm_t \right| \le C(u_s, v_t) \cdot |t - s|^{1/2}. \tag{33}$$

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

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

and thus

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

To proceed, we decompose the integrand into three terms

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

Let us first estimate the second term

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

due to the fact that $|\partial_r\mathcal{E}_r(w)|\le 3L\,\mathcal{E}_r(w)$ for each $w\in\mathcal{F}$. According to Theorem 2.12, the final expressions can be estimated (uniformly in δ) in terms of $\mathcal{E}_s(u_s)$ and $\mathcal{E}_t(v_t)+\|v_t\|^2_{L^2(m_t)}$. Thus we finally obtain

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

Now let us consider jointly the first and third terms

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

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

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

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

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

To summarize, we have

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

with $C:=Le^{3(L+1)T}$ 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* $\big(X, d_t, m_t\big)_{t\in I}$.

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 σ-finite Borel measure on X. We always assume that there exist constants $C, K, L, N' \in \mathbb{R}$ such that

- the metrics d_t are uniformly bounded and equivalent to each other with
$$\left|\log \frac{d_t(x,y)}{d_s(x,y)}\right| \le L\cdot|t-s| \tag{34}$$
for all s,t and all x,y ('log Lipschitz continuity in t');
- the measures m_t are mutually absolutely continuous with bounded, Lipschitz continuous logarithmic densities; more precisely, choosing some reference measure $m_\diamond$ the measures can be represented as $m_t = e^{-f_t}m_\diamond$ with functions f_t satisfying $|f_t(x)| \le C$, $|f_t(x) - f_t(y)| \le C\cdot d_t(x,y)$ and
$$|f_s(x) - f_t(x)| \le L\cdot|s-t| \tag{35}$$
for all s,t and all x,y;
- for each t the static space (X, d_t, m_t) is infinitesimally Hilbertian and satisfies a curvature-dimension condition $\mathrm{CD}(K,N')$ in the sense of [50], [38], [4].

In terms of the metric d_t for given t, we define the L^2-*Kantorovich-Wasserstein metric* W_t on the space of probability measures on X:
$$W_t(\mu,\nu) = \inf\left\{\int_{X\times X} d_t^2(x,y)\,dq(x,y) :\ q\in \mathrm{Cpl}(\mu,\nu)\right\}^{1/2}$$
where $\mathrm{Cpl}(\mu,\nu)$ as usual denotes the set of all probability measures on $X\times X$ with marginals μ and ν. In general, it is not really a metric but just a pseudo metric. Denote by $\mathcal{P} = \mathcal{P}(X)$ the set of all probability measures μ on X (equipped with its Borel σ-field) with $W_t(\mu,\delta_z) < \infty$ for some/all $z\in X$ and $t\in I$.

The log-Lipschitz bound (34) implies that for all $s,t\in I$ and all $\mu,\nu\in\mathcal{P}$
$$\left|\log\frac{W_t(\mu,\nu)}{W_s(\mu,\nu)}\right| \le L\cdot|t-s|, \tag{36}$$
see Corollary 2.2 in [51]. Note that the latter is equivalent to weak differentiability of $t\mapsto W_t(\mu,\nu)$ and $|\partial_t W_t(\mu,\nu)| \le L\cdot W_t(\mu,\nu)$ for all $\mu,\nu\in\mathcal{P}$.

A powerful tool is the dual representation of W_t^2:
$$\frac12 W_t^2(\mu,\nu) = \sup\left\{\int \varphi d\mu + \int \psi d\nu : \varphi(x) + \psi(y) \le \frac12 d_t^2(x,y)\right\},$$
where the supremum is taken among all continuous and bounded functions φ,ψ. Closely related to this is the d_t-Hopf-Lax semigroup defined on bounded Lipschitz functions φ by
$$Q_a^t\varphi(x) := \inf_{y\in X}\left\{\varphi(y) + \frac{1}{2a}d_t^2(x,y)\right\},\quad a>0,\ x\in X.$$
The map $(a,x)\mapsto Q_a^t\varphi(x)$ satisfies the Hamilton-Jacobi equation
$$\partial_a Q_a^t\varphi(x) = -\frac12(\mathrm{lip}_t Q_a^t\varphi)^2(x),\quad \lim_{a\to0} Q_a^t\varphi(x) = \varphi(x). \tag{37}$$
In addition, since (X,d_t) is assumed to be geodesic,
$$\mathrm{Lip}(Q_a^t\varphi) \le 2\mathrm{Lip}(\varphi),\quad \mathrm{Lip}(Q_\cdot^t f(x)) \le 2[\mathrm{Lip}(\varphi)]^2.$$
See for instance [6, Section 3] for these facts.

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

$$\frac{1}{2}W_t^2(\mu_0,\mu_1) = \sup_{\varphi}\left\{\int Q_1^t\varphi d\mu_1 - \int \varphi d\mu_0\right\}. \tag{38}$$

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

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

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

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

See for example [3, Theorem 1.1.2]. For continuous curves $\mu \in \mathcal{C}([0,1],\mathcal{P}(X))$ satisfying $\mu^a = u^a m$ with $u^a \le R$, μ belongs to $AC^2([0,1],\mathcal{P}(X))$ if and only if for each $t \in (0,T)$ there exists a velocity potential $(\Phi_t^a)_a$ such that $\int_0^1\int \Gamma_t(\Phi_t^a)d\mu^a da < \infty$ and

$$\int \varphi d\mu^{a_1} - \int \varphi d\mu^{a_0} = \int_{a_0}^{a_1}\int \Gamma_t(\varphi,\Phi_t^a)d\mu^a da, \text{ for every } \varphi \in Dom(\mathcal{E}). \tag{39}$$

Moreover we can express the metric speed in the following way

$$|\dot{\mu}^a|_t^2 = \int \Gamma_t(\Phi_t^a)d\mu^a. \tag{40}$$

See section 6 and 8 in [7] for a detailed discussion.

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

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

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

$$W_{s,t}(\mu,\nu) = \inf\left\{\int_0^1 |\dot{\mu}^a|^2_{W_{s+a(t-s)}}da\right\}^{1/2} \tag{41}$$

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

In the following we will make frequently use of the concept of regular curves, which has already been successfully used in [6, 17, 7]. 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 ρ is regular (w.r.t. m_t) if:*

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

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

We will use the following approximation result which is a combination of [7, Lemma 12.2] and [17, Lemma 4.11]. For this we define for a fixed time t the semigroup mollification h_ε^t given by

$$h_\varepsilon^t\psi = \frac{1}{\varepsilon}\int_0^\infty H_a^t\psi\kappa\left(\frac{a}{\varepsilon}\right)da, \tag{42}$$

where $(H^t_a)_{a\geq 0}$ denotes the semigroup associated to the Dirichlet form $\mathcal{E}_t$, and $\kappa \in \mathcal{C}^\infty_c((0,\infty))$ with $\kappa \geq 0$ and $\int_0^\infty \kappa(a)da = 1$. Recall that for $\psi \in L^2(m_t) \cap L^\infty(m_t)$, $h^t_\varepsilon\psi, \Delta_t(h^t_\varepsilon\psi) \in Dom(\Delta_t) \cap \mathrm{Lip}_b(X)$. Moreover $||h^t_\varepsilon\psi - \psi|| \to 0$ in $Dom(\mathcal{E})$ as $\varepsilon \to 0$ for $\psi \in Dom(\mathcal{E})$.

Lemma 3.2. *Let X be a RCD(K,∞) space. Let $\rho^0, \rho^1 \in \mathcal{P}(X)$ and $(\rho^a)_{a\in[0,1]}$ be the W_t-geodesic connecting them. Then there exists a sequence of regular curves $(\rho^a_n)_{a\in[0,1]}$, $n \in \mathbb{N}$, such that*

$$W_t(\rho^a_n, \rho^a) \to 0 \text{ for every } a \in [0,1], \tag{43}$$

$$\limsup_{n\to\infty} \int_0^1 |\dot\rho^a_n|^2_t da \leq W^2_t(\rho_0, \rho_1). \tag{44}$$

If we additionally impose that $\rho^0, \rho^1 \in Dom(S)$, then

$$S_t(\rho^a_n) \to S_t(\rho^a) \text{ for every } a \in [0,1], \tag{45}$$

and

$$\limsup_{n\to\infty} \sup_{a\in[0,1]} S_t(\rho^a_n) \leq \sup_{a\in[0,1]} S_t(\rho^a) = \max_{a\in[0,1]} S_t(\rho^a). \tag{46}$$

Proof. We follow the argumentation in [7, Lemma 12.2] and approximate ρ^0, ρ^1 by two sequences of measures $\{\sigma^i_n\}_n$ with bounded densities. Then as in [6, Proposition 4.11] one employs a threefold regularization procedure to the W_t-geodesic $(\nu^a_n)_a$ connecting σ^0_n and σ^1_n: Given $k \in \mathbb{N}$, we first define $\rho^a_{n,k,1} = H^t_{1/k}\nu^a_n$, where H^t denotes the static semigroup. Then we set $\rho^a_{n,k,2} = \int_{\mathbb{R}} \rho^{a-a'}_{n,k,1}\chi_k(a')da'$, where $\chi_k(a) = k\chi(ka)$ for some smooth kernel $\chi \in \mathcal{C}_c(\mathbb{R})$. Finally we set $\rho^a_{n,k} = h^t_{1/k}\rho^a_{n,k,2}$, where $h^t_{1/k}$ is given by (42). Then by a standard diagonal argument one obtains a sequence of regular curves in the sense of Definition 3.1 satisfying (43) and (44).

In order to show (45) and (46) note that since X is a RCD(K,∞) space we have that $a \mapsto S_t(\rho^a)$ is K-convex, where (ρ^a) denotes the W_t geodesic. Together with the lower semicontinuity of the entropy the map $a \mapsto S_t(\rho^a)$ is continuous. Using the convexity properties we follow the argumentation in [17, Lemma 4.11] and insert the explicit formulas of the regularization (ρ^a_n) to obtain

$$\begin{aligned} S_t(\rho^a_n) \leq S_t(\rho^a_{n,2}) &\leq \int_{\mathbb{R}} \chi_n(a') S_t(\rho^{a-a'})da' \\ &\leq S_t(\rho^a) + \int_{\mathbb{R}} \chi_n(a')|S_t(\rho^{a-a'}) - S_t(\rho^a)|da'. \end{aligned} \tag{47}$$

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 (45). One obtains (46) from (47) by exploiting the uniform continuity of the entropy along geodesics on compact intervals once more. □

Later on in this paper (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, 4, 5, 6] applies. In particular, for each t there is a well-defined energy functional

$$\mathcal{E}_t(u) = \int_X |\nabla_t u|^2 dm_t = \liminf_{\substack{v\to u \text{ in } L^2(X,m_t)\\ v\in\mathrm{Lip}(X,d_t)}} \int_X (\mathrm{lip}_t v)^2\, dm_t \tag{48}$$

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

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

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

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

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

Thanks to the RCD(K, ∞)-condition, for each t the domain of the Laplacian coincides with the domain of the Hessian [20], i.e. $Dom(\Delta_t) = W^{2,2}(X, d_t, m_t)$. Indeed, the 'self-improved Bochner inequality' implies that

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

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

$$\|\nabla_t^2 u\|^2 \leq (1 + K_-/2) \cdot \Big(\|\Delta_t u\|^2 + \|u\|^2 \Big) \tag{49}$$

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

Note that in general, $Dom(\Delta_t)$ may depend on t, see Example 2.7.

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 put $\mathcal{H} = L^2(X, m_\diamond)$ and $\mathcal{F} = \mathcal{D}(\mathcal{E}_\diamond)$ as well as

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

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

Theorem 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.

$$P_{t,s} h(x) = \int p_{t,s}(x, y) h(y)\, dm_s(y). \tag{50}$$

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

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

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

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

and has the propagator property

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

Proof. (i) It remains to verify the boundedness and regularity assumptions on f_t and Γ_t which were made for Theorem 2.2. Choose a reference point $t_0 \in I$ and put $\Gamma_\diamond = \Gamma_{t_0}$. Then $\mathcal{E}_\diamond(u) = \int \Gamma_{t_0}(u) e^{-f_{t_0}} dm_\diamond$. The uniform bounds on f_t and on $\Gamma_\diamond(f_t)$ are stated as assumption (35). The log Lipschitz bound (34) on d_t implies the requested uniform bound on Γ_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, $\Gamma_t(.)$ and $\Gamma_t(f_t)$ this allows to apply results of [36] which provides all the assertions of the Theorem. ☐

Remark 3.4. The formula (50) allows to give a pointwise definition for $P_{t,s}h(x)$ for each $h \in L^2(X, m_\diamond)$ (or, in other words, to select a 'nice' version) and, moreover, it allows to extend its definition to $h \in L^1 \cup L^\infty$.

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

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

$$\int u\, d(\hat{P}_{t,s}\mu) = \int (P_{t,s}u) d\mu \qquad (\forall u \in \mathcal{C}_b(X), \forall \mu \in \mathcal{P}(X)). \tag{51}$$

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

$$\partial_t v = \Delta_t v, \tag{52}$$

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

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

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

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

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

and the dual propagator is given by

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

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

Theorem 3.5. *(i) For each $0 \le \sigma < t \le T$ and each $g \in \mathcal{H}$ there exists a unique solution $v \in \mathcal{F}_{(0,t)}$ to the adjoint heat equation $\partial_s v_s = -\Delta_s v_s + (\partial_s f_s) v_s$ on $(\sigma, t) \times X$ with $v_t = g$.*

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

$$P^*_{t,s}g(y) = \int p_{t,s}(x, y) g(x)\, dm_t(x). \tag{53}$$

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

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

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

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

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

and

$$\hat{P}_{t,s}\big(g \cdot m_t\big) = \big(P^*_{t,s}g\big) \cdot m_s. \tag{54}$$

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 $W_\tau, r \in I$, see next section). More precisely, there exists a constant C such that for all $s, s' < t$, all τ and all μ

$$W_\tau^2(\mu_s, \mu_{s'}) \leq C \cdot |s - s'|. \tag{55}$$

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

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

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

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

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

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

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

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

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

According to [48, 36], the heat kernel admits upper Gaussian estimates of the form

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

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

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

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

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

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

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

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

(iv) In a similar calculation as in Proposition 2.8, we find for $\mu = v m_t$, $\mu_s = \hat{P}_{t,s}\mu$ since the dual heat flow is mass preserving,

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

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

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

Then, Theorem 7.3 in [1] yields

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

where the last conclusion is due to our previous calculation. □

Lemma 3.8. *Let $u, g \in Dom(\mathcal{E})$ and $t \in (0,T)$ with $g \in L^1(X, m_t)$. Then*

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

and for a.e. $s < t$

$$\lim_{h\searrow 0} \frac{1}{h}\left(\int uP^*_{t,s+h}gdm_{s+h} - \int uP^*_{t,s}gdm_s\right) = \int \Gamma_s(u, P^*_{t,s}g)dm_s$$

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

$$\frac{1}{h}\left(\int ugdm_t - \int uP^*_{t,t-h}gdm_{t-h}\right) = \int_0^1 \int \Gamma_{t-rh}(u, P^*_{t,t-rh}g)dm_{t-rh}dr. \tag{56}$$

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

$$\left|\int uP^*_{t,t-r_2h}gdm_{t-r_2h} - \int uP^*_{t,t-r_1h}gdm_{t-r_1h}\right| \leq \mathrm{Lip}(u)W_2(\hat{P}_{t,t-r_2h}(gm_t), \hat{P}_{t,t-r_1h}(gm_t)),$$

and hence, as a consequence of Lemma 3.7(ii), the map $r \mapsto \int uP^*_{t,t-rh}gdm_{t-rh}$ is absolutely continuous. Thus

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

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

Since we assume that the space has a lower Riemannian Ricci bound, we obtain equation (56) for every $u \in Dom(\mathcal{E})$ by approximating with Lipschitz functions u_n, satisfying $u_n \to u$ strongly in $(Dom(\mathcal{E}), \sqrt{||\cdot||^2_{L^2(X)} + \mathcal{E}(\cdot)})$, see [5, Proposition 4.10]. Hence

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

where the third inequality directly follows from Lemma 2.17 and the second equality follows from dominated convergence.

Similarly for the second claim we write for $h < t-s$

$$\begin{aligned}\frac{1}{h}\left(\int uP^*_{t,s+h}gdm_{s+h} - \int uP^*_{t,s}gdm_s\right) &= \frac{1}{h}\int_s^{s+h}\partial_r\int uP^*_{t,r}g\,dm_r\,dr\\ = \frac{1}{h}\int_s^{s+h}\int \Gamma_r(u, P^*_{t,r}g)dm_r\,dr,\end{aligned}$$

which converges for a.e. s to $\int \Gamma_s(u, P^*_{t,s}g)\,dm_s$ as $h \searrow 0$. □

To summarize:

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

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

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

4. Towards Transport Estimates

In the sequel, N always will denote an extended number in $(0,\infty]$. The assumptions from section 3.1 will always be in force (in particular, we assume $\mathrm{RCD}^*(K,N')$ and the bounds (34) and (35)). 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 $\big(X,d_t,m_t\big)_{t\in I}$ is a* super-N-Ricci flow *if the Boltzmann entropy S is* dynamical N-convex *on $I\times\mathcal{P}$ in the following sense: for a.e. $t\in I$ and every W_t-geodesic $(\mu^a)_{a\in[0,1]}$ in $\mathcal{P}$ with $\mu^0,\mu^1\in Dom(S)$*

$$\partial_a^+ S_t(\mu^a)\big|_{a=1-}-\partial_a^- S_t(\mu^a)\big|_{a=0+}\ \ \geq\ \ -\frac{1}{2}\partial_t^- W_{t-}^2(\mu^0,\mu^1)+\frac{1}{N}\Big|S_t(\mu^0)-S_t(\mu^1)\Big|^2. \tag{57}$$

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

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

Proposition 4.2. *Given probability measures $\mu,\nu\in\mathcal{D}\subset\mathcal{P}$, then the W_t-geodesic $(\eta^a)_{a\in[0,1]}$ connecting μ and ν has uniformly bounded densities $\frac{d\eta^a}{dm_t}\leq C$ and there exist W_t-Kantorovich potentials ϕ from μ to ν and ψ from ν to μ (both conjugate to each other) such that*

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

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 requested. □

Proposition 4.3. *Given $\tau\leq T$ and $\mu,\nu\in\mathcal{D}\subset\mathcal{P}$, put $\mu_t=\hat{P}_{\tau,t}\mu$ and $\nu_t=\hat{P}_{\tau,t}\nu$. For each $t\in(0,\tau)$, let ϕ_t and ψ_t be any conjugate W_t-Kantorovich potentials from μ_t to ν_t and vice versa. Then for every $0<r<t<s<\tau$*

$$\frac{1}{2}\partial_r^+ W_t^2(\mu_r,\nu_r)|_{r=t-}\leq\mathcal{E}_t(\phi_t,u_t)+\mathcal{E}_t(\psi_t,v_t), \tag{58}$$

and

$$\frac{1}{2}\liminf_{\delta\searrow 0}\frac{1}{\delta}\int_r^s\big[W_t^2(\mu_{t+\delta},\nu_{t+\delta})-W_t^2(\mu_t,\nu_t)\big]\,dt\geq\int_r^s\mathcal{E}_t(\phi_t,u_t)+\mathcal{E}_t(\psi_t,v_t)\,dt. \tag{59}$$

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

Proof. We closely follow the argumentation of the proof of Theorem 6.3 in [2]. According to Proposition 2.12, $u_t,v_t\in Dom(\mathcal{E})$. Moreover, due to boundedness of X, the Kantorovich potentials ϕ_t and ψ_t are Lipschitz and thus also lie in $Dom(\mathcal{E})$. Since ϕ_t and ψ_t are conjugate W_t-Kantorovich potentials from μ_t to ν_t and vice versa, we get

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

whereas

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

for $r\neq t$. Thus with the help of Lemma 3.8 and Theorem 2.5 (ii)

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

This proves the first claim. With the same notation as 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

$$\begin{aligned}
&\frac{1}{2}\liminf_{\delta\searrow 0}\frac{1}{\delta}\int_r^s \left[W_t^2(\mu_{t+\delta},\nu_{t+\delta}) - W_t^2(\mu_t,\nu_t)\right]\,dt \\
&\geq \liminf_{\delta\searrow 0}\frac{1}{\delta}\int_r^s \left[\int \phi_t[d\mu_{t+\delta}-d\mu_t] + \int \psi_t[d\nu_{t+\delta}-d\nu_t]\right]\,dt \\
&\geq \int_r^s \mathcal{E}_t(\phi_t,u_t) + \mathcal{E}_t(\psi_t,v_t)\,dt.
\end{aligned}$$

□

Theorem 4.4. *Assume that $\big(X, d_t, m_t\big)_{t\in(0,T)}$ is a super-Ricci flow and that $(\mu_t)_{t\leq\tau}$ and $(\nu_t)_{t\leq\tau}$ are dual heat flows started in probability measures $\mu_\tau, \nu_\tau \in \mathcal{D}$. Then for a.e. $t \in (0,T)$*

$$\partial_t W_t^2(\mu_t,\nu_t) \geq 0.$$

Proof. The assumptions on the densities are preserved by the dual heat flow, that is, μ_t and ν_t will have densities in $Dom(\mathcal{E})$ which are bounded from above and bounded away from 0, uniformly in t. Using the absolute continuity of $t \mapsto W_t^2(\mu_t,\nu_t)$, we obtain for all $r < s$

$$\begin{aligned}
W_s^2(\mu_s,\nu_s) - W_r^2(\mu_r,\nu_r) \;&\geq\; \limsup_{\delta\searrow 0}\int_r^s \frac{1}{\delta}\Big[W_t^2(\mu_{t+\delta},\nu_{t+\delta}) - W_t^2(\mu_t,\nu_t) \\
&\qquad\qquad + W_{t+\delta}^2(\mu_{t+\delta},\nu_{t+\delta}) - W_t^2(\mu_{t+\delta},\nu_{t+\delta})\Big]dt \\
&\geq\; \liminf_{\delta\searrow 0}\int_r^s \frac{1}{\delta}\Big(W_t^2(\mu_{t+\delta},\nu_{t+\delta}) - W_t^2(\mu_t,\nu_t)\Big)dt \\
&\qquad + \liminf_{\delta\searrow 0}\frac{1}{\delta}\int_r^s \Big(W_{t+\delta}^2(\mu_{t+\delta},\nu_{t+\delta}) - W_t^2(\mu_{t+\delta},\nu_{t+\delta})\Big)dt \\
&\geq\; \int_r^s 2\Big(\mathcal{E}_t(u_t,\phi_t) + \mathcal{E}_t(v_t,\psi_t)\Big)dt \\
&\qquad + \liminf_{\delta\searrow 0}\frac{1}{\delta}\int_r^s \Big(W_t^2(\mu_t,\nu_t) - W_{t-\delta}^2(\mu_t,\nu_t)\Big)dt \\
&\geq\; \int_r^s 2\Big(\mathcal{E}_t(u_t,\phi_t) + \mathcal{E}_t(v_t,\psi_t)\Big)dt \\
&\qquad - \int_r^s 2\Big(\mathcal{E}_t(u_t,\phi_t) + \mathcal{E}_t(v_t,\psi_t)\Big)dt \geq 0,
\end{aligned}$$

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.

$$-\frac{1}{2}\partial_r^- W_r^2(\mu_t,\nu_t)\big|_{r=t-} \leq \partial_a S(\eta_t^{1-}) - \partial_a S(\eta_t^{0+})$$

for every W_t-geodesic $(\eta_t^b)_{b\in[0,1]}$ connecting μ_t and ν_t. In the previous argumentation, we used in the third and fourth inequality that $\frac{1}{\delta}[W_{t+\delta}^2 - W_t^2]$ is uniformly bounded, which is due to the log-Lipschitz bound on the distances. □

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

$$W_s(\mu_s,\nu_s) \leq W_t(\mu_t,\nu_t). \tag{60}$$

Proof. For measures μ_τ, ν_τ with densities in $Dom(\mathcal{E})$ which are bounded from above and bounded away from 0 the estimate (60) immediately follows from the previous theorem and the fact that the map $t \mapsto W_t(\mu_t,\nu_t)$ 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 (60) carries over to all $\mu_\tau, \nu_\tau \in \mathcal{P}$. □

Theorem 4.6 (“$\mathbf{(I}_N\mathbf{)} \Rightarrow \mathbf{(II}_N\mathbf{)}$”). *Assume that $\big(X, d_t, m_t\big)_{t\in(0,T)}$ is a super-N-Ricci flow and that probability measures $\mu_\tau, \nu_\tau \in \mathcal{P}$ are given for some $\tau \in (0, T]$. Then the dual heat flows $(\mu_t)_{t\le\tau}$ and $(\nu_t)_{t\le\tau}$ starting in these points satisfy for all $0 \le s < t \le \tau$*

$$W_s^2(\mu_s, \nu_s) \le W_t^2(\mu_t, \nu_t) - \frac{2}{N}\int_s^t [S_r(\mu_r) - S_r(\nu_r)]^2\, dr. \tag{61}$$

Proof. For measures μ_τ, ν_τ within the subset $\mathcal{D}$ we follow the proof of the previous Theorem 4.4 line by line and finally use the enforcement of the super Ricci flow property to deduce

$$\begin{aligned}-\frac{1}{2}\liminf_{\delta\searrow 0}\frac{1}{\delta}\Big[W_{t+\delta}^2(\mu_{t+\delta}, \nu_{t+\delta}) - W_t^2(\mu_{t+\delta}, \nu_{t+\delta})\Big] \;&\le\; \partial_a S_t(\eta_t^{1-}) - \partial_a S_t(\eta_t^{0+})\\ &\qquad -\frac{1}{N}\,[S_t(\mu_t) - S_t(\nu_t)]^2\,.\end{aligned}$$

Together with the other estimates from the proof of the previous theorem this gives

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

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

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

as $t \uparrow \tau$. Here the convergence of the integrals is obvious. The convergence of the first term on the right-hand side follows from Lemma 3.7. □

4.2. **From Gradient Estimates to Transport Estimates.**

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

$$W_s^2(\mu_s^0, \mu_s^1) \le W_\tau^2(\mu_\tau^0, \mu_\tau^1) - \frac{2}{N}\int_s^\tau \big[S_t(\mu_t^0) - S_t(\mu_t^1)\big]^2\, dt. \tag{62}$$

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

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

of the curve $a \mapsto \mu_t^a = \hat{P}_{\tau,t}\mu_\tau^a$. Let $t \in (0, \tau]$ be given with $\mathcal{A}_t\big(\dot{\mu_t}\big) < \infty$. In other words, such that the curve $a \mapsto \mu_t^a$ is 2-absolutely continuous. (Obviously, this is true for $t = \tau$. The subsequent discussion indeed will show that this holds for all $t \le \tau$.) Let $(u_t^a)_{a\in[0,1]}$ and $(\Phi_t^a)_{a\in[0,1]}$ denote the densities and velocity potentials for the curve $(\mu_t^a)_{a\in[0,1]}$ (see [7, Theorem 8.2], or (39),(40)) in the static space (X, d_t, m_t). Then, in particular,

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

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

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

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

Then applying the continuity equation (39) and the Hamilton-Jacobi equation (37) yields

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

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

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

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

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

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

This is the claim.

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

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

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

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

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

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

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

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

Let us also briefly remark that the densities u_t^a of the measures μ_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' inequality $(a-b)^2 \geq \frac{\delta}{1+\delta}a^2 - \delta b^2$ where $\delta = N/\varepsilon$

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

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

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

Passing to the limit $s \nearrow t$ yields

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

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

$$\partial_{t-}^- \mathcal{A}_t(\dot{\mu_t}) \geq \frac{2}{N}\Big|S_t(\mu_t^1) - S_t(\mu_t^0)\Big|^2. \tag{63}$$

Recall that this holds for a.e. $t \in (0,\tau)$. Moreover, note that $t \mapsto \mathcal{A}_t(\mu_t^\cdot)$ is absolutely continuous. Indeed, by Lemma 3.7 and the log-Lipschitz assumption (34)

$$\begin{aligned}
\Big|W_{t+\epsilon}^2(\mu_{t+\epsilon}^a,\mu_{t+\epsilon}^b) - W_t^2(\mu_t^a,\mu_t^b)\Big| &\leq \Big|W_{t+\epsilon}^2(\mu_{t+\epsilon}^a,\mu_t^b) - W_t^2(\mu_t^a,\mu_t^b)\Big| \\
&\quad + \Big|W_t^2(\mu_{t+\epsilon}^a,\mu_{t+\epsilon}^b) - W_t^2(\mu_t^a,\mu_t^b)\Big| \\
&\leq 2L\epsilon\, e^{2L\epsilon}\, W_t^2(\mu_t^a,\mu_t^b) \\
&\quad + \frac{2\sqrt{\epsilon}}{1-2\sqrt{\epsilon}} W_t^2(\mu_t^a,\mu_t^b) + \frac{1}{\sqrt{\epsilon}} W_t^2(\mu_{t+\epsilon}^a,\mu_t^a) + + \frac{1}{\sqrt{\epsilon}} W_t^2(\mu_{t+\epsilon}^b,\mu_t^b) \\
&\leq C_0\sqrt{\epsilon}\, W_t^2(\mu_t^a,\mu_t^b) + C_1\sqrt{\epsilon}.
\end{aligned}$$

Thus we may integrate (63) from any $s \in (0,\tau)$ to τ to obtain

$$\mathcal{A}_s(\mu_s^\cdot) \leq \mathcal{A}_\tau(\mu_\tau^\cdot) - \frac{2}{N}\int_s^\tau \left[S_t(\mu_t^0) - S_t(\mu_t^1)\right]^2 dt. \tag{64}$$

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

$$\mathcal{A}_\sigma(\tilde{\mu}_\sigma^\cdot) \to W_\tau^2(\mu_\tau^0,\mu_\tau^1)$$

as $\sigma \nearrow \tau$. Carrying out the previous estimations, finally resulting in (64), with $(\tilde{\mu}_\sigma^a)_{a\in[0,1]}$ in the place of $(\mu_\tau^a)_{a\in[0,1]}$ yields

$$\begin{aligned}
W_s^2(\mu_s^0,\mu_s^1) &\leq \mathcal{A}_s(\tilde{\mu}_s^\cdot) \\
&\leq \mathcal{A}_\sigma(\tilde{\mu}_\sigma^\cdot) - \frac{2}{N}\int_s^\sigma \left[S_t(\mu_t^0) - S_t(\mu_t^1)\right]^2 dt \\
&\to W_\tau^2(\mu_\tau^0,\mu_\tau^1) - \frac{2}{N}\int_s^\tau \left[S_t(\mu_t^0) - S_t(\mu_t^1)\right]^2 dt.
\end{aligned}$$

This proves the claim. □

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

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

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

Choosing $t_n \nearrow \tau$ and $\delta_n \searrow 0$ yields a sequence of regular curves satisfying (43) - (46). *In addition, for these approximations the endpoints are simply given by the dual heat flow:*

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

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

Proof. The re-parametrization by means of ϑ forces the curve to be constant for some short interval around the endpoints and squeeze it in-between. The latter leads to a moderate increase of the metric speed. The former guarantees that the endpoints remain unchanged under the convolution. The convolution w.r.t. the kernel χ guarantees smooth dependence on a, i.e. (1) of Def 3.1. (43) 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 $u_t^a(x)$ for $(a,t,x)\in[0,1]\times(0,t]\times X$ for each $t<\tau$, i.e. (2) of Def 3.1. Moreover, $u_t^a(x)$ is uniformly bounded away from 0. Thus (3) of Def 3.1 is equivalent to a uniform bound for the energy $\mathcal{E}_t(u^a)$.

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

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

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 χ thus turns the integrable function $a \mapsto \mathcal{E}_t\left(u_t^{\vartheta(a)}\right)$ into a bounded function: $\int_{\mathbb{R}} \mathcal{E}_t\left(u_t^{\vartheta(a+\delta b)}\right)\chi(b)db \le C$. Since the energy $u \mapsto \mathcal{E}_t(u)$ is convex, Jensen's inequality implies

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

The action estimate (44) follows from part (i) of the previous proof. Indeed, the dual heat flow decreases the action. Also convolution in the a-parameter decreases the action. The re-parametrization increases the action by a factor bounded by $\frac{1}{(1-2\delta)^2}$.

The entropy estimates (45) and (46) follow as in the proof of Lemma 3.2 □

4.3. **Duality between Transport and Gradient Estimates in the Case $N = \infty$.** In the subsequent chapter, we will prove the implication $\mathbf{(II}_N\mathbf{)} \Rightarrow \mathbf{(III}_N\mathbf{)}$ by composing the results $\mathbf{(II}_N\mathbf{)} \Rightarrow \mathbf{(IV}_N\mathbf{)}$ and $\mathbf{(IV}_N\mathbf{)} \Rightarrow \mathbf{(III}_N\mathbf{)}$. Partly, these arguments are quite involved. (And actually, for the last one, we freely make use of the subsequent 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 ("**(II) ⇔ (III)**"). *For fixed $0 < s < t < T$ the following are equivalent:*

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

$$W_s(\hat{P}_{t,s}\mu, \hat{P}_{t,s}\nu) \le W_t(\mu, \nu) \tag{65}$$

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

$$\Gamma_t(P_{t,s}u) \le P_{t,s}(\Gamma_s(u)) \quad m\text{-a.e. on } X. \tag{66}$$

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

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

Thus following the argumentation from [5], Theorem 6.4, we obtain

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

The Hölder continuity of $z \mapsto P_{t,s}|\nabla_s u|^2(z)$, therefore, allows 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 $u \in Dom(\mathcal{E})$ follows as in [5].

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

5. From Transport Estimates to Gradient Estimates and Bochner Inequality

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

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

5.1. **The Bochner Inequality.**

The Time-Derivative of the Γ*-Operator.*

Definition 5.1. *Given an interval* $J \subset I$ *and* $u \in \mathcal{F}_J$ *with* $\Gamma_r(u_r)(x) \le C$ *uniformly in* $(r, x) \in J \times X$*. Then we define* $\overset{\bullet}{\Gamma}_r (u_r)(x)$ *as (one of the) weak subsequential limit(s) of*

$$\frac{1}{2\delta}\Big[\Gamma_{r+\delta}(u_r) - \Gamma_{r-\delta}(u_r)\Big](x) \tag{67}$$

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

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

as $n \to \infty$*.*

Actually, thanks to Banach-Alaoglu theorem, such a weak limit always exists since (67) – due to the log-Lipschitz continuity of the distances – defines a family of functions in $L^2(J \times X)$ with bounded norm. Thus in particular we will have

$$\begin{aligned}
&\liminf_{\delta\to 0} \frac{1}{2\delta}\int_J\int_X \Big[\Gamma_{r+\delta}(u_r) - \Gamma_{r-\delta}(u_r)\Big]\, g_r\, dm_r\, dr\\
&\le \int_J\int_X \overset{\bullet}{\Gamma}_r (u_r)\, g_r\, dm_r\, dr \qquad\qquad (68)\\
&\le \limsup_{\delta\to 0} \frac{1}{2\delta}\int_J\int_X \Big[\Gamma_{r+\delta}(u_r) - \Gamma_{r-\delta}(u_r)\Big]\, g_r\, dm_r\, dr.
\end{aligned}$$

Remark 5.2. All the subsequent statements involving $\overset{\bullet}{\Gamma}_r (u_r)$ will be independent of the choice of the sequence $(\delta_n)_n$ and of the accumulation point in $L^2(J \times X)$. For instance, the precise meaning of Theorem 1.7 is that each of the properties **(I)**, **(II)** or **(III)** will imply **(IV)** for *every* choice of the weak subsequential limit $\overset{\bullet}{\Gamma}_r (u_r)$. Conversely, if **(IV)** is satisfied for *some* choice of the weak subsequential limit $\overset{\bullet}{\Gamma}_r (u_r)$ then it implies properties **(I)**, **(II)** and **(III)**. Indeed, the only property of $\overset{\bullet}{\Gamma}_r (u_r)$ which enters the calculations is (68).

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

$$\Big|\overset{\bullet}{\Gamma}_r (u_r)\Big| \le 2L \cdot \Gamma_r(u_r). \tag{69}$$

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

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

In particular,

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

Proof.

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

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

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

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

The Distributional Γ_2*-Operator.*

Definition 5.4. *For* $r\in(0,T)$ *and* $u\in Dom(\Delta_r)$ *with* $|\nabla_r u|\in L^\infty$ *we define the distribution valued* Γ_2*-operator as a continuous linear operator*

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

by

$$\mathbf{\Gamma}_{2,r}(u)(g):=\int\Big[-\frac{1}{2}\Gamma_r\big(\Gamma_r(u),g\big)+(\Delta_r u)^2 g+\Gamma_r(u,g)\Delta_r u\Big]dm_r. \tag{70}$$

Note that

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

thanks to the fact that $\|\nabla_r^2 u\|_2^2\le(1+K_-)\cdot(\|\Delta_r u\|_2^2+\|u\|_2^2)$, cf. (49).

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

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

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

On the other hand, if $g\in Dom(\Delta_r)$ then in (70) we may replace the term $-\Gamma_r\big(\Gamma_r(u),g\big)$ by $\Gamma_r(u)\Delta_r g$.

The Bochner Inequality.

Definition 5.5. *(i) We say that $(X, d_t, m_t)_{t\in I}$ satisfies the dynamic Bochner inequality with parameter $N \in (0, \infty]$ if for all $0 < s < t < T$ and for all $u_s, g_t \in \mathcal{F}$ with $g_t \geq 0$, $g_t \in L^\infty$, $u_s \in \mathrm{Lip}(X)$ and for a.e. $r \in (s,t)$*

$$\mathbf{\Gamma}_{2,r}(u_r)(g_r) \geq \frac{1}{2}\int \overset{\bullet}{\Gamma}_r\,(u_r) g_r dm_r + \frac{1}{N}\Big(\int \Delta_r u_r g_r dm_r\Big)^2 \tag{71}$$

where $u_r = P_{r,s}u_s$ and $g_r = P^*_{t,r}g_t$, cf. (11).

(ii) We say that $(X, d_t, m_t)_{t\in I}$ satisfies property $(\mathbf{IV}_N)$ *if it satisfies the dynamic Bochner inequality with parameter N as above and in addition the regularity assumption* (7) *is satisfied, i.e. $u_r \in \mathrm{Lip}(X)$ for all $r \in (s,t)$ with $\sup_{r,x}\mathrm{lip}_r u_r(x) < \infty$.*

Note that in the case $N = \infty$ inequality (71) simply states that

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

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

5.2. **From Bochner Inequality to Gradient Estimates.**

Theorem 5.6 ("$(\mathbf{IV_N}) \Rightarrow (\mathbf{III_N})$"). *Suppose that the mm-space $(X, d_t, m_t)_{t\in I}$ satisfies the dynamic Bochner inequality* (71) *and the regularity assumption from Definition 5.5 (ii). Then for a.e. $x \in X$*

$$\Gamma_t(P_{t,s}u)(x) - P_{t,s}\Gamma_s(u)(x) \leq -\frac{2}{N}\int_s^t \big[P_{t,r}\Delta_r u_r(x)\big]^2 dr. \tag{72}$$

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

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

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

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

$$h_\tau \leq \liminf_{\delta\searrow 0}\frac{1}{\delta}\int_{\tau-\delta}^{\tau} h_r dr \quad \text{and} \quad h_\sigma \geq \limsup_{\delta\searrow 0}\frac{1}{\delta}\int_\sigma^{\sigma+\delta} h_r dr. \tag{73}$$

Note that by Lebesgue's density theorem, the latter is true at least for a.e. $\sigma \geq s$ and for a.e. $\tau \leq t$. (Moreover, at the end of this proof (as part (b)) we will present an argument which allows to conclude that (73) holds for $\sigma = s, \tau = t$.) Then

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

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

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

due Lebesgue's density theorem applied to $r\mapsto \Delta_r g_r e^{-f_r}$. Note that the latter function is in L^2 (Theorem 2.12) and the function $r\mapsto \Gamma_r(u_r)$ is in L^∞ thanks to Definition 5.5 (ii).

The second term can easily estimated in terms $\overset{\bullet}{\Gamma}_r$ according to Lemma 5.3:

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

The term (III') is transformed as follows

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

Here again we used Lebesgue's density theorem (applied to $r\mapsto\Delta_r u_r$) and the 'nearly continuity' of $r\mapsto g_r$ as map from (s,t) into $L^2(X,m)$ and as map into $\mathcal{F}$ (Lusin's theorem). Moreover, we used the boundedness (uniformly in r and x) of g_r and of $\nabla_r u_r$ as well as the square integrability of $\Delta_r u_r$.

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

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

Summarizing and then using (71), we therefore obtain

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

Thus

$$\int_X\Gamma_\tau(P_{\tau,\sigma}u)g\,dm_\tau-\int_X P_{\tau,\sigma}\Gamma_\sigma(u)\,g\,dm_\tau\le-\frac{2}{N}\int_\sigma^\tau\Big[\int_X P_{\tau,r}\Delta_r u_r\,g\,dm_\tau\Big]^2dr. \tag{74}$$

(b) Recall that, given u and g, this holds for a.e. τ and a.e. σ. Now let us forget for the moment the term with N. Choosing g's from a dense countable set one may achieve that the exceptional sets for σ and τ in (74) do not depend on g. Next we may assume that $\sigma,\tau\in[s,t]$

with $\sigma < \tau$ is chosen such that (74) with $N = \infty$ simultaneously holds for all u from a dense countable set $\mathcal{C}_1$ in $\mathrm{Lip}(X)$. Approximating arbitrary $u \in \mathrm{Lip}(X)$ by $u_n \in \mathcal{C}_1$ yields

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

due to lower semicontinuity of the weighted energy on L^2. In other words, we have derived the gradient estimate **(III)** for almost all times σ and τ. Thanks to Theorem 4.9 this implies the transport estimate **(II)** for these time instances. But both sides of the transport estimate are continuous in time (thanks to the continuity of $r \mapsto W_r$ and the continuity of the dual heat flow). This implies that the transport estimate holds for all $\sigma, \tau \in [s,t]$ with $\sigma < \tau$. In particular, it holds for $\sigma = s$ and $\tau = t$. Again by Theorem 4.9 it yields the gradient estimate for given s and t and thus our initial assumption (73) is satisfied for the choice $\sigma = s$ and $\tau = t$.

(c) Taking this into account, we may conclude that (74) (for given N) holds with the choice $\sigma = s$ and $\tau = t$. Finally, choosing sequences of g's which approximate the Dirac distribution at a given $x \in X$ then implies that for all $u \in \mathrm{Lip}(X)$

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

for a.e. $x \in X$. This proves the claim for bounded Lipschitz functions. The extension to $u \in \mathit{Dom}(\mathcal{E})$ follows as in [5]. ☐

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

Theorem 5.7 ("$\mathbf{(III_N)} \Rightarrow \mathbf{(IV_N)}$"). *Suppose that the mm-space $(X, d_t, m_t)_{t\in I}$ satisfies the gradient estimate* (72). *Then the dynamic Bochner inequality* (71) *holds true as well as the regularity assumption from Definition 5.5 (ii).*

Proof. Assume that the gradient estimate $\mathbf{(III}_N\mathbf{)}$ holds true. It immediately implies the regularity assumption (7). To derive the dynamic Bochner inequality, let $s, t \in (0,T)$ as well as $u \in \mathrm{Lip}(X)$ and $g \in \mathcal{F} \cap L^\infty$ with $g \ge 0$ be given. Put $u_r = P_{r,s}u$, $g_r = P^*_{t,r}g$ for $r \in [s,t]$ and as before consider the function

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

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

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

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

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

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

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

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

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

for a.e. $r\in(s,t)$. □

5.4. From Transport Estimates to Bochner Inequality.

Theorem 5.8 ("($\mathbf{II_N}$) ⇒ ($\mathbf{IV_N}$)")**.** *Suppose that the mm-space $(X,d_t,m_t)_{t\in I}$ satisfies the transport estimate* (9)=(61)*. Then the dynamic Bochner inequality* (10)=(71) *with parameter N holds true as well as the regularity assumption* (7).

Proof of the regularity assumption. Thanks to Theorem 4.9, we already know that the transport estimate ($\mathbf{II}_N$) implies the gradient estimate ($\mathbf{III}_N$) in the case $N=\infty$. This proves the requested regularity. □

Proof of the dynamic Bochner inequality. We follow the argumentation from [12] with significant modifications due to time-dependence of functions, gradients, and operators and mainly because of lack of regularity.

Let $0<s<t<T$ and $g_t\in\mathcal{F}\cap L^\infty$ with $g_t\geq 0$, $g_t\not\equiv 0$ as well as $u_s\in \mathrm{Lip}(X)$ be given and fixed for the sequel. Without restriction $\int g_t dm_t = 1$. For $\tau\in(s,t)$, put $u_\tau = P_{\tau,s}u_s$ and $g_\tau = P^*_{t,\tau}g_t$. Note that – thanks to the parabolic Harnack inequality – g is uniformly bounded from above and bounded from below, away from 0, on $(s',t')\times X$ for each $s<s'<t'<t$. In the beginning, let us also assume that $||u_s||_\infty\leq 1/4$.

For each $\tau\in(s,t)$, define a Dirichlet form $\mathcal{E}^g_\tau$ on $L^2(X,g_\tau m_\tau)$ with domain $Dom(\mathcal{E}^g_\tau):=Dom(\mathcal{E})$ by

$$\mathcal{E}^g_\tau(u) := \int \Gamma_\tau(u)g_\tau dm_\tau \quad \text{ for } u\in Dom(\mathcal{E}).$$

Associated with the closed bilinear form $(\mathcal{E}^g_\tau, Dom(\mathcal{E}^g_\tau))$ on $L^2(X,g_\tau m_\tau)$, there is the self-adjoint operator Δ^g_τ and the semigroup $(H^{\tau,g}_a)_{a\geq 0}$, i.e. $u_a = H^{\tau,g}_a u$ solves

$$\partial_a u_a = \Delta^g_\tau u_a \text{ on } (0,\infty)\times X, \qquad u_0 = u$$

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

$$g^{\sigma,a}_\tau := g_\tau(1+u_\sigma - H^{\tau,g}_a u_\sigma). \tag{76}$$

Note that these are probability densities w.r.t. m_τ. Indeed, for all $a>0$ and all $s<\sigma<\tau<t$

$$\int g_\tau^{\sigma,a} dm_\tau = 1 + \int u_\sigma (1 - H_a^{\tau,g} 1)\, g_\tau m_\tau = 1$$

thanks to conservativeness and symmetry of $H_a^{\tau,g}$ w.r.t. the measure $g_\tau m_\tau$. Moreover, $g_\tau^{\sigma,a} \geq 0$ for all a, σ and τ since the uniform bound $||u_s||_\infty \leq 1/4$ is preserved under the evolution of the time-dependent heat flow, thus $||u_\sigma||_\infty \leq ||P_{\sigma,s} u_s||_\infty \leq 1/4$, as well as under the heat flow in the static mm-space at fixed time τ, thus $||H_a^{\tau,g} u_\sigma||_\infty \leq ||u_\sigma||_\infty \leq 1/4$.

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

$$\begin{aligned} W_\sigma^2(\hat{P}_{\tau,s}(g_\tau m_\tau), \hat{P}_{\tau,\sigma}(g_\tau^{\sigma,a} m_\tau)) \quad \leq \quad & W_\tau^2(g_\tau m_\tau, g_\tau^{\sigma,a} m_\tau) \\ & -\frac{2}{N} \int_\sigma^\tau [S_r(\hat{P}_{\tau,r}(g_\tau m_\tau)) - S_r(\hat{P}_{\tau,r}(g_\tau^{\sigma,a} m_\tau))]^2 dr. \end{aligned}$$

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

$$\begin{aligned} & -\frac{1}{2} \int P_{\tau,\sigma}(\Gamma_\sigma(u_\sigma)) g_\tau dm_\tau + \int \Gamma_\tau(P_{\tau,\sigma} u_\sigma, u_\sigma) g_\tau dm_\tau \\ & \leq \frac{1}{2(1-2||u_\sigma||_\infty)} \int \Gamma_\tau(u_\sigma) g_\tau dm_\tau - \frac{1}{N} \int_\sigma^\tau \left[\int \Gamma_\tau\big(P_{\tau,r}(\log P_{\tau,r}^* g_\tau), u_\sigma\big)\, g_\tau dm_\tau \right]^2 dr. \end{aligned}$$

Replacing u_s by $\eta\, u_s$ for $\eta \in \mathbb{R}_+$ sufficiently small, we can get rid of the constraint $||u_s||_\infty \leq 1/4$. Then Lemma 5.9, Lemma 5.10 and Lemma 5.11 applied to ηu_s instead of u_s gives us

$$\begin{aligned} & -\frac{\eta^2}{2} \int P_{\tau,\sigma}(\Gamma_\sigma(u_\sigma)) g_\tau dm_\tau + \eta^2 \int \Gamma_\tau(P_{\tau,\sigma} u_\sigma, u_\sigma) g_\tau dm_\tau \\ & \leq \frac{\eta^2}{2(1-2\eta||u_\sigma||_\infty)} \int \Gamma_\tau(u_\sigma) g_\tau dm_\tau - \frac{\eta^2}{N} \int_\sigma^\tau \left[\int \Gamma_\tau\big(P_{\tau,r}(\log P_{\tau,r}^* g_\tau), u_\sigma\big)\, g_\tau dm_\tau \right]^2 dr. \end{aligned}$$

Dividing by η^2 and letting $\eta \to 0$ this inequality becomes

$$\begin{aligned} & -\frac{1}{2} \int P_{\tau,\sigma}(\Gamma_\sigma(u_\sigma)) g_\tau dm_\tau + \int \Gamma_\tau(P_{\tau,\sigma} u_\sigma, u_\sigma) g_\tau dm_\tau \\ & \leq \frac{1}{2} \int \Gamma_\tau(u_\sigma) g_\tau dm_\tau - \frac{1}{N} \int_\sigma^\tau \left[\int \Gamma_\tau\Big(P_{\tau,r}(\log P_{\tau,r}^* g_\tau), u_\sigma\Big)\, g_\tau dm_\tau \right]^2 dr. \end{aligned}$$

This can be reformulated into

$$\begin{aligned} & \frac{1}{2} \int \Gamma_\tau(u_\tau) g_\tau dm_\tau - \frac{1}{2} \int \Gamma_\sigma(u_\sigma) g_\sigma dm_\sigma \\ & -\frac{1}{2} \int \Gamma_\tau(u_\sigma) g_\tau dm_\tau - \frac{1}{2} \int \Gamma_\tau(u_\tau) g_\tau dm_\tau + \int \Gamma_\tau(u_\tau, u_\sigma) g_\tau dm_\tau \\ & \leq -\frac{1}{N} \int_\sigma^\tau \left[\int \Gamma_\tau\Big(P_{\tau,r}(\log P_{\tau,r}^* g_\tau), u_\sigma\Big)\, g_\tau dm_\tau \right]^2 dr. \end{aligned} \tag{77}$$

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

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

for $r \in (s,t)$. Recall that we already know from Theorem 4.9 that the transport estimate $(\mathbf{II}_N)$ implies the gradient estimate $(\mathbf{III})$ ('without N'). Thus for all $s<\sigma<\tau<t$

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

Arguing as in the proof of Theorem 5.7 we get

$$h_\tau - h_\sigma \geq \int_\sigma^\tau \Big[-2\mathbf{\Gamma}_{2,r}(u_r)(g_r) + \int \overset{\bullet}{\Gamma}_r (u_r)\, g_r\, m_r \Big]\, dr.$$

On the other hand, applying the previous estimate (77) (with $r+\delta$, r and q in the place of τ, σ and r) we obtain

$$\begin{aligned} h_\tau - h_\sigma \leq \liminf_{\delta \searrow 0} \frac{1}{\delta} \int_\sigma^{\tau-\sigma} \Big[-\frac{2}{N} \int_r^{r+\delta} \Big[\int \Gamma_{r+\delta}\Big(P_{r+\delta,q}(\log P^*_{r+\delta,q} g_{r+\delta}), u_r\Big)\, g_{r+\delta} dm_{r+\delta} \Big]^2 dq \\ + \int \Gamma_{r+\delta}(u_{r+\delta} - u_r) g_{r+\delta} dm_{r+\delta} \Big]\, dr. \end{aligned}$$

We estimate the term with the square from below using Young's inequality

$$\begin{aligned} &\Big[\int \Gamma_{r+\delta}\Big(P_{r+\delta,q}(\log P^*_{r+\delta,q} g_{r+\delta}), u_r\Big)\, g_{r+\delta} dm_{r+\delta} \Big]^2 \\ &\geq \frac{1}{1+\epsilon} \Big[\int \Gamma_r\Big(P_{r,q}(\log g_q), u_r\Big)\, g_r dm_r \Big]^2 \\ &- \frac{1}{\epsilon} \Big[\int \Gamma_{r+\delta}\Big(P_{r+\delta,q}(\log P^*_{r+\delta,q} g_{r+\delta}), u_r\Big)\, g_r dm_{r+\delta} - \int \Gamma_r\Big(P_{r,q}(\log g_q), u_r\Big)\, g_r dm_r \Big]^2, \end{aligned}$$

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

$$\begin{aligned} &\Big[\int \Gamma_{r+\delta}\Big(P_{r+\delta,q}(\log P^*_{r+\delta,q} g_{r+\delta}), u_r\Big)\, g_{r+\delta} dm_{r+\delta} - \int \Gamma_r\Big(P_{r,q}(\log g_q), u_r\Big)\, g_r dm_r \Big]^2 \\ &\leq 2 \Big[\int \Gamma_{r+\delta}\Big(P_{r+\delta,q}(\log g_q), u_r\Big)\, g_{r+\delta} dm_{r+\delta} - \int \Gamma_r\Big(P_{r+\delta,q}(\log g_q), u_r\Big)\, g_{r+\delta} dm_{r+\delta} \Big]^2 \\ &+ 2 \Big[\int \Gamma_r\Big(P_{r+\delta,q}(\log g_q), u_r\Big)\, g_{r+\delta} dm_{r+\delta} - \int \Gamma_r\Big(P_{r,q}(\log g_q), u_r\Big)\, g_{r+\delta} dm_{r+\delta} \Big]^2 \\ &+ 2 \Big[\int \Gamma_r\Big(P_{r,q}(\log g_q), u_r\Big)\, g_{r+\delta} dm_{r+\delta} - \int \Gamma_r\Big(P_{r,q}(\log g_q), u_r\Big)\, g_r dm_r \Big]^2 \\ &\leq 16 L^2 \delta^2 \Big[\int \Gamma_{r+\delta}\Big(P_{r+\delta,q}(\log g_q), u_r\Big)\, g_{r+\delta} dm_{r+\delta} + C \int \Gamma_{r+\delta}\Big(P_{r+\delta,q}(\log g_q) - u_r\Big)\, g_{r+\delta} dm_{r+\delta} \Big]^2 \\ &+ 2 \Big[\int \Gamma_r\Big(P_{r+\delta,q}(\log g_q) - P_{r,q}(\log g_q), u_r\Big)\, g_{r+\delta} dm_{r+\delta} \Big]^2 \\ &+ 2 \Big[\int \Gamma_r\Big(P_{r,q}(\log g_q), u_r\Big)\, d(g_{r+\delta} dm_{r+\delta} - g_r m_r) \Big]^2, \end{aligned}$$

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

$$\begin{aligned} \delta \int_r^{r+\delta} \Big| \int \Gamma_{r+\delta}\Big(P_{r+\delta,q}(\log P^*_{r+\delta,q} g_{r+\delta}), u_r\Big)\, g_{r+\delta} dm_{r+\delta} \Big|^2 dq \\ \leq C\delta \Big(\int_r^{r+\delta} \int \Gamma_q(\log g_q)\, g_q dm_q dr \Big) \mathcal{E}_r(u_r) \xrightarrow[\delta \to 0]{} 0, \end{aligned}$$

and Lemma 2.17 and Lebesgue differentiation theorem

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

while

$$\frac{1}{\delta}\int_r^{r+\delta}\Big[\int\Gamma_r\Big(P_{r,q}(\log g_q),u_r\Big)\,d(g_{r+\delta}dm_{r+\delta}-g_rm_r)\Big]^2\,dq\xrightarrow[\delta\to0]{}0.$$

Thus, since ϵ is arbitrary, and from the Lebesgue differentiation theorem we get

$$\begin{aligned}\liminf_{\delta\to0}\frac{1}{\delta}\int_r^{r+\delta}&\Big[\int\Gamma_{r+\delta}\Big(P_{r+\delta,q}(\log P^*_{r+\delta,q}g_{r+\delta}),u_r\Big)\,g_{r+\delta}dm_{r+\delta}\Big]^2\,dr\\&\geq\Big[\int\Gamma_r\Big(\log g_q,u_r\Big)\,g_rdm_r\Big]^2=\Big[\int(\Delta_ru_r)g_rdm_r\Big]^2.\end{aligned}$$

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

$$\begin{aligned}&\limsup_{\delta\to0}\frac{1}{\delta}\int_\sigma^{\tau-\delta}\int\Gamma_{r+\delta}(u_{r+\delta}-u_r)g_{r+\delta}dm_{r+\delta}\,dr\\&\leq\limsup_{\delta\to0}\frac{1}{\delta}\int_\sigma^{\tau-\delta}||g_{r+\delta}||_\infty\int\Gamma_{r+\delta}(u_{r+\delta}-u_r,u_{r+\delta})dm_{r+\delta}\,dr\\&\leq\limsup_{\delta\to0}\frac{1}{\delta}\int_\sigma^{\tau-\delta}e^{L|r+\delta-t|}||g_t||_\infty\Big(\int\Gamma_{r+\delta}(u_{r+\delta}-u_r,u_{r+\delta})dm_{r+\delta}-\int\Gamma_{r+\delta}(u_{r+\delta}-u_r,u_r)dm_{r+\delta}\Big)\,dr\\&=\limsup_{\delta\to0}\frac{1}{\delta}\int_\sigma^{\tau-\delta}e^{L|r+\delta-t|}||g_t||_\infty\Big(-\int\int_r^{r+\delta}\Delta_qu_qdq\Delta_{r+\delta}u_{r+\delta}dm_{r+\delta}-\int\Gamma_r(u_{r+\delta}-u_r,u_r)dm_r\Big)\,dr\\&=\limsup_{\delta\to0}\Big(\int_{\sigma+\delta}^{\tau}-e^{L|r-t|}||g_t||_\infty\int\frac{1}{\delta}\int_{r-\delta}^{r}\Delta_qu_qdq\Delta_ru_rdm_rdr\\&\qquad+\int_\sigma^{\tau-\delta}e^{L|r+\delta-t|}||g_t||_\infty\int\frac{1}{\delta}\int_r^{r+\delta}\Delta_qu_qdq\Delta_ru_rdm_r\,dr\Big)\\&\qquad=\int_\sigma^\tau e^{L|r-t|}||g_t||_\infty\Big(-\int(\Delta_ru_r)^2dm_r+\int(\Delta_ru_r)^2dm_r\Big)=0.\end{aligned}$$

Combining the previous estimates we get

$$h_\tau-h_\sigma\leq-\frac{2}{N}\int_\sigma^\tau\Big(\int\Delta_ru_r\,g_rdm_r\Big)^2dr,$$

and then

$$-\frac{2}{N}\int_\sigma^\tau\Big(\int\Delta_ru_r\,g_rdm_r\Big)^2dr\geq\int_\sigma^\tau\Big[-2\mathbf{\Gamma}_{2,r}(u_r)(g_r)+\int\overset{\bullet}{\Gamma}_r\,(u_r)\,g_r\,m_r\Big]\,dr,$$

which proves the claim. □

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

$$\liminf_{a\to0}\frac{W_\sigma^2(\hat P_{\tau,\sigma}(g_\tau^{\sigma,a}m_\tau),\hat P_{\tau,\sigma}(g_\tau m_\tau))}{2a^2}\geq-\int\frac{1}{2}P_{\tau,\sigma}(\Gamma_\sigma(u_\sigma))g_\tau dm_\tau+\int\Gamma_\tau(u_\tau,u_\sigma)g_\tau dm_\tau.$$

Proof. We denote by Q_a^σ the Hopf-Lax semigroup with respect to the metric d_σ. Note that $aQ_a^\sigma(\phi)=Q_1^\sigma(a\phi)$, so the Kantorovich duality (38) can be written as

$$\frac{W_\sigma^2(\nu_1,\nu_2)}{2a^2}=\frac{1}{a}\sup_\phi\Big[\int Q_a^\sigma\phi d\nu_1-\int\phi d\nu_2\Big].$$

We deduce

$$\begin{aligned}&\frac{W_\sigma^2(\hat P_{\tau,\sigma}(g_\tau^{\sigma,a}m_\tau),\hat P_{\tau,\sigma}(g_\tau m_\tau))}{2a^2}\geq\int\frac{Q_a^\sigma u_\sigma P^*_{\tau,\sigma}(g_\tau^{\sigma,a})-u_\sigma P^*_{\tau,\sigma}g_\tau}{a}dm_s\\&\geq\int\frac{Q_au_\sigma-u_\sigma}{a}P^*_{\tau,\sigma}(g_\tau^{\sigma,a}-g_\tau)dm_\sigma+\int\frac{Q_au_\sigma-u_\sigma}{a}P^*_{\tau,\sigma}g_\tau dm_\sigma+\int u_\sigma\frac{P^*_{\tau,\sigma}(g_\tau^{\sigma,a}-g_\tau)}{a}dm_\sigma.\end{aligned}$$

Note that, since u_s is a Lipschitz function, u_σ is a Lipschitz function as well. Indeed, from the dual representation of the Kantorovich-Rubinstein distance W_s^1 with respect to the metric d_s, we deduce

$$\begin{aligned}|u_\sigma(x)-u_\sigma(y)| &= \left|\int u_s(z)d\hat{P}_{\sigma,s}(\delta_x)(z)-\int u_s(z)d\hat{P}_{t,s}(\delta_y)(z)\right|\\ &\leq \mathrm{Lip}_s(u_s)W_s^1(\hat{P}_{\sigma,s}(\delta_x),\hat{P}_{t,s}(\delta_y)) \leq \mathrm{Lip}_s(u_s)W_s(\hat{P}_{\sigma,s}(\delta_x),\hat{P}_{t,s}(\delta_y))\\ &\leq \mathrm{Lip}_s(u_s)W_\sigma(\delta_x,\delta_y)=\mathrm{Lip}_s(u_s)d_\sigma(x,y),\end{aligned}$$

where the last inequality is a consequence of Theorem 4.9

Since $0 \geq (Q_a^\sigma u_\sigma(x)-u_\sigma(x))/a \geq -2\mathrm{Lip}(u_\sigma)^2$ and $g_\tau^{\sigma,a} \to g_\tau$ in $L^2(X)$ the first integral vanishes. For the second integral we use (37) and estimate by Fatou's Lemma

$$\liminf_{a\to 0}\int \frac{Q_a^\sigma u_\sigma - u_\sigma}{a}P^*_{\tau,\sigma}g_\tau dm_\sigma \geq -\frac{1}{2}\int \mathrm{lip}_\sigma(u_\sigma)^2P^*_{\tau,\sigma}g_\tau dm_\sigma.$$

For the last integral an argument similar to Lemma 3.8 for $H_a^{\tau,g}$ (compare Lemma 4.14 in [6]) yields

$$\lim_{a\to 0}\int \psi_\sigma\frac{P^*_{\tau,\sigma}(g_\tau^{\sigma,a}-g_\tau)}{a}dm_\sigma = \int \Gamma_\tau(P_{\tau,\sigma}u_\sigma,u_\sigma)g_\tau dm_\tau.$$

Combining the last two estimates we obtain

$$\begin{aligned}&\liminf_{a\to 0}\frac{W_\sigma^2(\hat{P}_{\tau,\sigma}(g_\tau^{\sigma,a}m_\tau),\hat{P}_{\tau,\sigma}(g_\tau m_\tau))}{2a^2} \geq -\frac{1}{2}\int \mathrm{lip}_\sigma(u_\sigma)^2P^*_{\tau,\sigma}g_\tau dm_\sigma + \int \Gamma_\tau(P_{\tau,\sigma}u_\sigma,u_\sigma)g_\tau dm_\tau\\ &= -\frac{1}{2}\int \Gamma_\sigma(u_\sigma)P^*_{\tau,\sigma}g_\tau dm_\sigma + \int \Gamma_\tau(P_{\tau,\sigma}u_\sigma,u_\sigma)g_\tau dm_\tau,\end{aligned}$$

where the last inequality follows from our static $\mathrm{RCD}(K,N')$ assumption, which implies Poincaré inequality and doubling property for the static space (X,d_σ,m_σ), and the fact that u_σ is a Lipschitz function (cf. [14]). ☐

Lemma 5.10. *For every $s<\sigma\leq\tau<t$,*

$$\limsup_{a\to 0}\frac{W_\tau^2(g_\tau^{\sigma,a}m_\tau,g_\tau m_\tau)}{2a^2} \leq \frac{1}{2(1-2||\psi_\sigma||_\infty)}\int \Gamma_\tau(u_\sigma)g_\tau dm_\tau.$$

Proof. Let $(Q_a^\tau)_{a\geq 0}$ be the d_τ Hopf-Lax semigroup and fix a bounded Lipschitz function ϕ. Note that

$$\begin{aligned}\partial_a\int Q_a^\tau(\phi)g_\tau^{\sigma,a}dm_\tau &\leq -\int\frac{1}{2}\mathrm{lip}_\tau(Q_a^\tau\phi)^2g_\tau^{\sigma,a}dm_\tau + \int \Gamma_\tau(Q_a^\tau\phi,H_a^{\tau,g}u_\sigma)g_\tau dm_\tau\\ &= \int\left[-\frac{1}{2}\mathrm{lip}_\tau(Q_a^\tau\phi)^2(1+u_\sigma-H_a^{\tau,g}u_\sigma)+\Gamma_\tau(Q_a^\tau\phi,H_a^{\tau,g}u_\sigma)\right]g_\tau dm_\tau,\end{aligned}$$

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

$$\int \Gamma_\tau(Q_a^\tau\phi,H_a^{\tau,g}u_\sigma)g_\tau dm_\tau \leq \sqrt{\mathcal{E}_g(Q_a^\tau\phi)\mathcal{E}_g(H_a^{\tau,g}u_\sigma)} \leq \sqrt{\int \mathrm{lip}_\tau(Q_a^\tau\phi)^2g_\tau dm_\tau\mathcal{E}_g(H_a^{\tau,g}u_\sigma)}.$$

Then, since $1+u_\sigma-H_a^{\tau,g}u_\sigma \geq 1-2||u_\sigma||_\infty$, we obtain using Young's inequality

$$\begin{aligned}\partial_a\int Q_a^\tau(\phi)g_\tau^{\sigma,a}dm_\tau &\leq \frac{1}{2(1-2||u_\sigma||_\infty)}\mathcal{E}_g(H_a^{\tau,g}u_\sigma) \leq \frac{1}{2(1-2||u_\sigma||_\infty)}\mathcal{E}_g(u_\sigma)\\ &= \frac{1}{2(1-2||u_\sigma||_\infty)}\int \Gamma_\tau(u_\sigma)g_\tau dm_\tau.\end{aligned}$$

Integrating over $[0,a]$,

$$\int Q_a^\tau\phi g_\tau^{\sigma,\tau}dm_\tau - \int \phi g_\tau dm_\tau \leq \frac{a}{2(1-2||u_\sigma||_\infty)}\int \Gamma_\tau(u_\sigma)g_\tau dm_\tau,$$

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

$$\frac{W_\tau^2(\nu_1,\nu_2)}{2a^2} = \frac{1}{a}\sup_\phi \left[\int Q_a^\tau \phi d\nu_1 - \int \phi d\nu_2\right]$$

and ϕ was an arbitrary bounded Lipschitz function. □

Lemma 5.11.

$$\liminf_{a\to 0} \int_s^\tau \left[\frac{S_r(\hat{P}_{\tau,r}(g_\tau^{\sigma,a} m_\tau)) - S_r(\hat{P}_{\tau,r}(g_\tau m_\tau))}{a}\right]^2 dr \geq \int_s^\tau \left[\int \Gamma_\tau\big(P_{\tau,r}(\log g_r), u_\sigma\big) g_\tau dm_\tau\right]^2 dr.$$

Proof. With the same estimates as in [12] we have

$$\begin{aligned}&[S_r(\hat{P}_{\tau,r}(g_\tau^{\sigma,a} m_\tau)) - S_r(\hat{P}_{\tau,r}(g_\tau m_\tau))]^2\\ &\geq \frac{1}{(1+\delta)}\left[\int (P_{\tau,r}^*(g_\tau^{\sigma,a}) - g_r)\log g_r dm_r\right]^2 - \frac{1}{\delta}\left[\int \frac{(P_{\tau,r}^* g_\tau^{\sigma,a} - g_r)^2}{g_r} dm_r\right]^2.\end{aligned}$$

Next we apply Jensen's inequality to the convex function $\alpha\colon \mathbb{R}\times\mathbb{R}_+ \to \mathbb{R}\cup\{+\infty\}$ defined by

$$\alpha(r,s) = \begin{cases} 0, & \text{if } r = 0 = s,\\ \frac{r^2}{s}, & \text{if } s \neq 0,\\ +\infty, & \text{if } s = 0 \text{ and } r \neq 0.\end{cases}$$

Recall that the map $dx \mapsto p_{\tau,r}(x,y)dm_\tau(x)$ is not Markovian, but Lemma 2.15 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

$$\begin{aligned}&\int \alpha(P_{\tau,r}^* g_\tau^{\sigma,a} - P_{\tau,r}^* g_\tau, P_{\tau,r}^* g_\tau) dm_r\\ &\leq \int\int \frac{\alpha((g_\tau^{\sigma,a}(x) - g_\tau(x))M_{\tau,r}(y), g_\tau(x)M_{\tau,r}(y))}{M_{\tau,r}} p_{\tau,r}(x,y)dm_\tau(x)dm_r(y)\\ &= \int\int \alpha((g_\tau^{\sigma,a}(x) - g_\tau(x)), g_\tau(x)) p_{\tau,r}(x,y)dm_\tau(x)dm_r(y)\\ &= \int \alpha((g_\tau^{\sigma,a}(x) - g_\tau(x)), g_\tau(x))dm_\tau(x) = \int g_\tau(\psi_\sigma - H_a^{\tau,g}u_\sigma)^2 dm_\tau,\end{aligned}$$

where we applied Jensen's inequality in the second, Fubini in the third, and the definition of $g_\tau^{\sigma,a}$ in the last line. Dividing by a and taking the lim sup we end up with

$$\begin{aligned}&\limsup_{a\to 0} \frac{1}{a}\int \frac{(P_{\tau,r}^* g_\tau^{\sigma,a} - P_{\tau,r}^* g_\tau)^2}{P_{\tau,r}^* g_\tau} dm_r \leq \limsup_{a\to 0}\frac{1}{a}\int g_\tau(u_\sigma - H_a^{\tau,g}u_\sigma)^2 dm_\tau\\ &\leq \limsup_{a\to 0} 2||u_\sigma||_\infty \int g_\tau\left(\frac{H_a^{\tau,g}u_\sigma - u_\sigma}{a}\right) dm_\tau = -2||u_\sigma||_\infty \int g_\tau \Gamma_\tau(u_\sigma, 1) dm_\tau = 0.\end{aligned}$$

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

Since $\delta > 0$ is arbitrary it suffices to show

$$\lim_{a\to 0}\frac{1}{a}\int P_{\tau,r}^*(g(H_a^{\tau,g}u_\sigma - u_\sigma))\log P_{\tau,r}^* g dm_r = \int \Gamma_\tau\big(P_{\tau,r}(\log P_{\tau,r}^* g), u_\sigma\big) g dm_\tau.$$

This, indeed, follows from the fact that $P_{\tau,r}(\log P_{\tau,r}^* g) \in \mathcal{F} = Dom(\mathcal{E}_\tau) = Dom(\mathcal{E}_\tau^g)$ (thanks to uniform boundedness of $P_{\tau,r}^* g$ from above and away from 0) and from the fact that $\frac{1}{a}(H_a^{\tau,g}u_\sigma - u_\sigma) \to \Delta_\tau^g u_\sigma$ weakly in $\mathcal{F}^*$ as $a \searrow 0$, more precisely (cf. Lemma 3.8)

$$\frac{1}{a}\int (H_a^{\tau,g}u_\sigma - u_\sigma)\phi g_\tau dm_\tau \to -\int \Gamma_\tau(u_\sigma, \phi) g_\tau dm_\tau$$

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 presumed that the Bakry-Émery gradient estimate **(III)** holds for the ('primal') heat equation. We will present the argument only in the case $N = \infty$. That is, we now assume that for all $u \in Dom(\mathcal{E})$ and $0 < s < t < T$

$$\Gamma_t(P_{t,s}u) \le P_{t,s}(\Gamma_s(u)) \quad m\text{-a.e. on } X. \tag{78}$$

For the notion of dynamic backward EVI$^{\pm}$-gradient flow we refer to the Appendix.

As in the previous chapters, the assumptions from section 3.1 will always be in force, in particular, we assume the RCD$^*(K, N')$-condition for each static mm-space (X, d_t, m_t) as well as boundedness and L-Lipschitz continuity (in t) for $\log d_t(x, y)$ and (in t and x) for $f_t(x)$.

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

$$\vartheta(a) = (1-a)s + ta \tag{79}$$

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*

$$\mathcal{A}_{s,t}(\mu) = \lim_{h\to 0} \sup \Big\{ \sum_{i=1}^n (a_i - a_{i-1})^{-1} W^2_{\vartheta(a_{i-1})}(\mu^{a_{i-1}}, \mu^{a_i}) \Big| \; 0 = a_0 < \cdots < a_n = 1, a_i - a_{i-1} \le h \Big\}.$$

For two probability measures $\mu, \nu \in \mathcal{P}(X)$ we define

$$W^2_{s,t}(\mu, \nu) = \inf \Big\{ \mathcal{A}_{s,t}(\mu) \Big| \mu \in AC^2([0,1], \mathcal{P}(X)) \text{ with } \mu_0 = \mu, \mu_1 = \nu \Big\}.$$

Lemma 6.2. *The following holds true.*

i) The action $\mu \mapsto \mathcal{A}_{s,t}(\mu)$ is lower semicontinuous, i.e. if $\mu^a_j \to \mu^a$ for every a as $j \to \infty$ we have

$$\mathcal{A}_{s,t}(\mu) \le \liminf_{j\to\infty} \mathcal{A}_{s,t}(\mu_j).$$

ii) For every absolutely continuous curve μ

$$\mathcal{A}_{s,t}(\mu) = \lim_{h\to 0} \inf \Big\{ \sum_{i=1}^n (a_i - a_{i-1})^{-1} W^2_{\vartheta(a_{i-1})}(\mu^{a_{i-1}}, \mu^{a_i}) | 0 = a_0 < \cdots < a_n = 1, a_i - a_{i-1} \le h \Big\}.$$

Proof. Since $\mu^j_a \to \mu_a$ for every $a \in [0,1]$ in the Wasserstein sense we have for every partition $0 = a_0 < \cdots < a_n = 1$

$$\sum_{i=1}^n (a_i - a_{i-1})^{-1} W^2_{\vartheta(a_{i-1})}(\mu^{a_{i-1}}, \mu^{a_i}) = \lim_{j\to\infty} \sum_{i=1}^n (a_i - a_{i-1})^{-1} W^2_{\vartheta(a_{i-1})}(\mu_j^{a_{i-1}}, \mu_j^{a_i}),$$

and hence

$$\sum_{i=1}^n (a_i - a_{i-1})^{-1} W^2_{\vartheta(a_{i-1})}(\mu^{a_{i-1}}, \mu^{a_i}) \le \liminf_{j\to\infty} \mathcal{A}_{s,t}(\mu_j).$$

Taking the supremum over each partition and letting $h \to 0$ proves

$$\mathcal{A}_{s,t}(\mu) \le \liminf_{j\to\infty} \mathcal{A}_{s,t}(\mu_j).$$

We prove the second assertion by contradiction. Assume that there exists a sequence $h_j \to 0$, and a partition $0 = a_0^j < \cdots < a_{n^j}^j = 1$ such that

$$a_i^j - a_{i-1}^j \le h \quad \text{and} \quad \lim_{j\to\infty} \sum_{i=1}^n (a_i^j - a_{i-1}^j)^{-1} W^2_{\vartheta(a_{i-1}^j)}(\mu^{a_{i-1}^j}, \mu^{a_i^j}) < \mathcal{A}_{s,t}(\mu).$$

For every $j \in \mathbb{N}$ we define the curve $(\mu_j^a)_{a\in[0,1]}$ by

$$\mu_j^a = \mu^a_{a_{i-1}^j, a_i^j}, \text{ if } a \in [a_{i-1}^j, a_i^j],$$

where $(\mu^a_{a_{i-1}^j, a_i^j})_{a\in[a_{i-1}^j, a_i^j]}$ denotes the $W_{\vartheta(a_{i-1}^j)}$-geodesic connecting $\mu^{a_{i-1}^j}$ and μa_i^j. Note that for every partition $\{\bar{a}_i\}_{i=1}^N$ with $\bar{a}_i - \bar{a}_{i-1} \ll h_j$

$$\sum_{i=1}^N (\bar{a}_i - \bar{a}_{i-1})^{-1} W^2_{\vartheta(\bar{a}_{i-1})}(\mu_j^{\bar{a}_i}, \mu_j^{\bar{a}_{i-1}}) \le e^{2Lh_j} \sum_{i=1}^n (a_i^j - a_{i-1}^j)^{-1} W^2_{\vartheta(a_{i-1}^j)}(\mu^{a_i^j}, \mu^{a_{i-1}^j}),$$

since for every $a_{i-1}^j \le \bar{a}_{k-1} < \bar{a}_k \le a_i^j$

$$W^2_{\vartheta(a_{i-1}^j)}(\mu_j^{\bar{a}_k}, \mu_j^{\bar{a}_{k-1}}) \le \frac{(\bar{a}_k - \bar{a}_{k-1})^2}{(a_i^j - a_{i-1}^j)^2} W^2_{\vartheta(a_{i-1}^j)}(\mu^{a_{i-1}^j}, \mu^{a_i^j}).$$

Hence

$$\mathcal{A}_{s,t}(\mu_j) \le e^{2Lh_j} \sum_{i=1}^n (a_i^j - a_{i-1}^j)^{-1} W^2_{\vartheta(a_{i-1}^j)}(\mu^{a_i^j}, \mu^{a_{i-1}^j}).$$

This is a contradiction since $\mu_j^a \to \mu_a$ for every a and hence

$$\liminf_{j\to\infty} \mathcal{A}_{s,t}(\mu_j) \ge \mathcal{A}_{s,t}(\mu).$$

□

Proposition 6.3. *For $s < t \in I$ and $\mu^0, \mu^1 \in \mathcal{P}$ we have*

$$W^2_{s,t}(\mu_0, \mu_1) = \inf\left\{ \int_0^1 |\dot{\mu}^a|^2_{s+a(t-s)} da \right\} \tag{80}$$

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

Proof. Choose an arbitrary partition $0 = a_0 < a_1 < \cdots < a_n = 1$ with $a_i - a_{i-1} \le h$. Let $(\mu^a)_{a\in[0,1]} \in AC^2([0,1], \mathcal{P}(X))$. Then, from the absolute continuity of (μ^a), and the log Lipschitz property (34) we deduce

$$\begin{aligned}\sum_{i=1}^n (a_i - a_{i-1})^{-1} W^2_{\vartheta(a_{i-1})}(\mu^{a_{i-1}}, \mu^{a_i}) &\le \sum_{i=1}^n (a_i - a_{i-1})^{-1} \left(\int_{a_i}^{a_{i-1}} |\dot{\mu}^a|_{\vartheta(a_{i-1})} da \right)^2 \\ &\le \sum_{i=1}^n \int_{a_i}^{a_{i-1}} |\dot{\mu}^a|^2_{\vartheta(a_{i-1})} da \\ &\le e^{2Lh} \int_0^1 |\dot{\mu}^a|^2_{\vartheta(a)} da.\end{aligned}$$

Taking the supremum over all partitions and letting $h \to 0$ we obtain

$$\mathcal{A}_{s,t}(\mu) \le \int_0^1 |\dot{\mu}^a|^2_{\vartheta(a)} da,$$

and consequently

$$W^2_{s,t}(\mu_0, \mu_1) \le \inf\left\{ \int_0^1 |\dot{\mu}^a|^2_{s+a(t-s)} da \right\}.$$

To verify the other inequality, we fix again a curve $(\mu_a)_{a\in[0,1]} \in AC^2([0,1],\mathcal{P}(X))$ with finite energy $\mathcal{A}_{s,t}(\mu)$. For each $h>0$ we consider the partition $0=a_0<a_1<\cdots<a_n\le 1<a_{n+1}$ with $a_i=ih$ and $nh\le 1$. We extend μ_a by μ_1 whenever $a>1$. We define μ_a^h to be the $W_{\vartheta(a_{i-1})}$-geodesic connecting $\mu_{a_{i-1}}$ with μ_{a_i} whenever $a\in[a_{i-1},a_i]$. Then we clearly have that $\mu^h\in AC^2([0,1],\mathcal{P}(X))$ and since μ is absolutely continuous, for each $a\in[0,1]$, $\mu_a^h\to\mu_a$ in $(\mathcal{P}(X),W)$. Note that $|\dot\mu_a^h|_{\vartheta(a)}$ is a uniformly bounded function in $L^2([0,1])$

$$\int_0^1 |\dot\mu_a^h|^2_{\vartheta(a)}da \le e^{2Lh}\sum_{i=1}^{n+1}\int_{a_{i-1}}^{a_i}|\dot\mu_a^h|^2_{\vartheta(a_{i-1})}da$$
$$\le e^{2Lh}\sum_{i=1}^{n+1}(a_i-a_{i-1})^{-1}W^2_{\vartheta(a_{i-1})}(\mu_{a_{i-1}},\mu_{a_i})<\infty,$$

since μ_a^h is a piecewise geodesic and $\mathcal{A}_{s,t}(\mu)<\infty$. Then, by the Banach-Alaoglu Theorem there exists a subsequence (not relabeled) $h\to 0$, and a function $A\in L^2([0,1])$ such that $|\dot\mu^h|_{\vartheta(.)}\rightharpoonup A$ in $L^2([0,1])$. Hence from the convergence of $\mu_a^h\to\mu_a$ we get

$$W_{\vartheta(a)}(\mu_a,\mu_{a+\delta}) = \lim_{h\to 0}W_{\vartheta(a)}(\mu_a^h,\mu_{a+\delta}^h)$$
$$\le \liminf_{h\to 0}\int_a^{a+\delta}|\dot\mu_b|_{\vartheta(a)}db \le \liminf_{h\to 0} e^{\delta(t-s)}\int_a^{a+\delta}|\dot\mu_b|_{\vartheta(b)}db$$
$$= e^{\delta(t-s)}\int_a^{a+\delta}A(b)db,$$

and hence

$$|\dot\mu_a|_{\vartheta(a)}\le A(a) \text{ for a.e. } a\in[0,1].$$

Consequently,

$$\int_0^1|\dot\mu_a|^2_{\vartheta(a)}da\le\int_0^1 A^2(a)da\le\liminf_{h\to 0}\int_0^1|\dot\mu_a^h|^2_{\vartheta(a)}da$$
$$\le \liminf_{h\to 0}e^{2Lh}\sum_{i=1}^{n+1}\int_{a_{i-1}}^{a_i}|\dot\mu_a^h|^2_{\vartheta(a_{i-1})}da\le\liminf_{h\to 0}e^{2Lh}\sum_{i=1}^{n+1}(a_i-a_{i-1})^{-1}W^2_{\vartheta(a_{i-1})}(\mu_{a_{i-1}},\mu_{a_i})$$
$$\le\mathcal{A}_{s,t}(\mu),$$

which proves the claim.

□

To conclude this section we define a dynamic 'dual distance' inspired by the dual formulation of the Kantorovich distance. We introduce the function space HLS_ϑ defined by

$$HLS_\vartheta := \left\{\varphi\in \mathrm{Lip}_b([a_0,a_1]\times X)\,\middle|\, \partial_a\varphi_a\le -\frac{1}{2}\Gamma_{\vartheta(a)}(\varphi_a)\quad L^1\times m \text{ a.e. in } (a_0,a_1)\times X\right\}.$$

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

$$\int\phi\varphi_{a_1}dm-\int\phi\varphi_{a_0}dm\le-\frac{1}{2}\int_{a_0}^{a_1}\int\phi\Gamma_{\vartheta(a)}(\varphi_a)dmda.$$

Definition 6.4. *Let $s<t$ and let $\vartheta\colon[a_0,a_1]\to[s,t]$ denote the linear interpolation. Define for two probability measures μ_0,μ_1*

$$\tilde W^2_\vartheta(\mu_0,\mu_1):=2\sup_\varphi\left\{\int\varphi_{a_1}d\mu_1-\int\varphi_{a_0}d\mu_0\right\},$$

where the supremum runs over all maps $\varphi(a,x)=\varphi_a(x)\in HLS_\vartheta$.

Note that $\tilde W_\vartheta$ does not necessarily define a distance. It does not even have to be symmetric. The next Lemma collects two essential properties of $\tilde W_\vartheta$.

Lemma 6.5. *The following holds true.*

(1) $\tilde W_\vartheta$ is lower semicontinuous with respect to the weak-*topology on $\mathcal{P}(X)\times\mathcal{P}(X)$.
(2) For every μ_0,μ_1

$$W_s^2(\mu_0,\mu_1)\le e^{2L|s-t|}(a_1-a_0)\tilde W_\vartheta^2(\mu_0,\mu_1). \tag{81}$$

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

$$\frac12\tilde W_\vartheta(\mu_0,\mu_1)\le\int\varphi_{a_1}d\mu_1-\int\varphi_{a_0}d\mu_0-\varepsilon,$$

where $\varepsilon>0$. Let $\mu_0^n\to\mu_0$, $\mu_1^n\to\mu$ be two sequences converging in duality with continuous bounded functions on X. then, since φ_{a_1} and φ_{a_0} belong to $\mathcal{C}_b(X)$,

$$\begin{aligned}\frac12\tilde W_\vartheta(\mu_0,\mu_1)&\le\int\varphi_{a_1}d\mu_{a_1}-\int\varphi_{a_0}-\varepsilon\\&=\lim_{n\to\infty}\left\{\int\varphi_{a_1}d\mu_1^n-\int\varphi_{a_0}d\mu_0^n\right\}-\varepsilon\\&\le\frac12\liminf_{n\to\infty}\tilde W_\vartheta(\mu_0^n,\mu_1^n)-\varepsilon.\end{aligned}$$

This proves, since $\varepsilon>0$ was arbitrary, that $\tilde W_\vartheta$ is lower semicontinuous with respect to the weak-*topology on $\mathcal{P}(X)\times\mathcal{P}(X)$. The second statement follows from the Kantorovich duality. Indeed, let $\varphi\in\mathrm{Lip}_b(X)$. As already mentioned above the Hopf-Lax semigroup $\varphi_b:=Q_b^s(\varphi)$ solves

$$\frac{d}{db}\varphi_b\le-\frac12\Gamma_s(\varphi_b)\le-\frac12e^{-2L|s-t|}\Gamma_{(1-b)s+bt}(\varphi_b)\quad L^1\times m\text{ a.e. in}(0,1)\times X. \tag{82}$$

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

$$\frac{d}{da}\tilde\varphi_a\le-\frac12\Gamma_{\vartheta(a)}(\tilde\varphi_a)\text{ in }(a_0,a_1)\times X,$$

and

$$e^{-2L|s-t|}(a_1-a_0)^{-1}\left(\int\varphi_1d\mu_1-\int\varphi_0d\mu_0\right)=\int\tilde\varphi_{a_1}d\mu_1-\int\tilde\varphi_{a_0}d\mu_0.$$

Hence

$$e^{-2L|s-t|}(a_1-a_0)^{-1}\left(\int\varphi_1d\mu_1-\int\varphi_0d\mu_0\right)\le\frac12\tilde W_\vartheta^2(\mu_0,\mu_1).$$

Taking the supremum among all φ the Kantorovich duality for the metric W_s implies

$$W_s^2(\mu_0,\mu_1)\le e^{2L|s-t|}(a_1-a_0)\tilde W_\vartheta^2(\mu_0,\mu_1).$$

□

Proposition 6.6. *Let $\vartheta\colon[0,1]\to[s,t]$ be the linear interpolation. Then we have $\tilde W_\vartheta\le W_{s,t}$.*

Proof. Fix $\varphi\in HJS_\vartheta$ and $(\mu)_{a\in[0,1]}$ 2-absolutely continuous curve. We subdivide $[0,1]$ into l intervals $[(k-1)/l,k/l]$ of length $\frac1l$. On each interval $[(k-1)/l,k/l]$ we approximate $(\mu_a)_{|[(k-1)/l,k/l]}$ by regular curves $(\rho_a^{n,k})_{a\in[(k-1)/l,k/l]}$. Obviously, for each k,n the map $[(k-1)/l,k/l]\ni a\mapsto\int\varphi_a d\rho_a^{k,n}$ is absolutely continuous;

$$\int\varphi_{a+h}d\rho_{a+h}-\int\varphi_ad\rho_a\le\mathrm{Lip}(\varphi_{a+h})W(\rho_{a+h},\rho_a)+||\varphi_{a+h}-\varphi_a||_\infty.$$

Let $u_a^{k,n}$ be the density of the regular curve $\rho_a^{k,n}$. Hence for fixed k,n

$$\frac{d}{da}\int\varphi_au_a^{k,n}dm\le\int\varphi_a\dot u_a^{k,n}dm-\frac12\int u_a^{k,n}\Gamma_{\vartheta(a)}(\varphi_a)dm$$

From Lemma 84 we deduce

$$\int\dot u_a^{k,n}\varphi_adm\le\frac12|\dot\rho_a^{k,n}|^2_{\vartheta(k-1/l)}+\frac12\int(\mathrm{lip}_{\vartheta(k-1/l)}\varphi_a)^2d\rho_a^{k,n}.$$

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

$$e^{-L\frac{|t-s|}{l}}(\mathrm{lip}_{\vartheta(k-1/l)}(\varphi_a))^2\le\Gamma_{\vartheta(a)}(\varphi_a)\qquad m\text{ a.e.},$$

we obtain

$$\begin{aligned}&\int \varphi_{k/l}u^{k,n}_{k/l}dm - \int \varphi_{k-1/l}u^{k,n}_{k-1/l}dm\\ &\leq \frac{1}{2}\int_{k-1/l}^{k/l}|\dot\rho^{k,n}_a|^2_{\vartheta(k-1/l)}da + \frac{1}{2}(1-e^{-L\frac{|t-s|}{l}})\int_{k-1/l}^{k/l}\int(\mathrm{lip}_{\vartheta(k-1/l)}\varphi_a)^2 d\rho^{k,n}_a da\\ &\leq \frac{1}{2}\int_{k-1/l}^{k/l}|\dot\rho^{k,n}_a|^2_{\vartheta(k-1/l)}da + \frac{C_1}{2l}(1-e^{-L\frac{|t-s|}{l}})\end{aligned}$$

Taking the limit $n\to\infty$ (and taking the scaling into account) gives

$$\int \varphi_{k/l}d\mu_{k/l} - \int \varphi_{k-1/l}d\mu_{k-1/l} \leq \frac{1}{2}lW^2_{\vartheta(k-1/l)}(\mu_{k-1/l},\mu_{k/l}) + \frac{C_1}{2l}(1-e^{-L\frac{|t-s|}{l}}).$$

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

$$\int \varphi_1 d\mu_1 - \int \varphi_0 d\mu_0 \leq \frac{1}{2}\sum_{k=1}^{l} lW^2_{\vartheta(k-1/l)}(\mu_{k-1/l},\mu_{k/l}) + \frac{C_1}{2}(1-e^{-L\frac{|t-s|}{l}}).$$

Letting $l\to\infty$ we obtain the desired estimate. □

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*

$$W_{s,t}(\mu_0,\mu_1) = \tilde W_\vartheta(\mu_o,\mu_1).$$

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

For this let $(\varphi_a)\in HLS_\vartheta$, and (μ_a) an absolutely continuous curve connecting μ_0 and μ_1.

Consider the Partition $0=a_0<a_1<\dots a_n=1$ with $a_i-a_{i-1}\leq h$ for some $h>0$. Set

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

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

$$\tilde W^2_{\vartheta_i}(\mu_{a_{i-1}},\mu_{a_i}) \leq 2\left\{\int \varphi_{a_i}d\mu_{a_i} - \int \varphi_{a_{i-1}}d\mu_{a_{i-1}}\right\}.$$

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

$$\begin{aligned}\sum_{i=1}^n (a_i-a_{i-1})^{-1}W^2_{\vartheta(a_{i-1})}(\mu_{a_{i-1}},\mu_{a_i}) &\leq e^{2Lh|s-t|}\sum_{i=1}^n \tilde W^2_{\vartheta_i}(\mu_{a_{i-1}},\mu_{a_i})\\ &\leq 2e^{2Lh|s-t|}\sum_{i=1}^n\left\{\int\varphi_{a_i}d\mu_{a_i} - \int\varphi_{a_{i-1}}d\mu_{a_{i-1}}\right\}\\ &= 2e^{2Lh|s-t|}\left\{\int\varphi_1 d\mu_1 - \int\varphi_0 d\mu_0\right\},\end{aligned}$$

where we made use of Lemma 6.5(ii) in the first inequality. Taking the supremum over all $(\varphi_a)\in HLS_\vartheta$ we deduce

$$\sum_{i=1}^n (a_i-a_{i-1})^{-1}W^2_{\vartheta(a_{i-1})}(\mu_{a_{i-1}},\mu_{a_i}) \leq e^{2Lh|s-t|}\tilde W^2_\vartheta(\mu_0,\mu_1), \tag{83}$$

We conclude

$$W^2_{s,t}(\mu_0,\mu_1)\leq \tilde W^2_\vartheta(\mu_0,\mu_1),$$

from taking the supremum in (83) over the partition $0=a_0<a_1<\dots<a_n=1$ with $a_i-a_{i-1}<h$ and subsequently letting $h\searrow 0$.

□

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

$$\left| \int \dot{u}^a \varphi dm_t \right| \leq \frac{1}{2} |\dot{\rho}^a|_t^2 + \frac{1}{2} \int \Gamma_t(\varphi) d\rho^a. \tag{84}$$

Actually, we have inequality (84) for each $\varphi \in Dom(\mathcal{E})$ since we assume that each (X, d_t, m_t) is a static RCD(K, ∞) which implies that Lipschitz functions are dense in the domain of the quadratic form $\mathcal{E}$ with respect to the norm $\sqrt{||\varphi||^2 + \mathcal{E}(\varphi)}$ (Proposition 4.10 in [5]).

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* $a \in (s,t)$

$$\frac{1}{h}(u_{a+h} - u_a) \to \partial_a u_a \text{ weakly}^* \text{ in } \mathcal{F}^*,$$

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

$$\int \frac{1}{h}(u_{a+h} - u_a) v dm_\diamond \to \langle \partial_a u_a, v \rangle.$$

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

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

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

$$\int u_{a+h} v dm_\diamond - \int u_a v dm_\diamond = \int_a^{a+h} \langle \partial_b u_b, v \rangle db.$$

Since $b \mapsto \langle \partial_b u_b, v \rangle$ is in $L^1(s,t)$ we apply the Lebesgue differentiation theorem and obtain that for almost every $a \in (s,t)$

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

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

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

Note that the term on the left-hand side converges to $\int_0^1 \int (u \circ \vartheta)(a) \partial_a \psi_a dm_\diamond da$ as $n \to \infty$ since

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

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

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

and consequently as $n \to \infty$

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

which is the assertion. □

For the following lemmas let $(\rho_a)_{a\in[0,1]}$ be a regular curve and let $\vartheta\colon[0,1]\to[0,\infty)$

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

Set $\rho_{a,\vartheta}:=\hat{P}_{t,\vartheta(a)}(\rho_a)=u_{a,\vartheta}m_{\vartheta(a)}$.

Lemma 6.9. *The curve $(u_{a,\vartheta})_{a\in[0,1]}$ belongs to $\mathrm{Lip}([0,1],\mathcal{F}^*)$ with $u_{a,\vartheta}\in L^2([0,1]\to\mathcal{F})$ and distributional derivative $\partial_a u_{a,\vartheta}\in L^\infty([0,1]\to\mathcal{F}^*)$ satisfying*

$$\partial_a u_{a,\vartheta}=-(t-s)\Delta_{\vartheta(a)}u_{a,\vartheta}+\partial_a f_{\vartheta(a)}u_{a,\vartheta}-P^*_{t,\vartheta(a)}(\dot{u}_a).$$

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

$$\int_0^1||u_{a,\vartheta}||^2_{L^2(m_{\vartheta(a)})}da\le e^{L(t-s)}\int_0^1||u_a||^2_{L^2(m_t)}da\le Re^{L(t-s)}\int_0^1||u_a||_{L^1(m_t)}da=Re^{L(t-s)},$$

and by Theorem 2.12

$$\begin{aligned}\int_0^1\mathcal{E}_{\vartheta(a)}(u_{a,\vartheta})da&\le e^{3L(t-s)}\int[\mathcal{E}_t(u_a)+||u_a||^2_{L^2(m_t)}]da\\&\le e^{3L(t-s)}\sqrt{R}[\int_0^1 2\mathcal{E}_t(\sqrt{u_a})da+R]\le e^{3L(t-s)}\sqrt{R}(2E+R).\end{aligned}$$

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

Next we show that $(u_{a,\vartheta})$ is contained in $\mathrm{Lip}([0,1],\mathcal{F}^*)$. For this let $\psi\in\mathcal{F}$. Then, for almost every $a_0,a_1\in(0,1)$, we obtain with Lemma 6.8, since $P^*_{t,\vartheta(a)}u_{a_0}\in\mathcal{F}_{(0,1)}$,

$$\begin{aligned}&\int\psi u_{a_1,\vartheta}dm_\diamond-\int\psi u_{a_0,\vartheta}dm_\diamond\\&=\int\psi(P^*_{t,\vartheta(a_1)}u_{a_0}-P^*_{t,\vartheta(a_0)}u_{a_0})dm_\diamond+\int\psi P^*_{t,\vartheta(a_1)}(u_{a_1}-u_{a_0})dm_\diamond\\&=(t-s)\int_{a_0}^{a_1}\mathcal{E}^\diamond_{\vartheta(a)}(P^*_{t,\vartheta(a)}u_{a_0},\psi)da+(t-s)\int_{a_0}^{a_1}\int\dot{f}_{\vartheta(a)}P^*_{t,\vartheta(a)}u_{a_0}\psi dm_\diamond da\\&\quad+\int P_{t,\vartheta(a_1)}(\psi e^{f_{\vartheta(a_1)}})(u_{a_1}-u_{a_0})dm_t\\&\le(t-s)\int_{a_0}^{a_1}\mathcal{E}_{\vartheta(a)}(P^*_{t,\vartheta(a)}u_{a_0})^{1/2}\mathcal{E}_{\vartheta(a)}(\psi e^{f_{\vartheta(a)}})^{1/2}da\\&\quad+(t-s)\int_{a_0}^{a_1}||\dot{f}_{\vartheta(a)}||_\infty||P^*_{t,\vartheta(a)}u_{a_0}||_{L^2(m_{\vartheta(a)})}||\psi e^{f_{\vartheta(a)}}||_{L^2(m_\diamond)}da\\&\quad+||e^{-f_t}||_\infty\mathcal{E}_\diamond(P_{t,\vartheta(a_1)}(\psi e^{f_{\vartheta(a_1)}}))^{1/2}\sup_a||\dot{u}_a||_{\mathcal{F}^*}(a_1-a_0)\\&\le(t-s)\mathcal{E}_{\vartheta(a)}(\psi)^{1/2}\int_{a_0}^{a_1}\mathrm{Lip}(f_{\vartheta(a)})\mathcal{E}_{\vartheta(a)}(P^*_{t,\vartheta(a)}u_{a_0})^{1/2}da\\&\quad+(t-s)\int_{a_0}^{a_1}||\dot{f}_{\vartheta(a)}||_\infty||P^*_{t,\vartheta(a)}u_{a_0}||_{L^2(m_{\vartheta(a)})}||\psi e^{f_{\vartheta(a)}}||_{L^2(m_\diamond)}da\\&\quad+||e^{-f_t}||_\infty\mathcal{E}_\diamond(P_{t,\vartheta(a_1)}(\psi e^{f_{\vartheta(a_1)}}))^{1/2}\sup_a||\dot{u}_a||_{\mathcal{F}^*}(a_1-a_0).\end{aligned}$$

Due to our assumptions on f we have that

$$\mathrm{Lip}(f_{\vartheta(a)})\le C,\ ||\dot{f}_{\vartheta(a)}||_\infty\le L,\ ||f_t||_\infty\le C,$$

while the energy estimate Theorem 2.12 and Corollary 2.15 yields

$$\begin{aligned}\mathcal{E}_{\vartheta(a)}(P^*_{t,\vartheta(a)}u_{a_0})&\le e^{3L(t-s)}[\mathcal{E}_t(u_{a_0})+||u_{a_0}||^2_{L^2(m_t)}],\\||P^*_{t,\vartheta(a)}u_{a_0}||_{L^2(m_{\vartheta(a)})}&\le e^{L(t-s)/2}||u_{a_0}||_{L^2(m_t)}.\end{aligned}$$

Note that the last two expressions are bounded since u is a regular curve. Moreover from (21), the gradient estimate (78) and Corollary 2.15 we find

$$\mathcal{E}_\diamond(P_{t,\vartheta(a_1)}(\psi e^{f_{\vartheta(a_1)}})) \leq Ce^{L(t-s)}\mathrm{Lip}(e^{f_{\vartheta(a_1)}})^2\mathcal{E}_{\vartheta(a_1)}(\psi)$$

Applying (21) once more we find that there exists a constant λ such that

$$\int \psi u_{a_1,\vartheta} dm_\diamond - \int \psi u_{a_0,\vartheta} dm_\diamond \leq (a_1 - a_0)\lambda||\psi||_{\mathcal{F}}, \tag{85}$$

and thus

$$||u_{a_1} - u_{a_0}||_{\mathcal{F}^*} \leq \lambda.$$

Note also that (85) holds for every a_0, a_1 by approximating with Lebesgue points. This implies the existence of $\partial_a u_{a,\vartheta} \in L^\infty([0,1], \mathcal{F}^*)$ such that

$$\int \psi u_{a_1,\vartheta} dm_\diamond - \int \psi u_{a_0,\vartheta} dm_\diamond = \int_{a_0}^{a_1} \langle \partial_a u_{a,\vartheta}, \psi\rangle_{\mathcal{F}^*,\mathcal{F}} da.$$

Fix $\psi \in \mathrm{Lip}_b(X)$. By a similar calculation as above it ultimately follows that

$$\begin{aligned}
&\lim_{h\to 0} \frac{1}{h}\Big(\int \psi u_{a+h,\vartheta} dm_\diamond - \int \psi u_{a,\vartheta} dm_\diamond\Big) \\
&= (t-s)\mathcal{E}^\diamond_{\vartheta(a)}(P^*_{t,\vartheta(a)} u_a, \psi) + (t-s)\int \dot f_{\vartheta(a)} P^*_{t,\vartheta(a)} u_a \psi dm_\diamond \\
&+ \lim_{h\to 0} \int P_{t,\vartheta(a+h)}(\psi e^{f_{\vartheta(a+h)}}) \frac{(u_{a+h} - u_a)}{h} dm_t
\end{aligned}$$

almost everywhere. To determine the last integral recall that $u \in \mathcal{C}^1([0,1], L^1(X))$. Then since $\psi \in \mathrm{Lip}_b(X)$

$$\begin{aligned}
\lim_{h\to 0} \int P_{t,\vartheta(a+h)}(\psi e^{f_{\vartheta(a+h)}}) \frac{(u_{a+h} - u_a)}{h} dm_t &= \int P_{t,\vartheta(a)}(\psi e^{f_{\vartheta(a)}}) \dot u_a dm_t \\
&= \int (\psi e^{f_{\vartheta(a)}}) P^*_{t,\vartheta(a)} \dot u_a dm_{\vartheta(a)} = \langle P^*_{t,\vartheta(a)} \dot u_a, \psi\rangle_{\mathcal{F}^*,\mathcal{F}}.
\end{aligned}$$

From the Lipschitz continuity of $(u_{a,\vartheta})$ we deduce that for almost every $a \in [0,1]$

$$\langle \partial_a u_{a,\vartheta}, \psi\rangle_{\mathcal{F}^*,\mathcal{F}} = \langle -(t-s)\Delta_{\vartheta(a)} u_{a,\vartheta} + \partial_a f_{\vartheta(a)} u_{a,\vartheta} - P^*_{t,\vartheta(a)}(\dot u_a), \psi\rangle_{\mathcal{F}^*,\mathcal{F}}.$$

We conclude the proof by approximating $\psi \in \mathcal{F}$ with bounded Lipschitz functions. □

Lemma 6.10. *For any map $\varphi \in HLS_\vartheta$ the map $a \mapsto \int \varphi_a d\rho_{a,\vartheta}$ is absolutely continuous and*

$$\begin{aligned}
\int \varphi_1 d\rho_{1,\vartheta} - \int \varphi_0 d\rho_{0,\vartheta} \leq \int_0^1 \Big[-\frac{1}{2}\int \Gamma_{\vartheta(a)}(\varphi_a) d\rho_{a,\vartheta} + \int P_{t,\vartheta(a)}(\varphi_a)\, \partial_a u_a\, dm_t \\
+ (t-s)\int \Gamma_{\vartheta(a)}(\varphi_a, u_{a,\vartheta}) dm_{\vartheta(a)}\Big] da.
\end{aligned}$$

Proof. Let us begin by showing that $a \mapsto \rho_{a,\vartheta}$ is 2-absolutely continuous. Indeed, let $a_0 < a_1$, we have with the equivalence of the gradient estimate (78) and the Wasserstein contraction (65)

$$\begin{aligned}
W_{\vartheta(a_0)}(\rho_{a_0,\vartheta}, \rho_{a_1,\vartheta}) &\leq W_{\vartheta(a_0)}(\hat P_{t,\vartheta(a_0)}\rho_{a_0}, \hat P_{t,\vartheta(a_0)}\rho_{a_1}) + W_{\vartheta(a_0)}(\hat P_{t,\vartheta(a_0)}\rho_{a_1}, \hat P_{t,\vartheta(a_1)}\rho_{a_1}) \\
&\leq W_t(\rho_{a_0}, \rho_{a_1}) + W_{\vartheta(a_0)}(\hat P_{t,\vartheta(a_0)}\rho_{a_1}, \hat P_{t,\vartheta(a_1)}\rho_{a_1}).
\end{aligned}$$

By virtue of Lemma 3.7(iv) we have that $\tilde\rho_a = \hat P_{t,\vartheta(a)}\rho_{a_1} = \tilde u_a m_{\vartheta(a)}$ is in $AC^2([0,1], \mathcal{P}(X))$. This proves that $a \mapsto \rho_{a,\vartheta}$ is 2-absolutely continuous.

To conclude that $a \mapsto \int \varphi_a d\rho_{a,\vartheta}$ is absolutely continuous we write

$$\begin{aligned}
&\int \varphi_{a_1} d\rho_{a_1,\vartheta} - \int \varphi_{a_0} d\rho_{a_0,\vartheta} \\
&= \int (\varphi_{a_1} - \varphi_{a_0}) d\rho_{a_1,\vartheta} + \int \varphi_{a_0} d\rho_{a_1,\vartheta} - \int \varphi_{a_0} d\rho_{a_0,\vartheta} \\
&\leq ||\varphi_{a_1} - \varphi_{a_0}||_\infty + \mathrm{Lip}(\varphi_{a_0}) W(\rho_{a_1,\vartheta}, \rho_{a_0,\vartheta}).
\end{aligned}$$

To compute its derivative we consider difference quotients. Since $\varphi \in \mathrm{Lip}([0,1], L^\infty(X))$ is in HLS_ϑ and $u_{a+h,\vartheta} \to u_{a,\vartheta}$ in $L^1(X)$ we have

$$\lim_{h\to 0} h^{-1} \int (\varphi_{a+h} - \varphi_a) d\rho_{a+h,\vartheta} \leq -\frac{1}{2} \int |\nabla_{\vartheta(a)} \varphi_a|^2 d\rho_{a,\vartheta}. \tag{86}$$

Now we need to determine

$$\lim_{h\to 0} \frac{1}{h} \Big(\int \varphi_a e^{-f_{\vartheta(a)}} (u_{a+h,\vartheta} - u_{a,\vartheta}) dm_\diamond + \int \varphi_a u_{a+h,\vartheta} d(m_{\vartheta(a+h)} - m_{\vartheta(a)}) \Big).$$

The expression on the right hand side clearly converges to

$$-\dot\vartheta(a) \int \varphi_a \dot f_{\vartheta(a)} u_{a,\vartheta} dm_{\vartheta(a)}, \tag{87}$$

while from Lemma 6.9 we deduce

$$\lim_{h\to 0} \int e^{-f_{\vartheta(a)}} \varphi_a \frac{1}{h} (u_{a+h,\vartheta} - u_{a,\vartheta}) dm_\diamond = \langle \partial_a u_{a,\vartheta}, e^{-f_{\vartheta(a)}} \varphi_a \rangle_{\mathcal{F},\mathcal{F}^*},$$

and after inserting

$$\begin{aligned} \langle \partial_a u_a, e^{-f_{\vartheta(a)}} \varphi_a \rangle_{\mathcal{F},\mathcal{F}^*} &= (t-s) \Big(\int \dot f_{\vartheta(a)} u_{a,\vartheta} \varphi_a e^{-f_{\vartheta(a)}} dm_\diamond + \mathcal{E}^\diamond_{\vartheta(a)}(u_{a,\vartheta}, \varphi_a e^{-f_{\vartheta(a)}}) \Big) && (88) \\ &= (t-s) \Big(\int \dot f_{\vartheta(a)} u_{a,\vartheta} \varphi_a dm_{\vartheta(a)} + \int \Gamma_{\vartheta(a)}(u_{a,\vartheta}, \varphi_a) dm_{\vartheta(a)} \Big). && (89) \end{aligned}$$

Then from the absolute continuity of $a \mapsto \int \varphi_a d\rho_{a,\vartheta}$ together with (86), (87) and (89), we obtain

$$\begin{aligned} &\int \varphi_1 d\rho_{1,\vartheta} - \int \varphi_0 d\rho_{0,\vartheta} = \int_0^1 \partial_a \int \varphi_a d\rho_{a,\vartheta} da \\ &\leq \int_0^1 \Big[-\frac{1}{2} \int |\nabla_{\vartheta(a)} \varphi_a|^2 d\rho_{a,\vartheta} + \int P_{t,\vartheta(a)} \varphi_a \dot u_a dm_t - (t-s) \int \varphi_a \dot f_{\vartheta(a)} u_{a,\vartheta} dm_{\vartheta(a)} \\ &\quad + (t-s) \int \dot f_{\vartheta(a)} u_{a,\vartheta} \varphi_a dm_{\vartheta(a)} + (t-s) \int \Gamma_{\vartheta(a)}(u_{a,\vartheta}, \varphi_a) dm_{\vartheta(a)} \Big] da \\ &\leq \int_0^1 \Big[-\frac{1}{2} \int |\nabla_{\vartheta(a)} \varphi_a|^2 d\rho_{a,\vartheta} + \int P_{t,\vartheta(a)} \varphi_a \dot u_a dm_t + (t-s) \int \Gamma_{\vartheta(a)}(u_{a,\vartheta}, \varphi_a) dm_{\vartheta(a)} \Big] da. \end{aligned}$$

□

We regularize the entropy functional by truncating the singularities of the logarithm. Define $e_\varepsilon \colon [0,\infty)$ by setting $e'_\varepsilon(r) = \log(\varepsilon + r) + 1$ and $e_\varepsilon(0) = 0$. Then e_ε is still a convex function and $e'_\varepsilon \in \mathrm{Lip}_b([0,R])$. For any t and $\rho = u m_t \in \mathcal{P}(X)$ we define

$$S^\varepsilon_t(\rho) = \int e_\varepsilon(u) dm_t.$$

Note that for any $\rho \in Dom(S)$ we clearly have $S^\varepsilon(\rho) \to S(\rho)$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 for any $\varepsilon > 0$*

$$\begin{aligned} S^\varepsilon_t(\rho_{1,\vartheta}) - S^\varepsilon_s(\rho_{0,\vartheta}) \geq \int_0^1 \int \dot u_a P_{t,\vartheta(a)}(e'_\varepsilon(u_{a,\vartheta})) dm_t + 4(t-s) \int e''_\varepsilon(u_{a,\vartheta}) \Gamma_{\vartheta(a)}(\sqrt{u_{a,\vartheta}}) d\rho_{a,\vartheta} \\ + (t-s) \int \dot f_{\vartheta(a)} (u_{a,\vartheta} e'_\varepsilon(u_{a,\vartheta}) - e'_\varepsilon(u_{a,\vartheta})) dm_{\vartheta(a)} da. \end{aligned}$$

Proof. From the convexity of e_ε we get for every $a_0, a_1 \in [0,1]$ by virtue of Lemma 6.9

$$\begin{aligned}
&S^\varepsilon_{\vartheta(a_1)}(\rho_{a_1,\vartheta}) - S^\varepsilon_{\vartheta(a_0)}(\rho_{a_0,\vartheta})\\
&= \int e_\varepsilon(u_{a_1,\vartheta}) - e_\varepsilon(u_{a_0,\vartheta}) e^{-f_{\vartheta(a_0)}} dm_\diamond + \int e_\varepsilon(u_{a_1,\vartheta})(e^{-f_{\vartheta(a_1)}} - e^{-f_{\vartheta(a_0)}}) dm_\diamond\\
&\geq \int e'_\varepsilon(u_{a_0,\vartheta})(u_{a_1,\vartheta} - u_{a_0,\vartheta}) e^{-f_{\vartheta(a_0)}} dm_\diamond + \int e_\varepsilon(u_{a_1,\vartheta})(e^{-f_{\vartheta(a_1)}} - e^{-f_{\vartheta(a_0)}}) dm_\diamond\\
&= \int_{a_0}^{a_1} (\langle \partial_a u_{a,\vartheta}, e^{-f_{\vartheta(a_0)}} e'_\varepsilon(u_{a_0,\vartheta})\rangle - \int e_\varepsilon(u_{a_1,\vartheta}) \dot\vartheta(a) \dot f_{\vartheta(a)} e^{-f_{\vartheta(a)}} dm_\diamond) da\\
&= \int_{a_0}^{a_1} (\langle -\dot\vartheta(a)\Delta_{\vartheta(a)} u_{a,\vartheta} + \dot\vartheta(a)\dot f_{\vartheta(a)} u_{a,\vartheta} + P^*_{t,\vartheta(a)}(\dot u_a), e^{-f_{\vartheta(a_0)}} e'_\varepsilon(u_{a_0,\vartheta})\rangle\\
&\quad - \int e_\varepsilon(u_{a_1,\vartheta}) \dot\vartheta(a) \dot f_{\vartheta(a)} e^{-f_{\vartheta(a)}} dm_\diamond) da\\
&= \int_{a_0}^{a_1} (-\dot\vartheta(a) \langle \Delta_{\vartheta(a)} u_{a,\vartheta}, e^{-f_{\vartheta(a_0)}} e'_\varepsilon(u_{a_0,\vartheta})\rangle + \int \dot\vartheta(a) \dot f_{\vartheta(a)} u_{a,\vartheta} e^{-f_{\vartheta(a_0)}} e'_\varepsilon(u_{a_0,\vartheta}) dm_\diamond\\
&\quad + \int P^*_{t,\vartheta(a)}(\dot u_a) e^{-f_{\vartheta(a_0)}} e'_\varepsilon(u_{a_0,\vartheta}) dm_\diamond - \int e_\varepsilon(u_{a_1,\vartheta}) \dot\vartheta(a) \dot f_{\vartheta(a)} e^{-f_{\vartheta(a)}} dm_\diamond) da.
\end{aligned}$$

Now fix $h > 0$ and choose a partition of $[0,1]$ consisting of Lebesgue points $\{a_i\}_{i=0}^n$ such that $0 \leq a_{i+1} - a_i \leq h$. Then

$$\begin{aligned}
&S^\varepsilon_t(\rho_{1,\vartheta}) - S^\varepsilon_s(\rho_{0,\vartheta}) = \sum_{i=1}^n (S^\varepsilon_{\vartheta(a_i)}(\rho_{a_i,\vartheta}) - S^\varepsilon_{\vartheta(a_{i-1})}(\rho_{a_{i-1},\vartheta}))\\
&\geq \sum_{i=1}^n \int_{a_{i-1}}^{a_i} (-\dot\vartheta(a) \langle \Delta_{\vartheta(a)} u_{a,\vartheta}, e^{-f_{\vartheta(a_{i-1})}} e'_\varepsilon(u_{a_{i-1},\vartheta})\rangle + \int \dot\vartheta(a) \dot f_{\vartheta(a)} u_{a,\vartheta} e^{-f_{\vartheta(a_{i-1})}} e'_\varepsilon(u_{a_{i-1},\vartheta}) dm_\diamond\\
&\quad + \int P^*_{t,\vartheta(a)}(\dot u_a) e^{-f_{\vartheta(a_{i-1})}} e'_\varepsilon(u_{a_{i-1},\vartheta}) dm_\diamond - \int e_\varepsilon(u_{a_i,\vartheta}) \dot\vartheta(a) \dot f_{\vartheta(a)} e^{-f_{\vartheta(a)}} dm_\diamond) da\\
&= \int_0^1 (-\dot\vartheta(a) \langle \Delta_{\vartheta(a)} u_{a,\vartheta}, \varsigma^h_a\rangle + \int \dot\vartheta(a) \dot f_{\vartheta(a)} u_{a,\vartheta} \varsigma^h_a dm_\diamond\\
&\quad + \int P^*_{t,\vartheta(a)}(\dot u_a) \varsigma^h_a dm_\diamond - \int \omega^h_a \dot\vartheta(a) \dot f_{\vartheta(a)} e^{-f_{\vartheta(a)}} dm_\diamond) da,
\end{aligned}$$

where

$$\begin{aligned}
\varsigma^h_a &= e^{-f_{\vartheta(a_{i-1})}} e'_\varepsilon(u_{a_{i-1},\vartheta}), \text{ for } a \in (a_{i-1}, a_i]\\
\omega^h_a &= e_\varepsilon(u_{a_i,\vartheta}), \text{ for } a \in (a_{i-1}, a_i].
\end{aligned}$$

Letting $h \to 0$ we obtain

$$\begin{aligned}
\varsigma^h_a &\to e^{-f_{\vartheta(a)}} e'_\varepsilon(u_{a,\vartheta}), \text{ in } L^1(X) \text{ for a.e. } a \in (0,1)\\
\omega^h_a &\to e_\varepsilon(u_{a,\vartheta}), \text{ in } L^1(X) \text{ for a.e. } a \in (0,1),
\end{aligned}$$

and thus from dominated convergence

$$\begin{aligned}
&S_t^\varepsilon(\rho_{1,\vartheta}) - S_s^\varepsilon(\rho_{0,\vartheta})\\
&\geq \limsup_{h\to 0}[\int_0^1 (-\dot\vartheta(a)\langle \Delta_{\vartheta(a)} u_{a,\vartheta}, \varsigma_a^h\rangle + \int \dot\vartheta(a)\dot f_{\vartheta(a)} u_{a,\vartheta}\varsigma_a^h dm_\diamond\\
&\quad + \int P^*_{t,\vartheta(a)}(\dot u_a)\varsigma_a^h dm_\diamond - \int \omega_a^h \dot\vartheta(a)\dot f_{\vartheta(a)} e^{-f_{\vartheta(a)}} dm_\diamond) da]\\
&\geq \limsup_{h\to 0}[\int_0^1 (-\dot\vartheta(a)\langle \Delta_{\vartheta(a)} u_{a,\vartheta}, \varsigma_a^h\rangle da]\\
&\quad + \int_0^1 (\int \dot\vartheta(a)\dot f_{\vartheta(a)} u_{a,\vartheta} e^{-f_{\vartheta(a)}} e'_\varepsilon(u_{a,\vartheta}) dm_\diamond\\
&\quad + \int P^*_{t,\vartheta(a)}(\dot u_a) e^{-f_{\vartheta(a)}} e'_\varepsilon(u_{a,\vartheta}) dm_\diamond - \int e_\varepsilon(u_{a,\vartheta})\dot\vartheta(a)\dot f_{\vartheta(a)} e^{-f_{\vartheta(a)}} dm_\diamond) da.
\end{aligned}$$

To see that $\langle \Delta_{\vartheta(a)} u_{a,\vartheta}, \varsigma_a^h\rangle \to \langle \Delta_{\vartheta(a)} u_{a,\vartheta}, e^{-f_{\vartheta(a)}} e'_\varepsilon(u_{a,\vartheta})\rangle$, recall that from Theorem 2.12 it suffices to show that

$$\varsigma_a^h \to e^{-f_{\vartheta(a)}} e'_\varepsilon(u_{a,\vartheta}) \text{ in } L^2(X).$$

This is a consequence of the boundedness of $u_{a,\vartheta}$ and $f_{\vartheta(a)}$. Then again by dominated convergence we have

$$\begin{aligned}
&S_t^\varepsilon(\rho_{1,\vartheta}) - S_s^\varepsilon(\rho_{0,\vartheta})\\
&\geq \int_0^1 [\dot\vartheta(a)\mathcal{E}^\diamond_{\vartheta(a)}(u_{a,\vartheta}, e^{-f_{\vartheta(a)}} e'_\varepsilon(u_{a,\vartheta})) + \int \dot\vartheta(a)\dot f_{\vartheta(a)} u_{a,\vartheta} e^{-f_{\vartheta(a)}} e'_\varepsilon(u_{a,\vartheta}) dm_\diamond\\
&\quad + \int P^*_{t,\vartheta(a)}(\dot u_a) e^{-f_{\vartheta(a)}} e'_\varepsilon(u_{a,\vartheta}) dm_\diamond - \int e_\varepsilon(u_{a,\vartheta})\dot\vartheta(a)\dot f_{\vartheta(a)} e^{-f_{\vartheta(a)}} dm_\diamond] da\\
&= \int_0^1 [\dot\vartheta(a)\mathcal{E}_{\vartheta(a)}(u_{a,\vartheta}, e'_\varepsilon(u_{a,\vartheta})) + \int \dot\vartheta(a)\dot f_{\vartheta(a)} u_{a,\vartheta} e'_\varepsilon(u_{a,\vartheta}) dm_{\vartheta(a)}\\
&\quad + \int P^*_{t,\vartheta(a)}(\dot u_a) e'_\varepsilon(u_{a,\vartheta}) dm_{\vartheta(a)} - \int e_\varepsilon(u_{a,\vartheta})\dot\vartheta(a)\dot f_{\vartheta(a)} dm_{\vartheta(a)}] da.
\end{aligned}$$

□

6.3. **The Dynamic EVI$^-$-Property.**

Proposition 6.12. *Let $\rho^a = u^a m_t$ be a regular curve. Then setting $\rho_{a,\vartheta} = \hat P_{t,\vartheta(a)}\rho^a$, it holds*

$$\begin{aligned}
&\frac{1}{2}\tilde W_\vartheta^2(\rho_{1,\vartheta}, \rho_{0,\vartheta}) - (t-s)(S_t(\rho_{1,\vartheta}) - S_s(\rho_{0,\vartheta}))\\
&\leq \frac{1}{2}\int_0^1 |\dot\rho_a|_t^2 da - (t-s)^2 \int_0^1 \int \dot f_{\vartheta(a)} d\rho_{a,\vartheta} da.
\end{aligned} \tag{90}$$

Proof. Applying Lemma 6.10 and Lemma 6.11, we find

$$\begin{aligned}
&\int \varphi_1 d\rho_{1,\vartheta} - \int \varphi_0 d\rho_{0,\vartheta} - (t-s)(S_t^\varepsilon(\rho_{1,\vartheta}) - S_s^\varepsilon(\rho_{0,\vartheta}))\\
&\leq \int_0^1 \Big[\int \dot u_a P_{t,\vartheta(a)}(\varphi_a - (t-s)e'_\varepsilon(u_{a,\vartheta})) dm_t\\
&- \frac{1}{2}\int \Gamma_{\vartheta(a)}(\varphi_a) d\rho_{a,\vartheta} + (t-s)\int \Gamma_{\vartheta(a)}(\varphi_a, u_{a,\vartheta}) dm_{\vartheta(a)} - 4(t-s)^2 \int e''_\varepsilon(u_{a,\vartheta})\Gamma_{\vartheta(a)}(\sqrt{u_{a,\vartheta}}) d\rho_{a,\vartheta}\\
&- (t-s)^2 \int (e_\varepsilon(u_{a,\vartheta}) - e'_\varepsilon(u_{a,\vartheta})u_{a,\vartheta})\dot f_{\vartheta(a)} dm_{\vartheta(a)}\Big] da.
\end{aligned} \tag{91}$$

Then since

$$4re''_\varepsilon(r) \geq 4r^2(e''_\varepsilon(r))^2 = r(p'_\varepsilon(\sqrt r))^2,$$

we can estimate

$$-4u_{a,\vartheta}e''_\varepsilon(u_{a,\vartheta})\Gamma_{\vartheta(a)}(\sqrt{u_{a,\vartheta}}) \le -u_{a,\vartheta}(p'_\varepsilon(\sqrt{u_{a,\vartheta}}))^2\Gamma_{\vartheta(a)}(\sqrt{u_{a,\vartheta}}) = -u_{a,\vartheta}\Gamma_{\vartheta(a)}(p_\varepsilon(\sqrt{u_{a,\vartheta}})),$$

and while, with $q_\varepsilon(r) := \sqrt{r}(2-\sqrt{r}p'_\varepsilon(\sqrt{r}))$,

$$\Gamma_{\vartheta(a)}(u_{a,\vartheta},\varphi_a) = 2\sqrt{u_{a,\vartheta}}\Gamma_{\vartheta(a)}(\sqrt{u_{a,\vartheta}},\varphi_a) = u_{a,\vartheta}\Gamma_{\vartheta(a)}(p_\varepsilon(\sqrt{u_{a,\vartheta}}),\varphi_a) + q_\varepsilon(u_{a,\vartheta})\Gamma_{\vartheta(a)}(\sqrt{u_{a,\vartheta}},\varphi_a)$$

we find

$$\begin{aligned}
&\int \varphi_1 d\rho_{1,\vartheta} - \int \varphi_0 d\rho_{0,\vartheta} - (t-s)(S^\varepsilon_t(\rho_{1,\vartheta}) - S^\varepsilon_s(\rho_{0,\vartheta}))\\
&\le \int_0^1 \Big[\int \dot u_a P_{t,\vartheta(a)}(\varphi_a - (t-s)e'_\varepsilon(u_{a,\vartheta}))dm_t\\
&- \frac12\int \Gamma_{\vartheta(a)}(\varphi_a)d\rho_{a,\vartheta} + (t-s)\int \Gamma_{\vartheta(a)}(\varphi_a, p_\varepsilon(\sqrt{u_{a,\vartheta}}))d\rho_{a,\vartheta} - (t-s)^2\int \Gamma_{\vartheta(a)}(p_\varepsilon(\sqrt{u_{a,\vartheta}}))d\rho_{a,\vartheta}\\
&+ (t-s)\int q_\varepsilon(u_{a,\vartheta})\Gamma_{\vartheta(a)}(\sqrt{u_{a,\vartheta}},\varphi_a)dm_{\vartheta(a)} - (t-s)^2\int (e_\varepsilon(u_{a,\vartheta}) - e'_\varepsilon(u_{a,\vartheta})u_{a,\vartheta})\dot f_{\vartheta(a)}dm_{\vartheta(a)}\Big]da.
\end{aligned} \tag{92}$$

Hence, by means of (84), the gradient estimate (78), and Young inequality $2xy \le \delta x^2 + y^2/\delta$ this yields

$$\begin{aligned}
&\int \varphi_1 d\rho_{1,\vartheta} - \int \varphi_0 d\rho_{0,\vartheta} - (t-s)(S^\varepsilon_t(\rho_{1,\vartheta}) - S^\varepsilon_s(\rho_{0,\vartheta}))\\
&\le \int_0^1 \Big[\frac12|\dot\rho_a|^2_t + \frac12\int \Gamma_t(P_{t,\vartheta(a)}(\varphi_a - (t-s)e'_\varepsilon(u_{a,\vartheta}))d\rho_a\\
&- \frac12\int P_{t,\vartheta(a)}\Gamma_{\vartheta(a)}(\varphi_a - (t-s)p_\varepsilon(\sqrt{u_{a,\vartheta}}))d\rho_a + (t-s)\int q_\varepsilon(u_{a,\vartheta})\Gamma_{\vartheta(a)}(\sqrt{u_{a,\vartheta}},\varphi_a)dm_{\vartheta(a)}\\
&- (t-s)^2\int (e_\varepsilon(u_{a,\vartheta}) - e'_\varepsilon(u_{a,\vartheta})u_{a,\vartheta})\dot f_{\vartheta(a)}dm_{\vartheta(a)}\Big]da\\
&\le \int_0^1 \Big[\frac12|\dot\rho_a|^2_t + +(t-s)\int |q_\varepsilon(u_{a,\vartheta})||\Gamma_{\vartheta(a)}(\sqrt{u_{a,\vartheta}},\varphi_a)|dm_{\vartheta(a)}\\
&- (t-s)^2\int (e_\varepsilon(u_{a,\vartheta}) - e'_\varepsilon(u_{a,\vartheta})u_{a,\vartheta})\dot f_{\vartheta(a)}dm_{\vartheta(a)}\Big]da\\
&\le \int_0^1 \Big[\frac12|\dot\rho_a|^2_t + \frac{(t-s)}{2\delta}\int (q_\varepsilon(u_{a,\vartheta}))^2\Gamma_{\vartheta(a)}(\varphi_a)dm_{\vartheta(a)} + \frac{(t-s)\delta}{2}\int \Gamma_{\vartheta(a)}(\sqrt{u_{a,\vartheta}})dm_{\vartheta(a)}\\
&- (t-s)^2\int (e_\varepsilon(u_{a,\vartheta}) - e'_\varepsilon(u_{a,\vartheta})u_{a,\vartheta})\dot f_{\vartheta(a)}dm_{\vartheta(a)}\Big]da.
\end{aligned}$$

We first pass to the limit $\varepsilon \to 0$,

$$\lim_{\varepsilon\to0} q^2_\varepsilon(r) = 0, \quad q^2_\varepsilon(r) = 4r(1-\frac{r}{\varepsilon+r})^2 \le 4r,$$

$$\begin{aligned}
&\lim_{\varepsilon\to0}(e_\varepsilon(r) - re'_\varepsilon(r)) = -r,\\
&\qquad |e_\varepsilon(r) - re'_\varepsilon(r)| \le 2(\varepsilon+r)|\log(\varepsilon+r)| + r + \varepsilon\log\varepsilon \le 2\sqrt{\varepsilon+r} + r + \varepsilon\log\varepsilon,
\end{aligned}$$

and then, $\delta \to 0$,

$$\begin{aligned}
&\int \varphi_1 d\rho_{1,\vartheta} - \int \varphi_0 d\rho_{0,\vartheta} - (t-s)(S_t(\rho_{1,\vartheta}) - S_s(\rho_{0,\vartheta}))\\
&\le \int_0^1 \Big[\frac12|\dot\rho_a|^2_t + (t-s)^2\int \dot f_{\vartheta(a)}d\rho_{a,\vartheta}\Big]da.
\end{aligned}$$

Taking the supremum over φ we obtain the desired estimate (90). □

Theorem 6.13. *Assume that the gradient estimate holds true for the time-dependent metric measure space $(X, d_t, m_t)_{t\in(0,T)}$. Then for every $\mu \in Dom(S)$ and every $\tau \in (0,T]$ the dual heat flow $\mu_t := \hat{P}_{\tau,t}\mu$ emanating in μ we have*

$$S_s(\mu_s) - S_t(\sigma) \leq \frac{1}{2(t-s)}(W_t^2(\mu_t,\sigma) - W_{s,t}^2(\mu_s,\sigma)) - (t-s)\int_0^1 \int \dot{f}_{\vartheta(a)} d\rho_{a,\vartheta} da \tag{93}$$

for all $s \in (0,\tau)$ and all $\sigma, \mu \in Dom(S)$. Here $(\rho_a)_{a\in[0,1]}$ denotes the W_t-geodesic connecting $\rho_0 = \mu_t$, $\rho_1 = \sigma$ and $\rho_{a,\vartheta} = \hat{P}_{t,\vartheta(a)}(\rho_a)$.

In particular μ_t is a dynamic upward EVI$^-$-gradient flow, i.e. for every $t \in (0,\tau)$ and every $\sigma \in Dom(S)$ we have

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

Proof. Let $(\rho_a)_{a\in[0,1]}$ be a W_t-geodesic connecting μ_t and σ, which exists and is unique. We approximate the geodesic $(\rho_a)_{a\in[0,1]}$ by regular curves $(\rho_a^n)_{a\in[0,1]}$. Proposition 6.12 states that for each $(\rho_a^n)_{a\in[0,1]}$

$$\begin{aligned} &\frac{1}{2}\tilde{W}_\vartheta^2(\rho_{1,\vartheta}^n, \rho_{0,\vartheta}^n) - (t-s)(S_t(\rho_{1,\vartheta}^n) - S_s(\rho_{0,\vartheta}^n)) \\ &\leq \frac{1}{2}\int_0^1 |\dot{\rho}_a^n|_t^2 da - (t-s)^2 \int_0^1 \int \dot{f}_{\vartheta(a)} d\rho_{a,\vartheta}^n da. \end{aligned} \tag{94}$$

Since for every $a \in [0,1]$ ρ_a^n converges to ρ_a in duality with bounded continuous functions, $\rho_{a,\vartheta}^n$ converges to $\rho_{a,\vartheta}$ in duality with bounded continuous functions as well. By virtue of Lemma 6.5 we obtain

$$\liminf_{n\to\infty} \tilde{W}_\vartheta^2(\rho_{1,\vartheta}^n, \rho_{0,\vartheta}^n) \geq \tilde{W}_\vartheta^2(\rho_{1,\vartheta}, \rho_{0,\vartheta}).$$

Note that (ρ_a^n) also converges to ρ_a in duality with L^∞ functions, since Lemma 3.2 provides $\sup_n S_t(\rho_a^n) < \infty$. The same argument applies then to $\rho_{a\vartheta}^n$. Hence

$$\lim_{n\to\infty} \int \dot{f}_{\vartheta(a)} d\rho_{a,\vartheta}^n = \int \dot{f}_{\vartheta(a)} d\rho_{a,\vartheta}.$$

Then we end up with

$$\begin{aligned} &\frac{1}{2}\tilde{W}_\vartheta^2(\mu_s,\sigma) - (t-s)(S_t(\sigma) - S_s(\mu_s)) \\ &\leq \frac{1}{2}W_t^2(\mu_t,\sigma) - (t-s)^2 \int_0^1 \int \dot{f}_{\vartheta(a)} d\rho_{a,\vartheta} da. \end{aligned} \tag{95}$$

Applying Corollary 6.7 we obtain

$$(t-s)(S_s(\mu_s) - S_t(\sigma)) \leq \frac{1}{2}W_t^2(\mu_t,\sigma) - \frac{1}{2}W_{s,t}^2(\mu_s,\sigma) - (t-s)^2 \int_0^1 \int \dot{f}_{\vartheta(a)} d\rho_{a,\vartheta} da.$$

Dividing by $t-s$ and letting $s \nearrow t$ we find

$$\begin{aligned} S_t(\mu_t) - S_t(\sigma) &\leq \liminf_{s\nearrow t} \frac{1}{2(t-s)} \left(W_t^2(\mu_t,\sigma) - W_{s,t}^2(\mu_s,\sigma)\right) \\ &= \frac{1}{2}\partial_s^- W_{s,t}^2(\mu_s,\sigma)_{|s=t-}. \end{aligned}$$

□

6.4. **Summarizing.** The precise integrated version (93) of the EVI$^-$-property indeed also implies a relaxed version of the EVI$^+$-property which then in turn allows to prove uniqueness of dynamic EVI-flows for the entropy.

Corollary 6.14. *The gradient estimate* **(III)** *implies the* $\mathbf{EVI}^+(-2L,\infty)$*-property. More precisely, for every* $\mu\in Dom(S)$ *and every* $\tau\le T$ *the dual heat flow* $\mu_t:=\hat P_{t,\tau}\mu$ *emanating in* μ *satisfies*

$$\frac12\partial_s^- W^2_{s,t}(\mu_s,\sigma)_{|s=t}\ge S_t(\mu_t)-S_t(\sigma)-L\,W_t^2(\mu_t,\sigma)$$

for all $t<\tau$ *and all* $\sigma\in\mathcal P(X)$.

Proof. Given $\mu_t:=\hat P_{t,\tau}\mu$ for $t\tau$, consider (93) for fixed $s<\tau$ and with $s\searrow t$. Then

$$\begin{aligned}
S_s(\mu_s)-S_s(\sigma) &= \lim_{s\searrow t} S_s(\mu_s)-S_t(\sigma)\\
&\le \lim_{s\searrow t}\frac{1}{2(t-s)}\Big[W_t^2(\mu_t,\sigma)-W^2_{s,t}(\mu_s,\sigma)\Big]\\
&\le \Big(\lim_{s\searrow t}\frac{1}{2(t-s)}\Big[W^2_{t,s}(\mu_t,\sigma)-W_s^2(\mu_s,\sigma)\Big]\\
&\qquad\qquad +\frac{L}{2}\Big[W_t^2(\mu_t,\sigma)+W_s^2(\mu_s,\sigma)\Big]\Big)\\
&= \frac12\partial_t^- W^2_{t,s}(\mu_t,\sigma)_{t=s+}+L\,W_s^2(\mu_s,\sigma)
\end{aligned}$$

where the last estimate follows from (98). □

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 in some* $\mu\in\mathcal P$*. Then*

$$\mu_t=\hat P_{t,\tau}\mu$$

for all $t\in(\sigma,\tau)$*. That is, the dual heat flow is the unique dynamic backward EVI*$^-$*-flow for the Boltzmann entropy.*

Proof. Corollary 7.8 together with Corollary 6.14 and Theorem 6.13. □

Theorem 6.16. *The gradient estimate* $(\mathbf{III}_N)$ *implies the dynamic* N*-convexity of the Boltzmann entropy* $(\mathbf{I}_N)$.

Proof. According to Theorem 4.7 and Theorem 6.13 the gradient estimate $(\mathbf{III}_N)$ implies both

- the transport estimate $(\mathbf{II}_N)$ and
- the $\mathbf{EVI}^-(0,\infty)$-property

According to Theorem 7.11 and Remark 7.12, both properties together imply dynamic N-convexity. □

7. Appendix

7.1. **Time-dependent Geodesic Spaces.** For this chapter, our basic setting will be a space X equipped with a 1-parameter family of complete geodesic metrics $(d_t)_{t\in I}$ where $I \subset \mathbb{R}$ is a bounded open interval, say for convenience $I = (0,T)$. (More generally, one might allow d_t to be pseudo metrics where the existence of connecting geodesics is only requested for pairs $x, y \in X$ with $d_t(x,y) < \infty$.) We always request that there exists a constant $L \in \mathbb{R}$ ('log-Lipschitz bound') such that

$$\left|\log \frac{d_t(x,y)}{d_s(x,y)}\right| \leq L \cdot |t-s| \tag{96}$$

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 7.1. *Given $s,t \in I$ and $x,y \in X$ we put*

$$d_{s,t}(x,y) := \inf \left\{ \int_0^1 |\dot\gamma^a|^2_{s+a(t-s)} da \right\}^{1/2} \tag{97}$$

where the infimum runs over all absolutely continuous curves $(\gamma^a)_{a\in[0,1]}$ in X connecting x and y.

Proposition 7.2. *(i) The infimum in the above formula is attained. Each minimizer $(\gamma^a)_{a\in[0,1]}$ is a curve of constant speed, i.e. $|\dot\gamma^a|_{s+a(t-s)} = d_{s,t}(x,y)$ for all $a \in [0,1]$.*

(ii) A point $z \in X$ lies on some minimizing curve γ with $z = \gamma^a$ if and only if

$$d_{s,t}(x,y) = d_{s,r}(x,z) + d_{r,t}(z,y)$$

with $r = s + a(t-s)$.

(iii) For all $s,t \in I$ and $x,y \in X$

$$\frac{1-e^{-L|t-s|}}{L|t-s|} \leq \frac{d_{s,t}(x,y)}{d_s(x,y)} \leq \frac{e^{L|t-s|}-1}{L|t-s|}.$$

Thus in particular,

$$\left|\partial_t d_{s,t}(x,y)\big|_{t=s}\right| \leq \frac{L}{2} d_s(x,y). \tag{98}$$

(iv) For all $s < t \in I$ and $x,y \in X$

$$d_{s,t}(x,y) = \lim_{\delta\to 0} \inf_{(t_i,x_i)_i} \left\{ \sum_{i=1}^k \frac{t-s}{t_i - t_{i-1}}\, d^2_{t_i}(x_i, x_{i-1}) \right\}^{1/2} \tag{99}$$

where the infimum runs over all $k \in \mathbb{N}$. all partitions $(t_i)_{i=0,\dots,k}$ of $[s,t]$ with $t_0 = s, t_k = t$ 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 "$\geq$"-inequality. Conversely, given any pair of minimizers for $d_{s,r}(x,z)$ and $d_{r,t}(z,y)$ by concatenation a curve connecting x and y can be constructed with action bounded by the scaled action of the two ingredients. This proves the "$\leq$"-inequality.

(iii) The log-Lipschitz continuity of the distance implies that for each absolutely continuous curve

$$e^{-La|t-s|} \int_0^1 |\dot\gamma^a|_s da \leq \int_0^1 |\dot\gamma^a|_{s+a(t-s)} da \leq e^{La|t-s|} \int_0^1 |\dot\gamma^a|_s da.$$

(iv) see section 6.1 for the argument in the case of $W_{s,t}$. ☐

7.2. **EVI Formulation of Gradient Flows.** For the subsequent discussion, a lower semi-bounded function $V : I \times X \to (-\infty, \infty]$ will be given with $V_s(x) \leq C_0 \cdot V_t(x) + C_1$ for all $s, t \in I$ and $x \in X$ (thus, in particular, $Dom(V) = \{x \in X : V_t(x) < \infty\}$ is independent of x) and such that for each $t \in I$ the function $x \mapsto V_t(x)$ is κ-convex along each d_t-geodesic (for some $\kappa \in \mathbb{R}$). We also assume that minimizing d_t-geodesics between pairs of points in $Dom(V)$ are unique.

In previous chapters, the following results will be applied

- to the Boltzmann entropy S_t on the time-dependent geodesic space $(\mathcal{P}, W_t)_{t \in I}$ as well as
- to the Dirichlet energy $\mathcal{E}_t$ on the time-dependent geodesic space $L^2(X, m_t)_{t \in I}$

in the place of the function V_t on the time-dependent geodesic space $(X, d_t)_{t \in I}$.

Definition 7.3. *Given a left-open interval $J \subset I$, an absolutely continuous curve $(x_t)_{t \in J}$ will be called* dynamic backward EVI$^-$-gradient flow *for V if for all $t \in J$ and all $z \in Dom(V_t)$*

$$\frac{1}{2} \partial_s^- d_{s,t}^2(x_s, z)\Big|_{s=t-} \geq V_t(x_t) - V_t(z) \tag{100}$$

where $d_{s,t}$ is defined in Definition 7.1.

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*

$$\frac{1}{2} \partial_s^- d_{s,t}^2(x_s, z)\Big|_{s=t+} \geq V_t(x_t) - V_t(z)$$

for all $t \in J$.

It is called dynamic backward EVI-gradient flow *if it is both, a* dynamic backward EVI$^+$-gradient flow *and a* dynamic backward EVI$^-$-gradient flow.

We say that the backward gradient flow $(x_t)_{t \in J}$ emanates *in $x' \in X$ if $\lim_{t \nearrow \sup J} x_t = x'$.*

Being a dynamic backward EVI$^{\pm}$-gradient flow for V obviously implies that $x_t \in Dom(V_t)$ for all $t < \tau$.

Remark. Note that these definitions are slightly different from a previous one presented in [51]. If d_s depends smoothly on s then

$$\partial_s^- d_{s,t}^2(x_s, z)\big|_{s=t-} = \partial_s^- d_t^2(x_s, z)\big|_{s=t-} + \partial_s^- d_{s,t}^2(x_t, z)\big|_{s=t-}$$

and always $\partial_s^- d_{s,t}^2(x_t, z)\big|_{s=t-} \geq \mathfrak{b}_t^0(\gamma)$ for any d_t-geodesic γ connecting x_t and z.

Often, we ask for an improved notion of dynamic backward EVI-gradient flows, involving parameters $N \in (0, \infty]$ (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 7.4. *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 $z \in Dom(V_t)$ and all $t \in (\sigma, \tau)$*

$$\frac{1}{2} \partial_s^- d_{s,t}^2(x_s, z)\Big|_{s=t} - \frac{K}{2} \cdot d_t^2(x_t, z) \geq V_t(x_t) - V_t(z) + \frac{1}{N} \int_0^1 \Big(\partial_a V_t(\gamma^a)\Big)^2 (1-a) da \tag{101}$$

where γ denotes the d_t-geodesic connecting x_t and z.

Analogously, we define dynamic backward EVI$^{\pm}(K, N)$-gradient flows *for V.*

In the case, $K = 0$, dynamic backward EVI(K, N)-gradient flows will be simply called dynamic backward EVI$_N$-gradient flows.

The concept of 'backward' gradient flows is tailor-made for our later application to the dual heat flow. This flow is running backward in time and on its way it tries to minimize the Boltzmann entropy. Regarded in positive time direction, it follows the 'upward gradient' of the entropy.

On the other hand, in calculus of variations mostly the 'downward' gradient flow will be considered where a curve tries to follow the negative gradient of a given functional.

Definition 7.5. *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* $z\in Dom(V_t)$ *and all* $t\in(\sigma,\tau)$

$$-\frac{1}{2}\partial_s^+ d_{s,t}^2(x_s,z)\Big|_{s=t}-\frac{K}{2}\cdot d_t^2(x_t,z)\ \geq V_t(x_t)-V_t(z)+\frac{1}{N}\int_0^1\Big(\partial_a V_t(\gamma^a)\Big)^2(1-a)da \tag{102}$$

where γ *denotes the* d_t*-geodesic connecting* x_t *and* z*.*

We say that a forward gradient flow emanates in a given point $x'\in X$ *if* $\lim_{t\searrow\sigma}x_t=x'$*.*

We will formulate all our results for 'backward' gradient flows and leave it to the reader to carry them over to the case of 'forward' gradient flows.

Lemma 7.6. *For each dynamic backward EVI*$^\pm(K,\infty)$*-gradient flow* $(x_t)_{t\in(\sigma,\tau)}$ *for* V

$$\int_\sigma^\tau V_t(x_t)dt<\infty.$$

Proof. Choose $z\in Dom(V)$, apply the EVI(K,∞)-property at time t, and then integrate w.r.t. time t

$$\begin{aligned}\int_\sigma^\tau V_t(x_t)dt &\leq \int_\sigma^\tau\Big[V_t(z)+\frac{1}{2}\partial_s d_{s,t}^2(x_s,z)\big|_{s=t}-\frac{K}{2}d_t^2(x_t,z)\Big]dt\\ &\leq (C_0\,V_\tau(z)+C_1)(\tau-\sigma)+\frac{1}{2}\int_\sigma^\tau\Big[\partial_t d_t^2(x_t,z)+(L-K)\,d_t^2(x_t,z)\Big]dt\\ &= (C_0\,V_\tau(z)+C_1)(\tau-\sigma)+\frac{1}{2}d_\tau^2(x_\tau,z)-\frac{1}{2}d_\sigma^2(x_\sigma,z)+\frac{L-K}{2}\int_\sigma^\tau d_t^2(x_t,z)dt.\end{aligned}$$

Obviously, the right hand side is finite which thus proves the claim. □

7.3. **Contraction Estimates.**

Theorem 7.7. *Given two curves* $(x_t)_{t\in(\sigma,\tau)}$ *and* $(y_t)_{t\in(\sigma,\tau)}$*, one of which is an is a dynamic backward EVI*$^-(K,N)$*-gradient flow for* V *and the other is a dynamic backward EVI*$^+(K,N)$*-gradient flow for* V*, then for all* $\sigma<s<t<\tau$

$$d_s^2(x_s,y_s)\leq e^{-2K(t-s)}\cdot d_t^2(x_t,y_t)-\frac{2}{N}\int_s^t e^{-2K(r-s)}\cdot\Big|V_r(x_r)-V_r(y_r)\Big|^2dr. \tag{103}$$

Proof. Assume that the curve $(x_t)_{t\in(\sigma,\tau]}$ is a dynamic backward EVI$^-$-gradient flow for V and $(y_t)_{t\in(\sigma,\tau]}$ is a dynamic backward EVI$^+$-gradient flow for V. It implies that $r\mapsto d_r(x_r,y_r)$ is absolutely continuous since

$$|d_t(x_t,y_t)-d_s(x_s,y_s)|\leq d_s(x_s,x_t)+d_s(y_s,y_t)+L(t-s)d_t(x_t,y_t).$$

Thus by the very definition of EVI flows

$$
\begin{aligned}
d_t^2(x_t,y_t)-d_s^2(x_s,y_s) &= \limsup_{\delta\searrow 0}\Big[\frac{1}{\delta}\int_{t-\delta}^t d_r^2(x_r,y_r)\,dr-\frac{1}{\delta}\int_s^{s+\delta} d_r^2(x_r,y_r)\,dr\Big]\\
&= \limsup_{\delta\searrow 0}\frac{1}{\delta}\int_{s+\delta}^t \Big[d_r^2(x_r,y_r)-d_{r-\delta}^2(x_{r-\delta},y_{r-\delta})\Big]\,dr\\
&\geq \liminf_{\delta\searrow 0}\frac{1}{\delta}\int_{s+\delta}^t \big[d_r^2(x_r,y_r)-d_{r,r-\delta}^2(x_r,y_{r-\delta})\big]\,dr\\
&\quad+\liminf_{\delta\searrow 0}\frac{1}{\delta}\int_{s+\delta}^t \big[d_{r,r-\delta}^2(x_r,y_{r-\delta})-d_{r-\delta}^2(x_{r-\delta},y_{r-\delta})\big]\,dr\\
&= \liminf_{\delta\searrow 0}\frac{1}{\delta}\int_{s+\delta}^t \big[d_r^2(x_r,y_r)-d_{r,r-\delta}^2(x_r,y_{r-\delta})\big]\,dr\\
&\quad+\liminf_{\delta\searrow 0}\frac{1}{\delta}\int_{s}^{t-\delta} \big[d_{r+\delta,r}^2(x_{r+\delta},y_r)-d_r^2(x_r,y_r)\big]\,dr\\
&\overset{(*)}{\geq} \int_s^t \liminf_{\delta\searrow 0}\frac{1}{\delta}\big[d_r^2(x_r,y_r)-d_{r,r-\delta}^2(x_r,y_{r-\delta})\big]\,dr\\
&\quad+\int_s^t \liminf_{\delta\searrow 0}\frac{1}{\delta}\big[d_{r+\delta,r}^2(x_{r+\delta},y_r)-d_r^2(x_r,y_r)\big]\,dr\\
&\geq 2\int_s^t\Big[\frac{K}{2}d_r^2(x_r,y_r)+V_r(y_r)-V_r(x_r)+\frac{1}{N}\int_0^1\Big(\partial_aV_r(\gamma_r^a)\Big)^2 a\,da\Big]dr\\
&\quad+2\int_s^t\Big[\frac{K}{2}d_r^2(x_r,y_r)+V_r(x_r)-V_r(y_r)+\frac{1}{N}\int_0^1\Big(\partial_aV_r(\gamma_r^a)\Big)^2 (1-a)\,da\Big]dr\\
&= 2K\int_s^t d_r^2(x_r,y_r)\,dr+\frac{2}{N}\int_s^t\int_0^1\Big(\partial_aV_r(\gamma_r^a)\Big)^2 da\,dr\\
&\geq 2K\int_s^t d_r^2(x_r,y_r)\,dr+\frac{2}{N}\int_s^t\Big|V_r(x_r)-V_r(y_r)\Big|^2 dr.
\end{aligned}
$$

Dividing by $t-s$ and passing to the limit $t-s\searrow 0$ yields

$$
\partial_t d_t^2(x_t,y_t)\geq 2K d_t^2(x_t,y_t)+\frac{2}{N}\Big|V_t(x_t)-V_t(y_t)\Big|^2
$$

for a.e. t. The claim now follows via 'variation of constants'.

It remains to justify the interchange of $\liminf_{\delta\searrow 0}$ and $\int\ldots dr$ in $(*)$ which requires quite some effort. Recall from Proposition 7.2 that $|\frac{d_{s,t}^2(x,y)}{d_s^2(x,y)}-1|\leq 2L\cdot|t-s|$ for all x,y,s,t with $|t-s|\leq\frac{1}{L}$. Thus we can estimate

$$
\begin{aligned}
&-\frac{1}{\delta}\Big[d_r^2(x_r,y_r)-d_{r,r-\delta}^2(x_r,y_{r-\delta})\Big]\\
&\leq\ -\frac{1}{\delta}\Big[d_r^2(x_r,y_r)-d_{r-\delta}^2(x_r,y_{r-\delta})\Big]+o_1\\
&=\ -\frac{1}{\delta}\int_{r-\delta}^r\partial_s d_s^2(x_r,y_s)\,ds+o_1\\
&\leq\ -\frac{1}{\delta}\int_{r-\delta}^r\partial_t d_{s,t}^2(x_r,y_t)\Big|_{t=s}ds+o_1+o_2\\
&\leq\ \frac{2}{\delta}\int_{r-\delta}^r\Big[V_s(x_r)-V_s(y_s)\Big]ds+o_1+o_2+o_3\\
&\leq\ 2C_0\cdot V_r(x_r)+2C_1+C+o_1+o_2+o_3
\end{aligned}
$$

where for the last inequality we used the growth estimate of $s\mapsto V_s(x)$ and the lower boundedness of V and where we put with $o_1(r,\delta)=2L\,d_r^2(x_r,y_{r-\delta})$, $o_2(r,\delta)=2L\,\frac{1}{\delta}\int_{r-\delta}^r d_r^2(x_r,y_\sigma)\,d\sigma$, $o_3(r)=K\,d_r^2(x_r,y_r)$. Continuity of $r\mapsto d_r$ and of $r\mapsto x_r$ as well as of $r\mapsto y_r$ imply that for any fixed

$z \in X$ the function $r \mapsto d_r^2(x_r, z)$ is bounded as well as $r \mapsto d_r^2(y_{r-\delta}, z)$ for $r \in (s,t)$, uniformly in $\delta \in (0,1)$. Thus $o_1(r,\delta) + o_2(r,\delta) + o_3(r,\delta) \le C'$ which finally justifies the interchange of limit and integral.

Similarly, we can estimate

$$\begin{aligned} & -\frac{1}{\delta}\Big[d^2_{r+\delta,r}(x_{r+\delta}, y_r) - d_r^2(x_r, y_r)\Big] \\ \le\ & -\frac{1}{\delta}\int_r^{r+\delta} \partial_s d_s^2(x_s, y_r)\, ds + o_1' \\ \le\ & 2C_0 \cdot V_r(y_r) + 2C_0 + C + o_1' + o_2' + o_3'. \end{aligned}$$

In both cases, the final expression is integrable w.r.t. $r \in [s,t]$ according to Lemma 7.6 since by assumption $V_t(x_t) < \infty$ as well as $V_t(y_t) < \infty$. □

Corollary 7.8. *Assume that $(x_t)_{t\in(\sigma,\tau)}$ is a dynamic backward EVI(K,N)-gradient flow for V and that $(y_t)_{t\in(\sigma,\tau)}$ is a dynamic backward EVI$^-(K,N)$- or EVI$^+(K,N)$-gradient flow for V emanating in the same point $x_\tau = y_\tau$. Then*

$$x_t = y_t$$

for all $t \le \tau$.

Corollary 7.9. *Assume that for given τ, a dynamic upward EVI(K,∞)-gradient flow terminating in x' exists for each x' in a dense subset $D \subset X$. Then this flow can be extended to a flow terminating in any $x' \in X$ and satisfying*

$$d_s(x_s, y_s) \le e^{-K(t-s)} \cdot d_t(x_t, y_t) \tag{104}$$

for any $s < t \le \tau$.

7.4. **Dynamic Convexity.** Let us recall the notion of dynamic convexity as introduced in [51].

Definition 7.10. *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 $(\gamma^a)_{a\in[0,1]}$ with $\gamma^0, \gamma^1 \in Dom(V_t)$*

$$\partial_a^+ V_t(\gamma_t^{1-}) - \partial_a^- V_t(\gamma_t^{0+}) \ge -\frac{1}{2}\partial_t^- d^2_{t-}(\gamma^0,\gamma^1) + \frac{K}{2} d_t^2(\gamma^0,\gamma^1) + \frac{1}{N}\left|V_t(\gamma^0) - V_t(\gamma^1)\right|^2. \tag{105}$$

Theorem 7.11. *Assume that for each $t \in I$ and each $x' \in Dom(V_t)$ there exists a dynamic backward EVI(K,N)-gradient flow $(x_s)_{s\in(\sigma,t]}$ for V emanating in x' and such that $\lim_{s\nearrow t} V_s(x_s) = V_t(x_t)$. Then V is strongly dynamically (K,N)-convex.*

Remark 7.12. To be more precise, we request the inequality (100) at the point t and the inequality (101) at all times before t.

Proof. Fix $t \in I$ and a d_t-geodesic $(\gamma^a)_{a\in[0,1]}$ with $\gamma^0, \gamma^1 \in Dom(V_t)$. The a priori assumption of κ-convexity implies $\gamma^a \in Dom(V_t)$ for all $a \in [0,1]$. For each a, let $(\gamma_s^a)_{s\le t}$ denote the EVI$_N$-gradient flow for V emanating in $\gamma^a = \gamma_t^a$. Then for all $a \in (0,\frac{1}{2})$

$$\begin{aligned} V_t(\gamma^a) - V_t(\gamma^0) &\le \frac{1}{2}\partial_s^- d^2_{s,t}(\gamma_s^a, \gamma^0)\Big|_{s=t-} \\ &\le \frac{1}{2}\partial_s^- d^2_s(\gamma_s^a, \gamma^0)\Big|_{s=t-} + a^2 L\, d_t^2(\gamma^0,\gamma^1) \end{aligned}$$

(due to the log-Lipschitz continuity of $s \mapsto d_s$) and

$$\begin{aligned} V_t(\gamma^{1-a}) - V_t(\gamma^1) &\le \frac{1}{2}\partial_s^- d^2_{s,t}(\gamma_s^{1-a}, \gamma^1)\Big|_{s=t-} \\ &\le \frac{1}{2}\partial_s^- d^2_s(\gamma_s^{1-a}, \gamma^1)\Big|_{s=t-} + a^2 L\, d_t^2(\gamma^0,\gamma^1). \end{aligned}$$

Moreover, the previous Theorem 7.7 implies

$$\begin{aligned} 0 \;&\leq\; \liminf_{s\nearrow t} \frac{1}{t-s}\Big[\frac12 d_t^2(\gamma^a,\gamma^{1-a})-\frac12 d_s^2(\gamma^a_s,\gamma^{1-a}_s)-K\,d_t^2(\gamma^a,\gamma^{1-a})-\frac1N\int_s^t\Big|V_r(\gamma^a_r)-V_r(\gamma^{1-a}_r)\Big|^2dr\\ &=\; \frac12\partial_s^- d_s^2(\gamma^a_s,\gamma^{1-a}_s)\Big|_{s=t-}-K\,d_t^2(\gamma^a,\gamma^{1-a})-\frac1N\Big|V_t(\gamma^a)-V_t(\gamma^{1-a})\Big|^2. \end{aligned}$$

(Here we used the requested continuity $V_r(\gamma^a_r)\to V_t(\gamma^a)$ for $r\nearrow t$.)

Adding up these inequalities (the last one multiplied by $\frac{1}{1-2a}$ and the previous ones by $\frac1a$) yields

$$\begin{aligned} &\frac1a\Big[V_t(\gamma^a)-V_t(\gamma^0)+V_t(\gamma^{1-a})-V_t(\gamma^1)\Big]\\ &\quad\leq\; \liminf_{s\nearrow t}\frac{1}{2(t-s)}\Big(\big[\frac1a d_t^2(\gamma^0,\gamma^a)+\frac{1}{1-2a}d_t^2(\gamma^a,\gamma^{1-a})+\frac1a d_t^2(\gamma^{1-a},\gamma^1)\big]\\ &\qquad\qquad -\big[\frac1a d_s^2(\gamma^0,\gamma^a_s)+\frac{1}{1-2a}d_s^2(\gamma^a_s,\gamma^{1-a}_s)+\frac1a d_s^2(\gamma^{1-a}_s,\gamma^1)\big]\Big)\\ &\qquad\quad +2aL\,d_t^2(\gamma^0,\gamma^1)-\frac{K}{1-2a}d_t^2(\gamma^a,\gamma^{1-a})-\frac{1}{N(1-2a)}\Big|V_t(\gamma^a)-V_t(\gamma^{1-a})\Big|^2\\ &\quad\leq\; \liminf_{s\nearrow t}\frac{1}{2(t-s)}\Big(d_t^2(\gamma^0,\gamma^1)-d_s^2(\gamma^0,\gamma^1)\Big)\\ &\qquad\quad -\big[(1-2a)K-2aL\big]\cdot d_t^2(\gamma^0,\gamma^1)-\frac{1}{N(1-2a)}\Big|V_t(\gamma^a)-V_t(\gamma^{1-a})\Big|^2. \end{aligned}$$

In the limit $a\to0$ this yields the claim. □

References

[1] Luigi Ambrosio, Matthias Erbar, and Giuseppe Savaré. Optimal transport, Cheeger energies and contractivity of dynamic transport distances in extended spaces. *Nonlinear Analysis: Theory, Methods & Applications*, 137:77–134, 2016.

[2] Luigi Ambrosio, Nicola Gigli, Andrea Mondino, and Tapio Rajala. Riemannian Ricci curvature lower bounds in metric measure spaces with σ-finite measure. *Trans. Amer. Math. Soc.*, 367(7):4661–4701, 2015.

[3] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the Space of Probabiliy Measures*. Birkhäuser, Basel, 2005.

[4] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below. *Invent. Math.*, 195(2):289–391, 2013.

[5] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. Metric measure spaces with Riemannian Ricci curvature bounded from below. *Duke Math. J.*, 163(7):1405–1490, 2014.

[6] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. Bakry-Émery curvature-dimension condition and Riemannian Ricci curvature bounds. *Ann. Probab.*, 43(1):339–404, 2015.

[7] Luigi Ambrosio, Andrea Mondino, and Giuseppe Savaré. Nonlinear diffusion equations and curvature conditions in metric spaces. *arXiv:1509.07273*, 2015.

[8] Marc Arnaudon, Koléhè Abdoulaye Coulibaly, and Anton Thalmaier. Brownian motion with respect to a metric depending on time; definition, existence and applications to Ricci flow. *Comptes Rendus Mathematique*, 346(13):773–778, 2008.

[9] Marc Arnaudon, Koléhè Abdoulaye Coulibaly, and Anton Thalmaier. Horizontal diffusion in C^1 path space. In *Séminaire de Probabilités XLIII*, volume 2006 of *Lecture Notes in Math.*, pages 73–94. Springer, Berlin, 2011.

[10] Kathrin Bacher and Karl-Theodor Sturm. Ricci bounds for Euclidean and spherical cones. In *Singular phenomena and scaling in mathematical models*, pages 3–23. Springer, Cham, 2014.

[11] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.

[12] Francois Bolley, Ivan Gentil, Arnaud Guillin, and Kazumasa Kuwada. Equivalence between dimensional contractions in Wasserstein distance and the curvature-dimension condition. *arXiv:1510.07793*, 2015.

[13] Huai-Dong Cao and Xi-Ping Zhu. A complete proof of the Poincaré and geometrization conjectures—application of the Hamilton-Perelman theory of the Ricci flow. *Asian J. Math.*, 10(2):165–492, 2006.

[14] Jeff Cheeger. Differentiability of Lipschitz functions on metric measure spaces. *Geom. Funct. Anal.*, 9(3):428–517, 1999.

[15] Zhen-Qing Chen and Masatoshi Fukushima. *Symmetric Markov processes, time change, and boundary theory*, volume 35 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2012.

[16] Koléhè Abdoulaye Coulibaly. Brownian motion with respect to time-changing Riemannian metrics, applications to Ricci flow. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(2):515–538, 2011.

[17] Matthias Erbar, Kazumasa Kuwada, and Karl-Theodor Sturm. On the equivalence of the entropic curvature-dimension condition and Bochner's inequality on metric measure spaces. *Invent. Math.*, 201:993–1071, 2015.

[18] Matthias Erbar and Karl-Theodor Sturm. Rigidity for Ricci flat metric cones. *In progress*, 2017.

[19] Tryphon Georgiou, Chris Sander, Allen Tannenbaum, and Liangjia Zhu. Ricci curvature and robustness of cancer networks. *arXiv:1502.04512*, 2015.

[20] Nicola Gigli. Nonsmooth differential geometry-An approach tailored for spaces with Ricci curvature bounded from below. *arXiv:1407.0809*, 2014.

[21] Piotr Hajłasz. Sobolev spaces on an arbitrary metric space. *Potential Analysis*, 5(4):403–415, 1996.

[22] Richard S. Hamilton. Three-manifolds with positive Ricci curvature. *J. Differential Geom.*, 17(2):255–306, 1982.

[23] Richard S. Hamilton. The formation of singularities in the Ricci flow. In *Surveys in differential geometry, Vol. II (Cambridge, MA, 1993)*, pages 7–136. Int. Press, Cambridge, MA, 1995.

[24] Robert Haslhofer and Aaron Naber. Weak solutions for the Ricci flow I. *arXiv:1504.00911*, 2015.

[25] Christian Ketterer. Cones over metric measure spaces and the maximal diameter theorem. *J. Math. Pures Appl. (9)*, 103(5):1228–1275, 2015.

[26] Bruce Kleiner and John Lott. Notes on Perelmans papers. *Geometry & Topology*, 12(5):2587–2855, 2008.

[27] Bruce Kleiner and John Lott. Singular Ricci flows I. *arXiv:1408.2271*, 2014.

[28] Eva Kopfer. Gradient flow for the Boltzmann entropy and Cheeger's energy on time-dependent metric measure spaces. *arXiv:1611.09522*, 2016.

[29] Eva Kopfer. Super-Ricci flows and improved gradient and transport estimates. *arXiv:1704.04177*, 2017.

[30] Eva Kopfer and Karl-Theodor Sturm. Super-Ricci flows and functional inequalities. *In progress*, 2017.

[31] Kazumasa Kuwada. Space-time Wasserstein controls and Bakry–Ledoux type gradient estimates. *Calculus of Variations and Partial Differential Equations*, 54(1):127–161, 2015.

[32] Kazumasa Kuwada and Robert Philipowski. Coupling of Brownian motions and Perelman's L-functional. *Journal of Functional Analysis*, 260(9):2742–2766, 2011.

[33] Kazumasa Kuwada and Robert Philipowski. Non-explosion of diffusion processes on manifolds with time-dependent metric. *Mathematische Zeitschrift*, 268(3-4):979–991, 2011.

[34] Sajjad Lakzian and Michael Munn. Super Ricci flow for disjoint unions. *arXiv:1211.2792*, 2012.

[35] Songzi Li and Xiang-Dong Li. The W-entropy formula for the Witten Laplacian on manifolds with time dependent metrics and potentials. *Pacific Journal of Mathematics*, 278(1):173–199, 2015.

[36] Janna Lierl and Laurent Saloff-Coste. Parabolic Harnack inequality for time-dependent non-symmetric Dirichlet forms. *arXiv:1205.6493*, 2012.

[37] Jacques Louis Lions and Enrico Magenes. *Non-homogeneous boundary value problems and applications. Vol. I.* Springer-Verlag, New York-Heidelberg, 1972. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181.

[38] John Lott and Cédric Villani. Ricci curvature for metric-measure spaces via optimal transport. *Annals of Mathematics*, 169(2):903–991, 2009.

[39] Robert J McCann and Peter M Topping. Ricci flow, entropy and optimal transportation. *American Journal of Mathematics*, 132(3):711–730, 2010.

[40] John Morgan and Gang Tian. *Ricci flow and the Poincaré conjecture*, volume 3 of *Clay Mathematics Monographs*. American Mathematical Society, Providence, RI; Clay Mathematics Institute, Cambridge, MA, 2007.

[41] Yann Ollivier. Ricci curvature of Markov chains on metric spaces. *J. Funct. Anal.*, 256(3):810–864, 2009.

[42] Grisha Perelman. The entropy formula for the Ricci flow and its geometric applications. *arXiv:math/0211159*, 2002.

[43] Grisha Perelman. Finite extinction time for the solutions to the Ricci flow on certain three-manifolds. *arXiv:math/0307245*, 2003.

[44] Grisha Perelman. Ricci flow with surgery on three-manifolds. *arXiv:math/0303109*, 2003.

[45] Michael Renardy and Robert C Rogers. *An introduction to partial differential equations*, volume 13. Springer-Verlag, New York, 2004.

[46] Romeil Sandhu, Tryphon Georgiou, and Allen Tannenbaum. Market fragility, systemic risk, and Ricci curvature. *arXiv:1505.05182*, 2015.

[47] Nageswari Shanmugalingam. Newtonian spaces: an extension of Sobolev spaces to metric measure spaces. *Rev. Mat. Iberoamericana*, 16(2):243–279, 2000.

[48] Karl-Theodor Sturm. Analysis on local Dirichlet spaces. II. Upper Gaussian estimates for the fundamental solutions of parabolic equations. *Osaka J. Math.*, 32(2):275–312, 1995.

[49] Karl-Theodor Sturm. Diffusion processes and heat kernels on metric spaces. *The Annals of Probability*, 26(1):1–55, 1998.

[50] Karl-Theodor Sturm. On the geometry of metric measure spaces. I and II. *Acta Math.*, 169(1):65–131, 2006.

[51] Karl-Theodor Sturm. Super Ricci flows for metric measure spaces. I. *arXiv:1603.02193*, 2016.

[52] Karl-Theodor Sturm. Synthetic upper Ricci bounds for metric measure spaces. *In progress*, 2017.

[53] Peter Topping. Ricci flow: The foundations via optimal transportation. 2013.